

NEUROSCIENCES ET ÉTHIQUE : QUE NOUS APPREND LE DILEMME DU WAGON FOU ?

Bernard Baertschi
Université de Genève
Institut d'éthique biomédicale
Bernard.Baertschi@unige.ch

RÉSUMÉ

L'éthique normative contemporaine est marquée par un débat entre deux positions antagonistes, le déontologisme et l'utilitarisme. Ce débat est resté jusqu'à récemment un débat usant d'arguments, d'objections, de contre-exemples et d'expériences de pensée essentiellement conceptuels. Il y a peu cependant, dans la foulée des progrès de l'imagerie cérébrale, Joshua Greene y a introduit des éléments empiriques : proposant à des sujets différents dilemmes, dont le dilemme du wagon fou (trolley problem), il a examiné comment leur cerveau réagissait. Il en a tiré la conclusion que le déontologisme s'appuyait sur des réponses émotionnelles, alors que l'utilitarisme utilisait des circuits rationnels, ce qui assurait à ce dernier une meilleure pertinence morale. Les raisons déontologistes seraient même des rationalisations a posteriori, voire des confabulations. Malheureusement, si le recours à des données empiriques est un progrès philosophique dans l'examen de ces questions, les conséquences que Greene en tire ne sont à mon sens pas probantes.

ABSTRACT

Contemporary normative ethics is characterized by a debate between two conflicting theories, deontologism and utilitarianism. Until recently, this debate has essentially used arguments, objections and thought-experiments of a conceptual nature. Following the progress in neuroimaging, Joshua Greene has introduced empirical elements in the discussion : submitting several subjects to different dilemmas, in particular the trolley problem, he has examined how their brain were reacting. He concluded that deontologists were drawing on their emotions, whereas utilitarians reacted more rationally, and therefore gave a better answer from a moral point of view. Deontologists reasons are even rationalizations and confabulations, rather than authentic moral reasons. Unfortunately, if Greene is on the right track in introducing empirical data in this debate, his conclusions are not warranted in my opinion.

MOTS-CLÉS

Déontologisme, Utilitarisme, Neuroéthique, Dilemme du wagon fou, Principe des actions à double effet

1 INTRODUCTION

L'éthique normative a pour tâche de répondre à des questions comme « Qu'est-ce qu'une action bonne ? » ou « Comment doit-on se comporter dans telle situation ? ». Les philosophes contemporains ont tendance à classer les réponses sous deux catégories opposées, exemplaires de deux théories morales en désaccord : soit on donne à ces questions une réponse *déontologiste* soit une réponse *conséquentialiste*. Une réponse déontologiste invoque une norme à laquelle l'action ou le comportement doivent se conformer, norme qui s'énonce généralement comme une obligation ou une interdiction catégoriques. Par exemple, chez Kant, la norme de non-instrumentalisation de la personne humaine : on doit se comporter en toute situation en traitant les personnes concernées comme des fins en soi et non comme de purs moyens. Une réponse conséquentialiste invoque une valeur qu'il est bon ou obligatoire de promouvoir. Par exemple, l'utilitarisme nous ordonne de maximiser le plaisir ou la satisfaction des préférences de toutes les personnes impliquées par nos actions, ce qui se mesure dans leurs conséquences.

Cela est bien connu, tout comme il est bien connu que ces deux théories se heurtent à de multiples difficultés¹. Les tenants des deux positions s'adressent de nombreuses objections et imaginent force contre-exemples, parfois très sophistiqués, pour se convaincre mutuellement d'erreur. Jusqu'à très récemment, le débat est resté conceptuel et normatif; mais depuis peu, quelque chose de nouveau s'est produit : les neurosciences, et plus particulièrement l'imagerie cérébrale, ont été convoquées dans le débat. L'un de ceux qui l'on fait, Joshua Greene, estime même que les neurosciences nous permettent maintenant de trancher, car les données empiriques obtenues montrent indubitablement que le conséquentialisme est vainqueur.

Dans cet article, je vais examiner l'argumentation déployée par Greene pour justifier sa thèse, ainsi que certaines objections qui lui ont été adressées, particulièrement celles de Marc Hauser, qui recourent elles aussi à des données empiriques. L'enjeu est de taille : si l'on pouvait donner raison à Greene, alors il faudrait admettre que l'éthique – ou du moins certaines questions morales – est devenue une discipline expérimentale grâce aux neurosciences, et que les données empiriques sont déjà suffisantes pour établir que Kant et les déontologistes sont dans l'erreur, en ce sens que leur théorie n'est pas une éthique normative adéquate.

Greene s'appuie sur ce qu'il considère comme une expérience de pensée cruciale, le *dilemme du wagon fou* (en anglais : *trolley problem*). C'est par conséquent autour ce dilemme que je vais disposer les arguments.

1. Je m'en suis occupé dans Baertschi, 1995, chap. 1-2.

2 LE DILEMME DU WAGON FOU ET SES LEÇONS SELON JOSHUA GREENE

Ce dilemme a été imaginé en 1967 par Philippa Foot dans un article célèbre à propos du problème de l'avortement (reproduit dans 1978). Depuis, il a connu une très grande fortune et a été utilisé pour tester nombre de nos intuitions morales. Comme je l'ai dit précédemment, je vais me limiter à son usage comme expérience cruciale pour départager le déontologisme et le conséquentialisme. Voici la situation que le dilemme met en jeu – je m'appuie sur l'exposé qu'en propose Hauser et ses collègues (2007). Un wagon dévale une pente à vive allure car ses freins sont hors-d'usage. Plus bas sur la voie travaillent cinq ouvriers qui vont être écrasés, d'autant qu'il n'y a aucun moyen de les prévenir. Toutefois, un aiguillage permettrait de faire aller le wagon sur une autre voie où un seul ouvrier travaille. Quelqu'un, qu'on appellera Denise, a la possibilité d'actionner cet aiguillage. A-t-elle le droit de le faire ?² 85% des personnes interrogées affirment qu'elle en a le droit. Cela ne surprend pas vraiment : ne vaut-il pas mieux qu'une seule personne meure plutôt que cinq ?³

On change maintenant le scénario (*le gros homme* ou *la passerelle*). Le wagon dévale la pente, mais il n'y a aucun aiguillage. Toutefois, un individu nommé Frank se trouve sur une passerelle qui enjambe la voie, à côté d'un homme suffisamment gros pour que, si Frank le pousse et qu'il tombe sur la voie, son corps arrête le wagon et l'empêche de poursuivre sa route meurtrière. A-t-il le droit de le faire ? Cette fois, 88% des mêmes personnes interrogées nient qu'il en ait le droit, alors même qu'ainsi une seule personne mourrait plutôt que cinq, tout comme dans le cas précédent.

Cette histoire est déroutante, car les mêmes personnes répondent différemment alors que la situation paraît moralement analogue et que le résultat est identique : un mort au lieu de cinq. Comment expliquer cette différence ? On peut risquer quelques hypothèses psychologiques :

- 1° L'être humain est irrationnel et se laisse détourner de son devoir par des éléments non pertinents moralement.
- 2° Pour des raisons évolutionnaires, l'être humain répugne à être la cause directe de la mort de quelqu'un (Levy, 2007, 291-292).
- 3° Loin des yeux, loin du cœur : l'ouvrier travaillant sur l'autre branche de l'aiguillage est moins visible que le gros homme.

2. Hauser et ses collègues disent (je souligne) : « Is it *morally permissible* for Denise to turn the train? », ce que je traduis par « A-t-elle le droit ».

3. Ces résultats ont été obtenus par sondage sur internet, ce qui a permis de toucher énormément de personnes, de toute origine, même s'il y a un léger biais, vu que seules les personnes ayant accès à internet ont pu voter. Aucune différence significative dans les résultats n'a été observée en fonction des groupes ethniques et religieux, des hommes et des femmes, ainsi que de l'âge (Hauser & al., 2007, 8-13).

On peut encore se laisser aller à d'autres supposition, et on peut aussi poser la question à ceux qui ont répondu différemment dans les deux cas si l'on estime que l'être humain a une vue suffisamment claire de ses motivations. Mais on peut aussi placer les sujets dans un scanner (IRMf) pour voir ce qui se passe dans leur cerveau lorsqu'on leur présente le dilemme – ce qui a été fait. On a constaté que ce ne sont pas les mêmes zones du cerveau qui sont activées dans les deux cas. Plus précisément, les expériences menées sous la direction de Greene montrent que plus la question morale qui se pose à nous est personnelle, c'est-à-dire plus nous sommes impliqués activement dans ce qu'il faut faire, et plus nous pouvons le faire directement et simplement en étendant le bras, plus les zones émotionnelles de notre cerveau sont actives ; or dans le *gros homme*, nous sommes bien plus impliqués personnellement et plus proches de l'action (Gazzaniga, 2005, 170-171). Bref, nos décisions morales sont influencées par la place que nous avons dans l'action, et plus nous sommes concernés, plus nos émotions sont vives, inhibant l'action lorsqu'elle implique un dommage sérieux fait à autrui.

Ces résultats sont corroborés par une étude menée par Nadelhoffer et Feltz (2008), qui ont présenté le dilemme à des sujets et leur ont demandé ce qu'il était permis à l'agent de faire, l'agent étant soit eux-mêmes, soit quelqu'un d'autre. Les résultats obtenus dans les deux cas sont bien différents : en position d'observateur – position qui est celle de l'exposé du dilemme que proposent Hauser et ses collègues –, 90% des sujets disent qu'il serait permis à l'acteur d'actionner l'aiguillage, alors qu'en position d'acteur, il ne sont plus que 65% à le penser⁴. On voit clairement en quoi ce résultat corrobore l'interprétation de Greene : l'implication du sujet dans l'action, source d'émotions, change sa perception du caractère permis de l'acte. Manifestement les émotions influencent et parfois inhibent.

L'explication de Greene peut être comprise comme étant du type des « excuses » : le résultat des deux dilemmes étant le même, nous devrions agir de la même manière ; toutefois des considérations émotionnelles nous empêchent de le faire. Elles nous détournent donc de notre devoir. Néanmoins, comme nous sommes ainsi câblés, il apparaît difficile de nous demander d'agir autrement ; notre constitution émotionnelle est en conséquence une excuse pour notre comportement irrationnel, et donc en définitive immoral. C'est ce que note encore Greene à l'occasion de l'étude d'autres dilemmes : « Peut-être existe-t-il quelque "bonne raison" qui explique pourquoi il n'est pas problématique de dépenser son argent pour des sushi ou des *power windows*, alors que des millions de personnes meurent de faim ou de maladie curable pourraient être sauvées. Peut-être ces intuitions morales n'ont-elles rien à voir avec de "bonnes raisons", mais concernent la manière dont notre cerveau est construit. » (2003, 224)

Toutefois, Greene ne pense pas que nous soyons définitivement les es-

4. Sur l'importance de la position d'observateur ou d'acteur, voir Tassy, 2011.

claves de notre câblage et estime qu'il nous est possible de surmonter cette irrationalité vers plus de rationalité et donc plus de moralité. En le soutenant, il élève la portée de l'étude du dilemme au statut d'expérience cruciale permettant de trancher entre la pertinence morale du déontologisme et celle de l'utilitarisme : « Nous faisons l'hypothèse que la pensée de causer la mort de quelqu'un en le poussant personnellement (comme dans le dilemme du *gros homme*) est plus forte émotionnellement que la pensée de causer les mêmes conséquences de manière plus impersonnelle (par exemple en manipulant un aiguillage, comme dans le dilemme du *wagon fou*) [...]. C'est-à-dire que les personnes tendent vers le conséquentialisme dans les cas où la réponse émotionnelle est faible, et tendent vers le déontologisme dans les cas où la réponse émotionnelle est forte » (2008, 43). Nous sommes tous tantôt déontologistes, tantôt conséquentialistes, tantôt des Frank, tantôt des Denise. Autrement dit, tantôt nous plaçons notre comportement sous la férule de commandement et d'interdits absolus, lorsque notre réponse émotionnelle est forte, tantôt sous l'égide de l'évaluation des conséquences de nos actes, lorsque notre réponse émotionnelle est faible – ce disant et de manière tout à fait paradoxale Greene range le déontologisme, et donc Kant, du côté de l'irrationalisme. Mais il n'y a là rien d'immuable. En effet, il nous est loisible de devenir plus utilitaristes ou, plus généralement, conséquentialistes, c'est-à-dire de faire preuve de plus de rationalité ; c'est-là une attitude que nous pouvons décider d'adopter, sinon en faisant taire nos émotions, du moins en ne les écoutant pas, bref en ne leur conférant pas de poids moral. C'est que les réactions émotionnelles sont évolutionnairement archaïques : « La tension entre les perspectives utilitaristes et déontologistes en philosophie morale reflète une tension plus fondamentale qui vient de la structure du cerveau humain. Les réponses socio-émotionnelles que nous avons héritées de nos ancêtres primates (dues sans doute à quelque avantage d'adaptation qu'elles conféraient), structurées et affinées par la culture, sous-tendent les interdictions absolues qui sont centrales au déontologisme. Par contraste, le "calcul moral" qui définit l'utilitarisme est rendu possible par des structures apparues plus récemment dans les lobes frontaux, soutenant la pensée abstraite et le contrôle cognitif de haut niveau » (Greene & al., 2004, 398).

Il existe donc deux types de théories morales – que Greene qualifie un peu bizarrement d'« espèces naturelles »⁵ –, le conséquentialisme et le déontologisme, où dominent respectivement la raison et les émotions, dont la pertinence morale n'est pas du tout la même. Seule celle qui s'appuie sur la raison est authentiquement morale. Est-ce là un nouveau chapitre de l'antique opposition entre notre nature rationnelle et notre nature animale, accompagnée

5. L'affirmation que conséquentialisme et déontologisme sont deux espèces naturelles est douteuse : d'une part ces deux doctrines sont d'invention récente et d'autre part il est très difficile d'en trouver des exemples purs. Kant, par exemple, s'il conteste que les conséquences comptent, affirme tout de même que certains biens (buts) sont sources de devoirs, notamment sa propre perfection et le bonheur d'autrui, cf. 1985, 56.

d'une exhortation à ne s'appuyer que sur la première si nous voulons décider moralement correctement? On en a parfois l'impression, mais cela demande à être nuancé, d'une part à cause du lien que Greene établit entre le rôle des émotions et le kantisme – qui, on le sait, les dénonce à la manière du stoïcisme (Kant, 1980, 186) –, et d'autre part parce qu'il ne nie pas que toute décision morale soit empreinte d'émotions : « Je n'affirme pas que le jugement conséquentialiste est dépourvu d'émotion. Au contraire, j'incline à être d'accord avec Hume que tout jugement moral doit posséder quelque composante affective, et je suspecte que la pesée conséquentialiste des dommages et des bénéfiques est un processus émotionnel » (2008, 64). Toutefois, Greene nie que ce processus soit de même nature que la démarche déontologiste, car « il s'agit d'un processus de pesée et non d'un processus d'"alarme". » Bref, il y a des émotions moralement appropriées, celles qui sont accompagnées par un processus rationnel de pesée, et des émotions moralement inappropriées, séparées d'un tel processus et qui jouent le rôle de simple « alarme ». La précision est intéressante, mais on ne peut pas dire qu'elle soit vraiment éclairante, d'une part parce qu'il s'agit seulement d'une indication et d'autre part parce qu'elle n'est fondée sur aucune étude empirique (Berker, 2009, 307-308). On a surtout l'impression que Greene cherche à la fois à se placer sous la bannière de Hume et à tenir compte des études qui, depuis Damasio, soulignent le rôle important que jouent et doivent jouer nos émotions si nous voulons pouvoir décider de manière appropriée (Baertschi, 2009, chap. 1).

Je résumerai la position de Greene ainsi :

1. Les émotions nous motivent (sans émotions, pas d'action).
2. Une motivation moralement adéquate est impossible sans émotions.
3. Toutefois, certaines émotions nous motivent de façon inadéquate sur le plan moral.
4. Seules les émotions qui sont en accord avec une norme morale rationnelle nous motivent de manière moralement adéquate.
5. La seule norme rationnellement adéquate pour la morale est la norme conséquentialiste (maximisation du bien dans les résultats).
6. Les émotions qui nous motivent dans le cas du *gros homme*, bien qu'elles soient explicables d'un point de vue évolutionnaire, sont donc moralement inadéquates.

L'explication de Greene est de nature *psychologique* : elle rend compte de notre comportement. Il en tire des considérations normatives qui impliquent que certaines causes psychologiques (les émotions-alarmes) nous font agir autrement que nous le devrions et donc qu'il n'existe pas de véritable *raison morale* qui puisse expliquer la différence de réaction observée dans le cas du *wagon fou* et dans celui de *gros homme*, raison qui montrerait qu'il est justifié de nous comporter différemment dans ces deux situations. Mais est-ce bien le cas? Deux considérations obligent à prendre cette question au sérieux :

- 1° Les personnes souffrant d'une lésion du lobe frontal ventromédian (VM-PC) – lésion qui cause notamment une inaptitude à prendre des déci-

sions morales adéquates – disent qu’elles n’hésiteraient pas à pousser le gros homme (Koenigs & al., 2007 ; Roskies 2006, 21). Par ailleurs, ces personnes décident de manière incorrecte dans différents tests, ce qui est provoqué par le fait que leurs émotions n’influencent pas normalement leurs décisions. Dans leurs cas, on parle de *sociopathie* acquise, terme qui souligne bien l’inadéquation morale de leur comportement. En outre, dans une étude récente, Liane Young et ses collègues (2010) ont montré que ces personnes ont un problème dans la perception et dans l’évaluation des intentions morales : lorsqu’on leur demande de comparer deux situations, l’une qui débouche sur un malheur accidentel (une personne est empoisonnée suite à une confusion entre deux produits), et une autre qui, par bonheur, ne débouche sur aucun malheur malgré une intention malveillante (une personne veut empoisonner une autre mais se trompe de produit), les patients VMPC considèrent que l’action de l’agent malveillant pose moins de problème que celle de l’agent qui s’est trompé, contrairement aux participants à l’étude qui ne souffrent pas de cette lésion cérébrale. C’est comme si, pour ces patients, seul le résultat comptait – reproche que l’on entend parfois en ce qui concerne justement les utilitaristes.

- 2° L’interprétation de Greene n’est pas normativement neutre, comme on le voit bien dans le résumé de sa position que je viens de proposer : elle s’appuie sur une conception utilitariste ou conséquentialiste de l’action morale et du rôle de la raison, ce qui a pour autre effet que les distinctions qu’il trace, entre raison et émotion d’une part, et entre émotions adéquates et inadéquates de l’autre, pourraient, elles aussi, être imprégnées de théorie ou biaisées⁶.

3 LES VARIANTES IMAGINÉES PAR MARC HAUSER : UN TEST DÉCISIF ?

Traditionnellement, lorsqu’un philosophe envisage d’examiner une thèse comme celle que Greene propose, il a recours à une argumentation conceptuelle aidée par des expériences de pensée et la construction de contre-exemples. Et c’est bien ainsi qu’il lui convient de procéder. Toutefois, de même que Greene a testé empiriquement une expérience de pensée, le dilemme du *wagon fou*, je vais ici aussi recourir à un test expérimental pour évaluer la thèse de Greene, test que Marc Hauser et ses collègues ont réalisé.

6. Très récemment, les résultats de Greene ont été contestés : ce qu’il aurait testé ne serait pas l’opposition entre les réactions utilitaristes et les réactions déontologistes, mais l’opposition entre les réactions à des dilemmes dont la solution est intuitive d’une part, contre-intuitive de l’autre (G. Kahane & al., 2011). Je laisserai cette question de côté, tout en soulignant que c’est à mon sens une des forces de l’approche expérimentale en philosophie que de pouvoir être contestée au niveau expérimental lui-même. De manière analogue, Popper ne disait-il pas que la réfutabilité est ce qui fait la force de la méthode scientifique ?

Ces derniers ont mené une enquête pour tenter de voir si, contrairement à l'affirmation de Greene, nos décisions divergentes dans le cas du *wagon fou* (Denise) et dans celui du *gros homme* (Frank) n'avaient pas un fondement authentiquement normatif – en fait, il s'agit de la même enquête qui a donné les résultats que l'on a vu pour Denise et Frank (Hauser & al., 2007)⁷. À cet effet, ils ont proposé deux nouvelles variantes du dilemme, les cas de *Ned* et d'*Oscar*. Tout comme Denise, Ned se trouve à proximité d'un aiguillage qui, actionné, tuera une personne au lieu de cinq. Toutefois, il est heureux que cette personne soit un gros homme, car sinon le wagon ne serait pas stoppé et reviendrait sur la voie principale pour tuer les cinq ouvriers, vu que la voie secondaire rejoint la voie principale un peu plus loin. Ned a-t-il le droit d'actionner l'aiguillage ? 56% des sondés répondent par l'affirmative. Soulignons en passant qu'ici aussi, la question est posée du point de vue du spectateur et non de l'acteur et que, si elle est posée sur le plan normatif, elle ne l'est pas dans les termes d'une obligation ; ce serait donc une seconde question de savoir si actionner l'aiguillage serait moralement obligatoire. Mais il est tout à fait justifié de poser la question ainsi, car ce qu'il s'agit de savoir, c'est si les actions de Denise, de Frank et de Ned (et bientôt d'Oscar) sont moralement *justifiables*, par conséquent s'ils ont d'authentiques ou de bonnes raisons morales d'agir comme ils le font ; ce que Greene nie dans certains cas, on l'a vu.

Dans le cas de Ned, 56% des sondés estiment justifié d'actionner l'aiguillage. Qu'en est-il dans le cas d'Oscar (le second scénario nouveau introduit) ? Oscar se trouve dans la même configuration que Ned, mais avec une petite différence : l'homme que le train va écraser s'il actionne l'aiguillage n'a pas besoin d'être gros, car il y a un objet pesant posé sur la voie juste derrière lui qui arrêtera le train – on peut penser à un rocher. Cette fois, 72% des sondés estiment qu'il est moralement permis à Oscar d'actionner l'aiguillage.

Classons maintenant les quatre scénarios en fonction du degré d'approbation que les sondés ont attribué à l'action. On obtient :

Denise (85%) → Oscar (72%) → Ned (56%) → Frank (12%).

Les résultats ne sont pas très tranchés et marquent une progression. Pourquoi ? Hauser et ses collègues n'ont pas systématiquement demandé aux sondés quelles étaient les raisons qui ont motivé leurs choix, car ils considèrent que, souvent, nous n'avons pas d'accès conscient aux principes moraux qui nous guident, que nous réagissons fréquemment intuitivement, quitte à rationaliser après coup notre jugement, ce qui est particulièrement

7. Le propos de l'article de Hauser & al. est de montrer que ce n'est pas parce qu'une réponse est rapide et intuitive, qu'elle n'exprime pas un principe moral. C'est ce que je veux dire lorsque je parle de « fondement authentiquement normatif ». Ainsi, une personne n'a pas besoin de pouvoir dire quel principe elle suit et pourquoi elle le suit pour que son action ait un fondement authentiquement moral ; il suffit qu'elle suive un principe proprement moral. Cela peut paraître trop peu exigeant, mais suffit à mon sens pour la question qui m'occupe ici (merci à un des relecteurs pour avoir attiré mon attention sur ce point).

vrai dans les dilemmes où les émotions jouent un rôle. Dans de tels cas, il existe « une dissociation entre le jugement et la justification » (2007, 2). On l'a vu et on le verra encore, c'est aussi exactement ce que pense Joshua Greene. Cette opinion est d'ailleurs répandue chez les neuropsychologues, et certains comme Jonathan Haidt (2001) vont même jusqu'à en tirer des conclusions sceptiques par rapport à notre discernement moral. Nous n'avons toutefois pas besoin d'examiner cette question, car l'introduction des cas de Ned et d'Oscar, ainsi que la comparaison des quatre versions du dilemme vont nous permettre de tirer des conclusions pertinentes indépendamment de la question de la validité des justifications morales que les sujets pourraient donner⁸.

Revenons à ce classement. Que peut-on en tirer? On remarque en premier lieu que les cas les plus problématiques sont constitués par ceux où l'homme sacrifié est gros, c'est-à-dire où il est lui-même un *instrument* pour sauver d'autres personnes. Or cela contrevient directement au principe kantien de l'interdiction d'instrumentaliser une personne. Si l'on en restait là, on se rendrait déjà compte qu'il existe une raison morale de ne pas pousser le gros homme, sauf bien sûr pour un utilitariste strict. Il est donc faux de n'y voir qu'une irrationalité, qu'une excuse : il est possible de pointer une différence morale pertinente entre la situation de Denise et celle de Frank, du moins pour ceux, fort nombreux dans un contexte libéral comme le nôtre, qui admettent le principe de non instrumentalisation. En outre, ce n'est pas la seule différence morale discernable. En effet, être un instrument, c'est être un moyen ; or il existe des raisons morales liés au choix des moyens pour atteindre les fins qu'on se propose. C'est du moins ce que pensent tous ceux qui estiment que la fin ne justifie pas les moyens et qui estiment que la structure causale de nos actions joue un rôle dans leur évaluation.

Ainsi, Greene a tort lorsqu'il estime qu'un Frank ne pourrait alléguer aucune raison proprement morale pour son refus. Mais ce n'est pas tout : la comparaison des deux nouveaux cas (Ned et Oscar) montre en second lieu que Greene n'a pas correctement saisi la nature de la différence qui explique les réactions opposées que les personnes sondées ont par rapport aux cas de Denise et de Frank. C'est ce que soulignent Hauser et ses collègues : « Si, ainsi que Greene et ses collègues le suggèrent, nos jugements moraux étaient guidés par la distinction personnel-impersonnel, médiatisée par des processus émotionnels, alors tous les sujets auraient jugé les scénarios 3 et 4 comme semblables. [...] La conclusion que nous tirons de ces résultats est que la dimension personnelle-impersonnelle, quoique pouvant être importante pour expliquer certains jugements moraux, ne dit pas tout. La pièce manquante de ce puzzle théorique est la partie de notre psychologie qui évalue les causes et les conséquences de l'action, spécialement sa structure intentionnelle » (2007,

8. Notons encore que dans l'expérience menée par Hauser et ses collègues, le fait que les sondés aient eu ou non des connaissances en philosophie morale ne changeait rien à leur difficulté à justifier leurs jugements et leurs intuitions (2007, 16).

17). Ned et Oscar sont exactement dans la même situation de proximité personnelle par rapport à l'événement, pourtant leur action n'est pas évaluée de la même manière. Pourquoi ? Parce que les propriétés causales et intentionnelles de l'action comptent aussi, elles sont moralement pertinentes, ce que Greene n'a pas vu.

Ce dernier ne s'est toutefois pas rendu à ces arguments. Selon lui, le fait que les sondés, représentant la morale commune, jugent d'une certaine manière et qu'on peut trouver des raisons morales à leurs jugements n'entraîne ni qu'ils ont raison de juger comme ils le font, ni même que leurs raisons sont d'authentiques raisons morales. Plus généralement, Greene taxe de « rationalisations », voire même de « confabulations », les raisons déontologistes avancées. On pourrait penser qu'il s'agit là d'une manœuvre rhétorique – déconsidérer l'adversaire –, mais ce n'est pas le cas. Comme on l'a vu, Greene développe une approche de la décision morale dans laquelle les interdits absolus – et celui de l'instrumentalisation en est un – ressortissent à une structure cérébrale archaïque, émotionnellement très chargée, par opposition à la pesée rationnelle des intérêts dépendant de structures plus évoluées et donc plus récentes. Richard Dean développe ce point de la manière suivante : « Les êtres humains ont développé de fortes aversions émotionnelles contre l'acte d'infliger des dommages personnels à autrui, mais aucune réaction émotionnelle similaire dans les situations morales "impersonnelles". La force de ces aversions vient du fait qu'elles "aident les individus à disséminer leurs gènes dans un contexte social", et non parce qu'elles "reflètent des vérités morales profondes que l'on pourrait découvrir rationnellement". Ainsi, cet "aspect contingent et non moral de notre histoire" est "moralement non pertinent" [. . .]. Si la théorie déontologiste est fondée sur ces réactions, alors elle est fondée sur des facteurs moralement non pertinents et est par conséquent non fiable. » (2010, 46) Conformément à ce qu'il a affirmé plus haut, Greene continue à considérer le déontologisme et l'utilitarisme non comme deux théories morales concurrentes et de même niveau, mais comme deux modes de traitement de la prise de décision morale qu'il est possible de comparer et d'évaluer, la première étant inadéquate car archaïque (d'où les qualificatifs de « rationalisation » et de « confabulation »), et la seconde adéquate, fondée sur d'authentiques raisons.

En quoi, toutefois, le fait d'être archaïque serait-il un signe de non pertinence normative ? En ce que, nous dit Greene, cela nous rend sensible à des facteurs qui ne sont actuellement plus pertinents. Ainsi que le relève Frances Kamm : « Il semble clair que causer un dommage à quelqu'un de manière immédiate par un contact physique [le gros homme] ou par le biais de moyens mécaniques [actionner un aiguillage] n'est moralement pas pertinent. C'est pourquoi, dit Greene, les tentatives des philosophes déontologistes pour justifier leur interdit en faisant appel à d'autres facteurs apparemment pertinents ne sont que pures "confabulations" ou rationalisations » (2009, 334). Le problème cependant, comme Kamm encore le souligne, c'est que cette manière de présenter la position déontologiste n'est pas correcte – aucun déontolo-

giste n'a jamais affirmé que la distance entre l'agent et sa victime comptait moralement. Si l'on ajoute à cela l'objection de Hauser que l'invocation de la distance entre l'agent et la victime n'explique pas certaines de nos réactions (Ned et Oscar), il apparaît clairement que la distinction que fait Greene entre « raison » et « rationalisation » n'a pas de fondement solide, qu'elle est entièrement dépendante d'un biais, comme nous le soupçonnions : seules les raisons conséquentialistes sont de véritables raisons, car celles que les déontologistes invoquent sont des constructions *a posteriori* inventées pour justifier des émotions (d'alarme) suscitées par le caractère personnel de l'action envisagée, fonction de la proximité de la victime, qu'on doit expliquer par des facteurs évolutionnaires contingents. L'ennui pour Greene, c'est que si tout cela constitue une jolie histoire, cela peut difficilement compter comme de bonnes raisons pour appuyer sa thèse.

Dans ce contexte, le recours à l'expérimentation neuroscientifique relève même du rideau de fumée, puisque l'une des théories est présentée de manière erronée. C'est à cette même conclusion qu'en vient Sim Berker : la force de l'argument de Greene repose uniquement sur l'identification qu'il fait d'une part entre les dilemmes personnels et ceux qui provoquent des réactions déontologistes, et d'autre part entre les dilemmes impersonnels et ceux qui provoquent des réactions conséquentialistes. Or, non seulement cette manière de présenter les choses est hautement contestable, mais encore l'imagerie cérébrale n'est d'aucune utilité pour en traiter : elle est établie « en fauteuil », et non à l'aide de l'examen du siège cérébral des émotions et de la raison (2009, 326). J'ajouterai qu'il n'est pas très étonnant que l'on puisse rapprocher le déontologisme et les émotions d'alarme : le déontologisme est une éthique qui repose sur des principes, c'est-à-dire sur des prescriptions et des interdictions à caractère assez absolu (mais l'utilitarisme de la règle est-il très différent sur ce point ?) ; or, on le sait, le rôle de tels principes est de simplifier la complexité : un principe dit quoi faire même dans une situation embrouillée. Mais les émotions jouent aussi ce rôle : elles nous permettent d'agir et de réagir rapidement, avant que notre raison puisse entrer en action. Seulement, du fait que principes et émotions ont une caractéristique en commun, il ne s'ensuit pas qu'on puisse les mettre sur le même plan dans le champ normatif : les premiers simplifient en attirant l'attention sur les éléments moralement et rationnellement pertinents, les seconds étant, disons, le fruit d'adaptations utiles.

Greene s'est donc mépris sur l'interprétation à donner aux réactions des sujets dans le cas du *wagon fou* et dans celui du *gros homme*, ainsi que sur leur portée théorique. Ces différences peuvent s'expliquer par des raisons authentiquement morales, même si elles s'accompagnent de réactions cérébrales et émotionnelles différentes. Pourquoi s'étonner d'ailleurs de ces différences, puisqu'elles caractérisent des décisions différentes fondées sur des raisons différentes ? En outre, les raisons proprement morales qui peuvent être alléguées pour rendre compte des réactions des personnes sondées ne se limitent pas aux raisons utilitaristes (minimisation du nombre de morts)

et déontologistes (interdit de l'instrumentalisation) évoquées. Nous avons vu que Hauser et ses collègues reprochaient à Greene de ne pas tenir compte de la structure causale et intentionnelle de l'action ; cela indique une nouvelle piste qui, à mon sens, établit définitivement qu'il existe des raisons non-utilitaristes qui sont de véritables raisons morales.

4 LE PADE ET LA STRUCTURE INTENTIONNELLE DE L'ACTION

Selon Hauser et ses collègues, « les jugements observés [dans les cas de Denise et de Frank] sont compatibles avec au moins trois distinctions morales :

- 1° Dommage prévu contre dommage voulu (Principe du double effet) : il est moins permis de causer un dommage si c'est un moyen voulu pour atteindre une fin que si c'est une conséquence prévue.
- 2° Redirection d'une menace contre introduction d'une menace : il est moins permis de causer un dommage en introduisant une nouvelle menace (p. ex. en poussant un homme) qu'en redirigeant une menace déjà présente (p. ex. en faisant dévier un wagon hors de contrôle vers un homme).
- 3° Personnel contre impersonnel : il est moins permis de causer un dommage par contact physique direct que par un moyen indirect. » (2007, 15)

À quoi on pourra ajouter l'interdit kantien et libéral de l'instrumentalisation, déjà mentionné.

S'agit-il d'authentiques distinctions morales ? L'argumentation de Greene est dirigée contre la troisième, ce qui a suscité l'objection des variantes de Ned et d'Oscar. Greene accepterait-il la seconde ? Probablement pas, car il paraît qu'un utilitariste doive la rejeter. Toutefois, elle semble justifiable par des raisons indirectes ; en effet, pour prendre un autre exemple proposé par Hauser et ses collègues, on tombera aisément d'accord qu'on n'a pas le droit de rendre une personne aveugle afin que la vue de deux autres soit restaurée, mais qu'on a le droit de ne pas rendre la vue à un aveugle si cela permet de la rendre à deux autres personnes (allocation de ressources médicales rares). Raison morale donc, même pour un conséquentialiste. Je ne vais toutefois pas l'examiner, car la première distinction est bien plus intéressante pour mon propos, puisqu'elle heurte le calcul utilitariste et combine de manière assez complexe causalité et intentionalité. Il s'agit de ce qu'on appelle souvent le principe des actions à double effet (PADE).

Ce principe a été beaucoup débattu et critiqué dans la littérature, tant par des déontologistes que par des conséquentialistes. C'est en le discutant que Philippa Foot a inventé le dilemme du *wagon fou* et c'est pour le tester que Hauser et ses collègues ont imaginé les scénarios de Ned et d'Oscar. Sa formulation et son sens sont déjà sources de désaccords, mais je ne vais pas

m'engager dans une analyse détaillée de ces débats⁹. Ce qui m'intéresse, c'est la question du statut de ce principe comme principe authentiquement moral et pour cela, une présentation sommaire suffira.

Ce principe, dont l'énoncé remonte à la scolastique du XVI^e siècle, et dont on trouve des éléments chez Thomas d'Aquin déjà (*Somme théologique*, IIa-IIae, q. 64, a. 7), concerne des actions qui ont deux effets, l'un bon et l'autre mauvais. Il énonce quatre conditions dont la satisfaction simultanée est nécessaire pour que l'action soit permise. On comprend qu'il soit susceptible d'éclairer les dilemmes qui nous occupent, puisqu'ils contiennent justement deux effets, l'un bon (sauver des vies), l'autre mauvais (sacrifier une vie). Voici ces quatre conditions (appelées souvent « règles ») :

- 1° L'action doit être en elle-même bonne ou indifférente.
- 2° L'intention (préalable) de l'agent doit être bonne ; notamment, elle ne doit pas porter sur l'effet mauvais.
- 3° L'effet bon ne doit pas provenir de l'effet mauvais, ce qui signifie que l'effet mauvais ne doit pas être un moyen pour obtenir l'effet bon.
- 4° Il faut une raison proportionnée pour agir ainsi : qu'il n'y ait pas d'autre moyen d'obtenir l'effet bon et que l'effet bon soit supérieur ou au moins d'égale importance à l'effet mauvais qui est permis.

Dans les quatre variantes du dilemme, la condition 4 est satisfaite. La condition 2 est violée par Frank, vu que son intention est aussi dirigée sur l'effet mauvais. Les deux cas impliquant un gros homme (Frank et Ned) violent la condition 3, même si Ned utilise simplement la présence préalable du gros homme, sans le pousser lui-même activement. Denise et Oscar respectent les conditions 1, 2 et 3 : ils n'utilisent en rien l'ouvrier-victime, qui ne joue aucun rôle causal dans le sauvetage des cinq autres ; c'est un mal prévu, non un mal voulu (comme moyen). Pourtant, ils savent que l'ouvrier va mourir et agissent en connaissance de cause. Distinguer la valeur d'un mal parce qu'il est voulu ou parce qu'il est prévu, est-ce vraiment une distinction proprement morale ou une rationalisation ? On peut diverger sur cette question, mais si on veut aller au-delà du conflit d'intuitions, il faut examiner la condition 1.

À première vue, cette condition est une sorte de tautologie : une action n'est permise que si elle est bonne ou indifférente ; c'est de plus une pétition de principe, puisque la bonté d'une action sera définie différemment selon la doctrine morale qu'on adopte. Mais les apparences sont trompeuses, car bien des choses dépendent de la manière dont on définit une action. La doctrine du PADE a été élaborée dans une tradition téléologiste – par conséquent ni déontologiste, ni conséquentialiste –, pour laquelle ce qui définit l'action et détermine sa valeur, ainsi que la valeur de l'intention qui l'habite, c'est le but qu'elle vise. Je précise.

9. Deux analyses du PADE m'apparaissent particulièrement éclairantes : Goffi (2004) et Timmons (2002, 77-85).

Une action intentionnelle est caractérisée par son objet intentionnel ou son contenu, que j'appellerai son *but interne*. Ainsi, par exemple, ce qui fait qu'un homicide est un *meurtre*, c'est le fait qu'il est une action intentionnelle dont le but interne et voulu est la mort d'un être humain ; il devient un *assassinat* lorsque son but interne est la mort *préméditée* d'un être humain. Mais la fonction du but interne ne se limite pas à cet aspect conceptuel : il permet encore de déterminer la valeur de l'action : la mort d'un être humain est toujours un mal, et quand elle est visée ou voulue, c'est un mal moral. Un acte qui vise formellement un mal est *intrinsèquement* mauvais (comme est *intrinsèquement* bon un acte qui vise formellement un bien), cela, quelles que soient les intentions de l'agent – c'est peut-être pourquoi l'on dit que le chemin de l'enfer est pavé de bonnes intentions.

Dans la philosophie contemporaine, cette conception de l'action morale a été défendue notamment par Elisabeth Anscombe. On peut la résumer ainsi :

- 1° Une action intentionnelle est caractérisée par son but interne.
- 2° La valeur morale de l'action et la droiture de l'intention de l'agent sont déterminées par la valeur du but interne et non par les conséquences.
- 3° Une action est intrinsèquement mauvaise si son but interne est mauvais.
- 4° Une action est intrinsèquement bonne si son but interne est bon.

Il s'ensuit, nous dit Anscombe, – et ici on retrouve le PADE – que « choisir de tuer un innocent comme un moyen pour atteindre ses propres fins est toujours un meurtre. Naturellement, tuer un innocent comme une fin en elle-même est aussi un meurtre. [Mais] tuer un innocent, même si vous savez d'une certitude statistique que ce que vous faites l'implique, n'est pas nécessairement un meurtre. » (1957, 66)

Cette conception de l'action est bien différente de celle que défend l'utilitarisme. Pour ce dernier, nous dit Mark Timmons : « Une manière directe de répondre à cette question [que faire en présence de deux effets, l'un bon et l'autre mauvais] est de voir *combien* de bien et *combien* de mal résulterait de l'action et d'agir si la quantité de bien prévu l'emporte sur la quantité de mal prévu » (2002, 76). Contrairement à cela, commente Jean-Yves Goffi, « ce que permet le PADE, c'est de comparer les conséquences bonnes et mauvaises d'un acte reconnu comme licite de son espèce [c'est-à-dire qui n'est pas intrinsèquement mauvais], afin de savoir s'il est approprié de l'accomplir dans une situation embrouillée » (2004, 238).

À la lumière de cette conception de l'action morale, il faut dire que, si chaque fois, ce que font les quatre agents (Denise, Oscar, Ned et Frank) a deux conséquences identiques, l'une bonne (sauver cinq vies), l'autre mauvaise (la mort d'une personne), l'action elle-même doit cependant être qualifiée fort différemment. Denise, Oscar et Ned actionnent un aiguillage, ce qui est une action moralement indifférente, alors que Frank pousse un gros homme, ce qui est un meurtre, donc une action intrinsèquement mauvaise (violation des conditions 1 et 2). Comme Ned viole la condition 3, seuls Denise et Oscar s'en

sortent moralement selon le PADE. La différence des jugements concernant Oscar et Denise reste alors à expliquer ; peut-être vient-elle de ce que le cas d'Oscar a été présenté juste après celui de Ned, avec lequel il a bien des affinités et qui paraît (mais faussement) ne différer de lui que par un élément de peu d'importance (la présence d'un lourd obstacle sur la voie) ? Peut-être vient-elle de ce qu'Oscar demeure une sorte d'instrument en ce qu'il fait corps avec l'obstacle, ce qui n'est pas du tout le cas de Denise ? Il est impossible de le savoir sans de plus amples investigations.

Quoiqu'il en soit de la validité de ces explications, la différence entre les deux cas pourrait aussi tenir au fait que la morale commune fait intervenir d'autres principes que le PADE. Non seulement tous les sondés ne l'appliquent pas et certains jugent Denise et Oscar différemment, mais encore, moyennant d'autres expériences, on se rend compte que la morale commune n'utilise pas ce principe dans toutes les situations, notamment lorsque l'agent est lui-même menacé. On connaît l'histoire imaginée par Robert Nozick et qu'Helen Frowe appelle *Le fusil à rayons* : quelqu'un est au fond d'un puits ; une tierce personne, poussée par-dessus la margelle, tombe et va l'écraser. Celui qui est au fond n'a qu'un moyen de sauver sa vie : utiliser son fusil à rayon pour désintégrer la personne qui constitue une menace vitale pour lui (2008, 278). Nombreux sont ceux qui estiment qu'il est permis de désintégrer la personne qui tombe pour sauver sa propre vie, alors même que c'est une personne innocente. Or cela viole le PADE, puisque l'action est intrinsèquement mauvaise – tuer directement un innocent –, et que l'effet bon est obtenu par le moyen de l'effet mauvais. Mais peut-être ici l'acte est-il moins permis qu'excusable ? Voilà encore une distinction morale importante lorsqu'il s'agit d'évaluer des actions intentionnelles.

Une autre piste intéressante à explorer pour expliquer les réactions différentes des personnes sondées est celle du réseau d'héritage des propriétés¹⁰. L'idée est que, lorsqu'on passe d'un état *A* à un état *B*, *B* étant un état indésirable qui sera le résultat de notre action, et qu'on peut le faire par divers chemins, on préfère le chemin le plus long. Les dilemmes mentionnés partent tous d'un wagon hors de contrôle et aboutissent à la mort d'une personne si l'on décide d'agir. La situation de Denise comporte deux étapes intermédiaires : actionner l'aiguillage et la bifurcation du tram, alors que celle de Frank n'en comporte qu'une : pousser le gros homme (l'arrêt du tram étant postérieur à la mort du gros homme). Le chemin de Denise est plus long et donc moins problématique. Mais pourquoi préfère-t-on un chemin plus long ? Pour deux raisons sans doute : d'abord parce qu'il éloigne la responsabilité, et ensuite parce qu'il multiplie les possibilités d'interventions tierces capables de se mettre au travers du résultat indésirable. Par exemple, l'action de l'aiguillage pourrait faire dérailler le wagon et donc éviter toute victime. Cela toutefois ne compte pas forcément comme raisons morales – il ne s'agit peut-

10. Cela m'a été suggéré par Pierre Livet.

être que de raisons prudentielles – et n'explique pas tout, notamment pas pourquoi les réponses concernant Denise, Oscar et Ned sont différentes, car dans les trois cas, il y a deux étapes.

5 CONCLUSION

Nous voulions savoir si une investigation empirique des réponses données au dilemme du *wagon fou* permettait, comme le pense Greene, de trancher entre le déontologisme et l'utilitarisme. Nous avons vu, en nous basant sur l'approche proposée par Hauser et ses collègues, que ce n'est pas le cas : l'opposition de ces deux théories morales est de nature *authentiquement* normative – et par conséquent toutes les deux sont capables de fournir des justifications morales –, il ne s'agit pas du contraste entre une théorie normative et un ensemble de réactions émotionnelles rationalisées *a posteriori*. Certes, bien des considérations morales que nous formulons constituent de telles rationalisations, mais cela ne concerne pas plus une théorie morale qu'une autre – d'autant que nous donnons nos raisons lorsqu'on nous les demande, et on nous les demande généralement *après* que nous avons agi, d'où la possibilité récurrente du soupçon. En outre, et c'est plus important, au-delà de tout ce que chacun peut invoquer pour justifier une décision qu'il a prise, la *nature* des raisons alléguées est ce qui compte. Un agent peut rationaliser ou confabuler ; savoir si c'est le cas ou non est pertinent si l'on veut évaluer sa bonne foi ou sa lucidité sur lui-même, mais n'a aucun impact sur le caractère authentiquement moral et la validité de ce qu'il allègue pour justifier sa conduite. Dans le cas des variantes du dilemme, on peut rendre compte de la différence entre les actions projetées en faisant appel à des considérations de ce genre. Le PADE en fait partie ; il est sous-tendu par une théorie morale et une théorie de l'action complexes, qui a l'avantage pour notre question d'être « neutre » en ce qu'il n'est ni déontologiste ni conséquentialiste. Si de telles théories n'ont pas droit au label « moralement authentique », on voit mal qui y aurait droit et selon quel critère.

Cela dit, considérer le PADE ou tout autre principe moral comme un authentique principe normatif n'implique pas qu'il soit un principe *adéquat* ou *correct*. J'ai mentionné qu'il existe de nombreux auteurs qui l'ont critiqué au nom de considérations morales, utilitaristes souvent, déontologistes parfois. Mais il s'est agi et il s'agit encore de débats proprement normatifs, dont on n'est pas étonné qu'ils divisent la communauté des philosophes. Il est certes vrai que les situations « conséquentialistes » (Denise) sont accompagnées d'émotions différentes que les situations « déontologistes » (Frank) et donnent lieu à des décisions différentes, mais ces différences ne se laissent pas expliquer par l'opposition de l'authentique et de l'inauthentique. Il reste qu'il est très instructif d'utiliser l'imagerie cérébrale pour étudier nos réactions morales, particulièrement à des dilemmes, car ils mettent à l'épreuve nos théories ; ce faisant, ils en révèlent la structure et les modes d'argumentation. Greene en attendait plus, pensant qu'ils permettraient de trancher entre

les diverses théories ; mais c'est probablement trop leur demander.

BIBLIOGRAPHIE

- Anscombe G. E. M. (1957), « Mr Truman's Degree », in *Ethics, Religion and Politics (Collected Philosophical Papers, vol. III)*, Oxford, Blackwell, 1981, p. 62-71.
- Baertschi B. (1995), *La valeur de la vie humaine et l'intégrité de la personne*, Paris, PUF.
- Baertschi B. (2009), *La Neuroéthique*, Paris, La Découverte.
- Berker S. (2009), « The Normative Insignificance of Neuroscience », *Philosophy and Public Affairs*, n° 37, p. 293-329.
- Dean R. (2010) « Does Neuroscience Undermine Deontological Theory? », *Neuroethics*, n° 3, p. 43-60.
- Foot P. (1978) « The Problem of Abortion and the Doctrine of the Double Effect » in *Virtues and Vices and Other Essays in Moral Philosophy*, Oxford, Blackwell, p. 19-32.
- Frowe H. (2008), « Equating Innocent Threats and Bystanders », *Journal of Applied Philosophy*, n° 25, p. 277-290.
- Gazzaniga M. (2005), *The Ethical Brain*, New York, Dana Press.
- Goffi J.-Y. (2004), « Le principe des actions à double effet », in J.-Y. Goffi, dir., *Hare et la philosophie morale*, Paris, Vrin, p. 231-248.
- Greene J. (2003), « From Neural "Is" to Moral "Ought" », in W. Glannon, dir., *Defining Right and Wrong in Brain Science*, New York, Dana Press, 2007, p. 221-229.
- Greene J., Nystrom L., Engell A. and Cohen J. (2004), « The Neural Bases of Cognitive Conflict and Control in Moral Judgment », *Neuron*, n° 44, p. 389-400.
- Greene J. (2008), « The Secret Joke of Kant's Soul », in W. Sinnott-Armstrong, *Moral Psychology*, Cambridge Mass., MIT Press, vol. 3, p. 35-79.
- Haidt J. (2001), « The Emotional Dog and its Rational Tail : A Social Intuitionist Approach to Moral Judgment », *Psychological Review*, vol. 108, p. 814-834.
- Hauser M., Cushman F., Young L., Kang-Xing R. and Mikhail J. (2007), « A Dissociation Between Moral Judgments and Justification », *Mind and Language*, vol. 22, p. 1-21.
- Kahane G., Wiech K., Schakel N., Farias M., Savulescu J. and Tracey I. (2011), « The Neural Basis of Intuitive and Counterintuitive Moral Judgment », *Social Cognitive and Affective Neuroscience*, advance access published March 18.
- Kamm F. (2009), « Neuroscience and Moral Reasoning : A Note on Recent Research », *Philosophy and Public Affairs*, n° 37, p. 330-345.
- Kant E. (1980), *Fondements de la métaphysique des mœurs*, Paris, Vrin.
- Kant E. (1985), *Doctrine de la vertu*, Paris, Vrin.

Koenigs M., Young L., Adolphs R., Tranel D., Cushman F., Hauser M. and Damasio A. (2007), « Damage to the Prefrontal Cortex Increases Utilitarian Moral Judgements », *Nature*, n° 446, p. 908-911.

Levy N. (2007), *Neuroethics*, Cambridge University Press, Cambridge.

Nadelhoffer T. & Feltz A. (2008), « The actor-observer bias and moral intuitions : Adding fuel to Sinnott-Armstrong's fire », *Neuroethics* n° 1, p. 133-144.

Roskies A. (2006), « A Case Study of Neuroethics : the Nature of Moral Judgment », in J. Illes, dir., *Neuroethics*, Oxford, Oxford University Press, p. 17-32.

Tassy S. (2011), « La nécessité de distinguer le jugement et le choix subjectif dans les neurosciences cognitives de la morale », *Médecine/sciences*, à paraître.

Timmons M. (2002), *Moral Theory*, Lanham, Rowman & Littlefield.

Young L., Bechara A., Tranel D., Damasio H., Hauser M. and Damasio A., (2010), « Damage to Ventromedial Prefrontal Cortex Impairs Judgment of Harmful Intent », *Neuron*, n° 65, p. 845-851.