



FRANS DE WAAL

PRIMATES
ET PHILOSOPHES

Le Pommier





Titre original: *Primates and Philosophers. How Morality Evolved*

© Princeton University Press, 2006

Pour la traduction française

© Le Pommier, 2008 ; 2020, pour la présente édition

Tous droits réservés

ISBN 978-2-7465-2228-2

170 bis, boulevard du Montparnasse, 75014 Paris

www.editions-lepommier.fr

Frans de Waal

Primates et philosophes

Traduction de Sarah Gurcel

Le Pommier

Introduction

Josiah Ober et Stephen Macedo

Dans les Tanner Lectures, série de conférences sur les valeurs humaines proposées par la Fondation Tanner, qui sont devenues l'essai central de ce livre, Frans de Waal met à profit ses décennies de recherche sur les primates, ainsi que la capacité qu'on lui connaît à réfléchir en profondeur à la signification de l'évolution, pour aborder une question fondamentale relative à la moralité humaine. Trois philosophes distingués ainsi qu'un éminent spécialiste de la psychologie évolutionniste réagissent à la manière dont de Waal a défini son sujet et à sa réponse même. À son tour, de Waal répond à leurs critiques dans une troisième partie. Si le désaccord entre les cinq essayistes est grand, tant sur la formulation du problème que sur la réponse à y apporter, il existe néanmoins entre eux de nombreux terrains d'entente. Notons pour commencer que tous acceptent l'opinion scientifique reçue selon laquelle l'évolution biologique est basée sur un processus de sélection

naturelle aléatoire. Aucun d'eux ne laisse entendre qu'on aurait des raisons de considérer les êtres humains comme essentiellement différents des animaux d'un point de vue métaphysique ; aucun ne fonde en tout cas son argumentation sur le fait que les êtres humains seraient les seuls à posséder une âme qui les transcende.

Autre hypothèse de départ commune à de Waal et à ses quatre commentateurs : le bien est un phénomène réel sur lequel on peut tenir un discours de vérité. Elle requiert, au minimum, de tenir compte des autres. Le mal, du même coup, participe d'une forme d'égoïsme qui pousse à se conduire envers les autres de manière inadéquate en ignorant leurs intérêts ou en les traitant comme de simples instruments. C'est sur cette double base, d'une science évolutionniste et d'une réalité morale, que s'établissent les limites du débat sur les origines du bien dont fait état le présent livre. N'y sont donc pas représentés ceux qui, par foi religieuse, adhèrent à l'idée que, seuls parmi les créatures vivantes, les hommes ont reçu certains attributs (tels que le sens moral) par la grâce de Dieu. Ne sont pas non plus représentés les penseurs des sciences sociales pour qui l'essence de la nature humaine, conformément à une certaine version de la théorie de l'agent rationnel, se définit par une tendance irréductible à choisir l'égoïsme (profiter des situations, tricher) plutôt que la coopération volontaire. Sont enfin également exclus du débat les relativistes, qui estiment qu'une action ne peut être jugée moralement bonne ou mauvaise que localement, au regard de considérations contingentes et contextuelles. Ce que nous vous proposons donc ici, c'est un débat entre cinq chercheurs qui partagent certaines conceptions de base sur la science et la morale,

une conversation sérieuse et animée entre des scientifiques fermement convaincus de la valeur et de la validité de la science comme d'une morale tournée vers autrui.

Le problème auquel sont confrontés de Waal et ses commentateurs est le suivant : étant donné que nous avons toutes les raisons scientifiques de penser que l'égoïsme (au moins au niveau génétique) est un mécanisme primaire de sélection naturelle, comment expliquer l'attachement des êtres humains à la valeur morale du Bien ? Pour formuler la question différemment, comment se fait-il que nous ne trouvions pas acceptable d'être mauvais ? Pour ceux qui croient en la réalité du sens moral, mais n'ont pas recours, pour l'expliquer, à l'hypothèse théologique d'une grâce divine accordant aux hommes et à eux seuls la propension à faire le bien, il s'agit là d'une question ardue et cruciale.

De Waal s'oppose à certaines réponses au problème de la genèse de la moralité qu'il rassemble sous le terme de « théorie du vernis », selon laquelle notre morale ne serait qu'une mince couche de vernis appliqué à une base immorale ou amoral – théorie assez largement répandue, du moins jusqu'à tout récemment, d'après de Waal. Il s'attaque d'abord à Thomas Huxley, dit « le bouledogue de Darwin », surnom que lui valut une défense acharnée de la théorie de l'évolution à la fin du XIX^e siècle. Pour de Waal, Huxley trahit ses propres convictions darwiniennes en comparant la morale à du « jardinage », à savoir à une bataille de chaque instant contre les mauvaises herbes de l'immoralisme menaçant toujours d'envahir la psyché humaine. De Waal attaque également des théoriciens du contrat social (notamment Thomas Hobbes), dont le postulat de départ tient l'homme pour fondamentalement asocial, voire antisocial, et certains

biologistes évolutionnistes qui, à son avis, ont une tendance outrancière à généraliser le rôle avéré de l'égoïsme dans le processus de sélection naturelle.

Aucun des cinq essayistes de cet ouvrage ne se réclame de la « théorie du vernis » telle que la définit de Waal. Comme il en existe toutefois diverses conceptions possibles, il sera peut-être utile d'établir une sorte d'idéal-type de cette théorie, quelles que soient les limites inhérentes à l'exercice. L'idéal-type de la théorie du vernis part du principe que les hommes sont par nature bestiaux et donc mauvais – c'est-à-dire impitoyablement égoïstes – et qu'on doit par conséquent s'attendre à ce qu'ils se conduisent mal – c'est-à-dire à ce qu'ils traitent les autres de façon inadéquate. Tout observateur sera pourtant contraint d'admettre qu'il arrive que les hommes traitent autrui convenablement et fassent le bien, exactement comme s'ils étaient bons. Étant donné que, selon la logique de cette théorie, les hommes sont intrinsèquement mauvais, leur bon comportement ne saurait alors s'expliquer que comme le produit d'un vernis de moralité qui recouvrirait mystérieusement leur mauvais fond. De Waal reproche à la théorie du vernis de ne pas identifier la *source* de ce vernis : si donc il n'existe qu'hors du règne naturel, il doit, à ce titre, être considéré comme un mythe par tous ceux qui sont attachés à une explication scientifique des phénomènes naturels.

Et, si la théorie du vernis repose sur un mythe, le phénomène de la bonté humaine doit trouver une autre explication. De Waal commence par renverser la proposition de départ : les êtres humains sont, selon lui, naturellement bons. Nous aurions hérité cette « bonne nature », entre autres choses, de nos ancêtres non humains par le biais du processus de sélection naturelle identifié par Darwin. Afin

de tester cette hypothèse, de Waal nous invite à observer avec lui le comportement de nos plus proches parents non humains – d’abord les chimpanzés, puis d’autres primates moins proches de nous et enfin d’autres animaux sociaux non primates. Si nos plus proches parents agissent effectivement comme s’ils étaient bons, et si nous autres, êtres humains, agissons aussi comme si nous étions bons, le principe méthodologique de parcimonie nous pousse à supposer que cette bonté est bien réelle, qu’il est naturel de vouloir faire le bien et que la moralité des hommes et celle de leurs parents phylogéniques ont une source commune.

Si les comportements humains manifestent un degré de bonté plus développé que les comportements non humains, de Waal nous invite néanmoins à considérer la moralité non humaine, plus simple, comme le *fondement*, au sens profond, d’une moralité humaine plus complexe. Les preuves empiriques en faveur d’une théorie de l’« anti-vernis » qui lie les morales humaine et non humaine reposent sur une observation minutieuse du comportement de nos cousins primates. Au cours de sa longue et brillante carrière, de Waal a en effet étudié avec le plus grand soin le comportement des primates, observant et consignait de nombreuses manifestations de bonté. Il a acquis par là un immense respect et une grande tendresse pour ses sujets d’observation. Le plaisir de lire ses écrits sur les primates – un plaisir patent chez chacun des commentateurs – vient en partie de la joie que lui-même tire de son travail auprès des chimpanzés, des bonobos et des capucins, et de ce qu’il les considère comme ses collaborateurs dans une entreprise capitale.

De Waal conclut que la capacité humaine à bien agir, de temps en temps du moins, plutôt qu’à toujours mal agir,

trouve son origine, en termes d'évolution, dans certaines émotions que nous partageons avec d'autres animaux, réactions involontaires (non choisies, prérationnelles) et physiologiquement évidentes (et donc observables) à ce qui arrive aux autres. Parmi ces réactions émotionnelles, l'empathie est d'une importance fondamentale. De Waal explique que, dans un premier temps, la réaction empathique est affaire de « contagion émotionnelle ». La créature A s'identifie directement à la créature B et à ce qu'elle vit, jusqu'à, pour ainsi dire, « ressentir sa détresse ». À ce niveau, l'empathie est encore, dans une certaine mesure, égoïste : A cherche à consoler B parce que, A ayant « attrapé » la détresse de B, il a lui-même besoin d'être consolé. À un niveau supérieur, l'empathie émotionnelle peut laisser place à la sympathie, c'est-à-dire à la reconnaissance du fait que B a des envies ou des besoins particuliers liés à sa situation propre, besoins et envies qui sont donc différents de ceux de A. De Waal cite l'exemple révélateur, et charmant, d'un chimpanzé essayant d'aider un oiseau blessé à voler. Le chimpanzé serait bien sûr incapable de voler : il lui a donc fallu prendre en compte les besoins spécifiques de l'oiseau et sa manière singulière d'être au monde.

Si l'on constate couramment des exemples de contagion émotionnelle chez de nombreuses espèces, les manifestations de sympathie ne s'observent, elles, que chez certains grands singes. Parmi les réactions émotionnelles liées à la sympathie et propices à un bon comportement, on trouve l'altruisme réciproque et peut-être même un certain sens de l'équité, encore que ce dernier point soit contesté (comme le rappelle Philip Kitcher). Encore une fois, on n'observe les formes les plus complexes, ou les plus sophistiquées, que

revêtent ces comportements à motifs émotionnels (selon de Waal) que chez les grands singes et quelques rares autres espèces – éléphants, dauphins et singes capucins.

Les réactions émotionnelles sont, pour de Waal, les composantes fondamentales (ou « blocs de construction ») de la moralité humaine. Le comportement moral humain est beaucoup plus sophistiqué que celui de n'importe quel autre animal, mais de Waal estime qu'il y a *continuité* entre les comportements moraux humain et non humain – de la même manière qu'il y a continuité entre la compassion chez les chimpanzés et la contagion émotionnelle chez d'autres animaux. Cette continuité rend inutile l'idée d'une moralité se surajoutant mystérieusement à une base immorale. De Waal nous invite à nous représenter les êtres humains non comme des nains de jardin en terre cuite recouverts d'un vernis de peinture criarde, mais comme des « poupées russes » présentant une continuité ontologique entre le moi moral externe et toute une série de « moi préhumains » internes – les différentes couches successives, de l'enveloppe extérieure à la minuscule figurine du centre, étant toutes, et de façon homogène, moralement bonnes.

Comme en témoigne la vigueur des réactions qu'elle a suscitées, la conception de la genèse et de la nature de la moralité humaine défendue par de Waal est à la fois stimulante et sujette à controverse. Les quatre commentateurs sont d'accord avec de Waal pour dire que l'idéal-type de la théorie du vernis est *a priori* peu alléchant, bien qu'ils divergent légèrement sur ce que recouvre exactement cette notion, de même que sur la question de savoir si l'on peut raisonnablement y souscrire, du moins dans la version musclée évoquée plus haut. Pourtant, en définitive, chacun

d'entre eux défend un point de vue qui s'apparente plus ou moins à cette théorie du vernis. Robert A. Wright ne s'en cache pas, puisqu'il appelle sa position la « théorie naturaliste du vernis ». De fait, comme le souligne Peter Singer, de Waal lui-même reconnaît combien l'effort humain pour élargir le « cercle de l'obligation morale » aux étrangers à la communauté est « fragile » : une formule qui semble impliquer qu'on puisse considérer certaines formes reculées de la morale humaine comme une sorte de vernis.

Les inquiétudes exprimées par de Waal quant à l'extensibilité du « cercle de l'obligation morale » participent de ce qui pousse ses commentateurs à établir une ligne de démarcation claire entre la moralité humaine et le comportement animal. Ils sont fermement convaincus que l'« authenticité » (Kitcher) de la morale tient à son application potentiellement universelle. Une telle conviction exclut les animaux de la sphère des êtres authentiquement moraux ; elle les place « au-delà du jugement moral », selon la formule de Christine M. Korsgaard, puisque les animaux non humains ne se comportent pas universellement bien. La tendance à la partialité en faveur des membres du groupe est une constante parmi les animaux sociaux non humains. Il est vrai que les êtres humains peuvent aussi connaître cette tendance, comme l'indique de Waal – Wright va jusqu'à penser qu'il s'agirait là d'une menace endémique pour la moralité humaine. Mais Kitcher, Korsgaard et Singer soulignent tous que les hommes traitent l'universalisation des êtres à qui sont dus des devoirs moraux (« êtres » au sens de personnes ou, dans la perspective de Singer, de toute créature ayant des intérêts) comme *conceptuellement possible* (certains philosophes humains allant jusqu'à traiter cette

universalisation comme conceptuellement essentielle). En tout cas, ils la mettent parfois en application.

Les commentateurs posent tous une question similaire, encore que dans des registres philosophiques passablement différents : si même les plus avancés des animaux non humains se contentent généralement de bien se comporter envers leurs proches (parents ou membres de la même communauté), peut-on véritablement qualifier leur comportement de *moral* ? En cas de réponse négative (conclut chacun d'entre eux), il faudra postuler une *discontinuité* entre les capacités humaines et les capacités naturelles des espèces non humaines. De Waal reconnaît le problème, puisqu'il écrit (comme Singer le note, que « ce n'est *que* lorsque nous émettons ces jugements généraux, impartiaux, que l'on peut vraiment parler d'approbation ou de désapprobation *morales*¹. »

La discontinuité la plus flagrante entre les capacités des hommes et celles des animaux non humains touche au langage et à la mise en œuvre consciente de la raison que nous associons de près à l'usage strictement humain du langage. Parole, usage du langage et raison sont bien entendu liés à la cognition. Que pouvons-nous donc dire de la cognition non humaine ? Aucun des auteurs de ce livre ne tient une espèce non humaine pour l'*égale* cognitive de l'homme, mais la question de savoir si *seuls* les hommes sont capables de raisonnement moral demeure.

C'est à ce stade du débat que la définition de l'*anthropomorphisme* devient un enjeu brûlant ; Wright, notamment, se concentre sur ce point. De Waal est un partisan

1. C'est Peter Singer qui souligne. (*Note du traducteur.*)

ardent et réfléchi d'une version critique et parcimonieuse de l'anthropomorphisme scientifique – qu'il distingue fermement de l'anthropomorphisme sentimental employé à tort et à travers dans de nombreux ouvrages sur les animaux à destination du grand public (aussi attachants que soient ces documents). Il ne serait pas juste de qualifier l'un ou l'autre des commentateurs de tenant de l'« anthropodéni », terme proposé par de Waal pour désigner le refus (peut-être motivé par un rejet esthétique de la nature) de reconnaître toute continuité entre les hommes et les autres animaux. Le débat tenu chez les philosophes et les éthologues sur la spécificité humaine s'est beaucoup concentré sur la question de savoir si un animal non humain était capable de développer quelque chose du type théorie de l'esprit, c'est-à-dire si la capacité d'imaginer ce qui se passe dans la tête d'une autre créature, en tant que ce qui s'y passe est différent de ce qui se passe dans la tête du sujet, était strictement humaine ou pas. Les données expérimentales existantes sont susceptibles de corroborer les deux options. De Waal remarque que certains chimpanzés se reconnaissent dans un miroir (apportant ainsi la preuve d'une conscience de soi souvent considérée comme la condition nécessaire à une théorie de l'esprit) et rappelle à plusieurs reprises qu'il est singulièrement anthropocentrique d'exiger des grands singes qu'ils soient capables de formuler une théorie de l'esprit *humain*. La question d'une théorie de l'esprit non humain reste néanmoins ouverte : de nouvelles recherches s'imposent dans ce domaine.

Kitcher et Korsgaard distinguent vigoureusement un comportement animal motivé par l'émotion de la moralité humaine, dont ils estiment qu'elle repose sur une conscience

cognitive permettant de juger de l'« à-propos moral » de ce qu'on envisage de faire. Kitcher définit son critère en faisant de la théorie humienne-smithienne du spectateur impartial une forme de conscience de soi qui nécessiterait l'aptitude au langage. Korsgaard en appelle à la conception kantienne d'un autogouvernement autonome comme fondement nécessaire d'une moralité authentique. Kitcher et Korsgaard décrivent tous deux les animaux non humains comme des créatures fluctuantes (*wanton*), dominées par leur corps et leurs émotions, reprenant à leur compte le concept élaboré dans un autre contexte par l'éthicien Harry Frankfurt. Il manque aux créatures fluctuantes de Frankfurt un mécanisme leur permettant de choisir de façon homogène parmi les diverses motivations qui, de temps à autre, sont susceptibles de les pousser à agir. On ne peut donc pas dire qu'elles soient guidées par un raisonnement conscient appliqué à la bienséance ou à la justesse des actions qu'elles envisagent. À ce stade, cependant, se pose la question de savoir si Kitcher et Korsgaard ne placent pas la barre de la moralité si haut que la plupart des actions *humaines* ne sauraient être qualifiées de morales. Chacun des deux philosophes propose de la morale une *version normative*, la définissant comme la façon dont les gens devraient se comporter, plutôt qu'une *description* de la manière dont la plupart d'entre nous se comportent effectivement la plupart du temps. Si les comportements effectifs de la grande majorité des êtres humains indiquent qu'ils agissent en créatures fluctuantes, cela ôte pas mal de piquant à l'assertion selon laquelle tous les animaux non humains se comportent aussi en créatures fluctuantes.

La même question se pose lorsque Singer analyse ce que les éthiciens appellent le « problème du trolley ». Selon

une perspective conséquentialiste¹ de l'agglomération des intérêts, la raison morale exigerait que, dans certaines circonstances, il faille pousser quelqu'un sous les roues d'un trolley sans conducteur pour sauver cinq autres vies (par hypothèse, on est soi-même trop léger pour espérer bloquer le trolley de son propre corps, alors que l'individu que l'on pousse a une masse corporelle suffisante). Singer se réfère aux études de l'activité cérébrale d'individus confrontés à des questions du type : « Faut-il tuer une personne pour en sauver cinq ? » Ceux qui répondent par la négative prennent une décision rapide et l'activité cérébrale au moment de la prise de décision apparaît surtout dans les zones associées aux émotions. Chez ceux qui répondent par la positive, on remarque en revanche une plus grande activité dans les zones cérébrales associées à la cognition rationnelle. Singer en conclut que ce qu'il estime être la réponse juste moralement (il faut tuer une personne pour en sauver cinq) est aussi la réponse rationnelle. Il reconnaît toutefois que seule une minorité de gens donnent la bonne réponse : la plupart d'entre nous ne disent pas qu'ils seraient eux-mêmes prêts à tuer quelqu'un pour sauver cinq vies. Singer ne cite aucun cas répertorié de gens qui auraient *effectivement* poussé quelqu'un sous les roues d'un trolley.

Le fait est que les arguments avancés par de Waal quant aux réactions émotionnelles des primates reposent *exclusivement* sur l'observation de comportements effectifs, quelle que soit la forme des données, quantitative ou récit

1. Le conséquentialisme est un point de vue moral qui prend les conséquences pour seul critère normatif, c'est-à-dire que seules les conséquences d'une action donnée doivent constituer la base d'un jugement moral de ladite action. (*Note du traducteur.*)

d'anecdotes. De Waal est obligé de rendre compte de la morale des primates en termes de ce qu'ils font, puisqu'il n'a pas accès à leur notion de ce qu'ils *devraient* faire idéalement pour répondre aux exigences de leur sens moral ou à la manière dont ils pensent qu'ils devraient se comporter dans une situation hypothétique. À confronter le *comportement* des primates (données quantitatives et récits d'anecdotes) aux *idéaux normatifs* des êtres humains, on risque donc de comparer des choses incomparables. Ses critiques rétorqueraient bien sûr à de Waal que c'est bien le problème : les animaux non humains n'ont pas la notion de la manière dont ils devraient se comporter, et d'ailleurs pas de notion tout court, puisqu'ils n'ont d'aptitude ni à la parole ni au langage ni à la raison.

Les animaux non humains sont incapables de formuler et de communiquer des idéaux normatifs. Doit-on en conclure qu'il existe un fossé entre le type de comportement « moral » émotionnellement motivé que de Waal et d'autres ont pu observer chez les primates et les actions humaines « authentiquement » morales, fondées sur la raison ? Si le secrétaire d'édition du présent ouvrage connaissait la réponse, il saurait auquel des mots de la phrase précédente, « moral » et « authentiquement », il conviendrait d'ôter les guillemets. Sur ce choix repose dans une large mesure notre compréhension de nous-mêmes et des espèces avec lesquelles nous partageons la Terre. L'un des objectifs de ce livre est d'inciter ses lecteurs à réfléchir à la réponse qu'eux-mêmes apporteraient. Nous vous proposons d'assister à une conversation entre des savants dévoués à l'étude du comportement des primates et d'autres tout aussi dévoués à celle de la morale humaine. L'existence même de ce livre est la preuve que

ces deux groupes ne sont pas totalement distincts l'un de l'autre. Nous souhaitons encourager une réflexion interdisciplinaire et inviter tous ceux qu'intéressent le bien et sa genèse, tant chez les animaux humains que non humains, à poursuivre le débat.