

Cet article est disponible en ligne à l'adresse :

[http://www.cairn.info/article.php?ID\\_REVUE=RDM&ID\\_NUMPUBLIE=RDM\\_031&ID\\_ARTICLE=RDM\\_031\\_0242](http://www.cairn.info/article.php?ID_REVUE=RDM&ID_NUMPUBLIE=RDM_031&ID_ARTICLE=RDM_031_0242)

---

## Les fondements utilitaristes et anti-utilitaristes de la coopération en biologie

par Bruno KESTEMONT

| Éditions La Découverte | Revue du Mauss

2008/1 - n° 31

ISSN 1247-4819 | ISBN 978-2-7071-5491-0 | pages 242 à 269

---

Pour citer cet article :

— Kestemont B., Les fondements utilitaristes et anti-utilitaristes de la coopération en biologie, Revue du Mauss 2008/1, n° 31, p. 242-269.

---

Distribution électronique Cairn pour les Éditions La Découverte.

© Éditions La Découverte. Tous droits réservés pour tous pays.

La reproduction ou représentation de cet article, notamment par photocopie, n'est autorisée que dans les limites des conditions générales d'utilisation du site ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Toute autre reproduction ou représentation, en tout ou partie, sous quelque forme et de quelque manière que ce soit, est interdite sauf accord préalable et écrit de l'éditeur, en dehors des cas prévus par la législation en vigueur en France. Il est précisé que son stockage dans une base de données est également interdit.

# Les fondements utilitaristes et anti-utilitaristes de la coopération en biologie

*Bruno Kestemont*

## Introduction

Alors que les sciences naturelles ont semblé donner raison à l'utilitarisme pendant plus d'un demi-siècle, il devient aujourd'hui possible de contredire cette hypothèse : l'homme n'est pas *uniquement* égoïste, en plus du fait qu'il n'est pas parfaitement rationnel. Ce non-utilitarisme de l'être humain s'applique aux formes les plus primitives de la vie : la coopération est inscrite dès les origines et pourrait être, avec la survie du plus apte, un des moteurs essentiels de l'évolution. La découverte récente de mécanismes permettant la naissance de la coopération dans un contexte de gènes égoïstes remet cent ans de débats sociologiques sur la table de travail.

L'idée de sélection naturelle en biologie est venue à Darwin dans le cadre culturel de l'Angleterre victorienne, au travers des écrits de l'économiste Malthus à propos de la pression de la population sur des ressources naturelles limitées [Gowdy et Seidl, 2004]. Les conclusions méticuleuses de Darwin furent ensuite popularisées, notamment par Herbert Spencer, comme « la survie du plus fort » [*ibid.*]. Les termes de compétitivité, de survie du plus fort, etc., ont donné ce que l'on appelle le darwinisme social<sup>1</sup>, repris par la théorie néoclassique jusque dans ses applications

---

1. Le darwinisme social est la transposition des théories darwiniennes de l'évolution dans le cadre des sociétés humaines, en particulier la compétition entre groupes, entreprises, etc.

politiques récentes<sup>2</sup>. Darwin n'était pourtant pas si catégorique, et il reconnaissait déjà par exemple le rôle de la sélection de groupe [Darwin, 1871 ; Corning, 1997].

Le darwinisme social s'est profondément enraciné dans nos esprits, et les premiers anthropologues étaient évolutionnistes, au même titre que les successeurs d'Adam Smith, de Karl Marx et des autres économistes classiques. Aujourd'hui, on parle encore de sociétés « développées » par opposition aux sociétés appelées pudiquement « en voie de développement » ou « pays les moins avancés », pour ne plus utiliser le terme de « sociétés primitives ». Vue sous cette perspective, la notion même de « développement durable » peut apparaître comme imprégnée de ce darwinisme social à cause du mot « développement ». La perspective néoclassique se base également sur ce modèle de sélection du plus compétitif dans un contexte de concurrence. On peut donc avancer que le darwinisme social est un trait marquant de la culture moderne.

L'analyse scientifique ne permet pas de classer hiérarchiquement une espèce biologique ou une société par rapport à une autre en termes de « supériorité », mais seulement en termes de succession d'apparition ou de disparition dans le temps, de degré de complexité, etc. Encore faudrait-il postuler que les espèces les plus « vieilles », popularisées sous le nom de « fossiles vivants » – tels les crocodiles –, ont cessé d'évoluer depuis leur apparition, ce qui semble fort éloigné de la réalité. Les êtres les plus simples, comme les virus, n'ont pu apparaître que tardivement dans l'échelle du temps, étant une forme dégénérée d'organismes plus complexes. Quelque 14 millions d'espèces [PNUE, 2002] existent aujourd'hui. Elles sont toutes le résultat d'une évolution de plusieurs milliards d'années. 24 % des espèces de mammifères et 14 % des espèces d'oiseaux sont aujourd'hui menacées de disparition [*ibid.*], ce qui est un signe d'inadaptation au monde *moderne*. On ne peut pas conclure pour autant sans anthropomorphisme qu'elles sont « inférieures » au virus de la grippe le plus basique, qui lui se porte très bien. Si la sélection naturelle se limitait à la survie du plus fort, il

---

2. En Belgique, l'objectif du ministère des Affaires économiques est aujourd'hui de « favoriser le fonctionnement compétitif et durable du marché des biens et services » [Verjus, 2006]. Les objectifs de Lisbonne pour l'Union européenne se basent sur le même leitmotiv.

n'y aurait plus aujourd'hui, dans le meilleur des cas, qu'une seule espèce, et cette espèce ne comporterait qu'une super-race, parmi laquelle une superfamille hyper consanguine. Dans le domaine économique, au lieu des 700 000 entreprises comptabilisées en Belgique actuellement [INS, 2006], il ne resterait qu'une super-entreprise qui mettrait bien à mal l'hypothèse d'infinité des acteurs de la théorie néoclassique.

On voit que, dans le domaine social comme dans le domaine biologique, le concept de « sélection du plus fort » ne reflète pas fidèlement la théorie darwinienne et n'explique pas la diversité observée. La théorie de l'évolution de Darwin explique l'un des grands moteurs de la création d'espèces, sans plus. Outre l'absence d'évolution définitive vers une super-espèce ou un monopole qui écraserait tout, il reste cependant à expliquer pourquoi des comportements altruistes ont pu résister à la sélection naturelle chez différentes espèces. En effet, si la théorie (utilitariste) de l'évolution est valable, les comportements non égoïstes devraient en première analyse avoir disparu depuis longtemps.

Les principales explications biologiques de la coopération ou de l'altruisme<sup>3</sup> qui ont été jusqu'ici avancées sont les suivantes<sup>4</sup> :

- la sélection de parentèle (*kin selection*) [Hamilton, 1964] ;
- le signalement coûteux ou « barbe verte » [Dawkins, 1976] ;
- l'altruisme réciproque [Trivers, 1971 ; Axelrod et Hamilton, 1981 ; Axelrod, 1984] ;
- la réciprocité indirecte (ou le phénomène de réputation) [Nowak et Sigmund, 1998 ; Leimar et Hammerstein, 2001 ; Milinski, Semmann *et alii*, 2002] et son signalement coûteux [Gintis, Smith *et alii*, 2001] ;
- la réciprocité dure [Fehr, Fischbacher *et alii*, 2002] ;
- finitude, milieu visqueux, flèche du temps [Nowak et May, 1992 ; Nowak, Asaki *et alii*, 2004] ;
- l'effet des fluctuations [Sanchez et Cuesta, 2005] ;
- la non-viabilité de la concurrence parfaite ;
- la sélection de groupe [Darwin, 1871 ; Wynne-Edwards, 1962 ; Henrich, 2004] ;

---

3. L'altruisme est défini comme la capacité d'accomplir des actes coûteux qui donnent un bénéfice à un tiers [Fehr et Fischbacher, 2003].

4. Voir M. A. Novak et K. Sigmund [1968], A. Sanchez et J. A. Cuesta [2005].

— et enfin la sélection culturelle ou la coévolution groupe-individu [Darwin, 1871 ; Bowles, Choi *et alii*, 2003 ; Boyd et Richerson, 2003 ; Hammerstein, 2003 ; Henrich, 2004].

Ces explications se complètent à peu près dans l'ordre. La première explication laisse des comportements altruistes inexpliqués qu'arrive à expliquer la deuxième théorie, etc.

Pour un détail de ces théories, il faudra se rapporter à la littérature la plus récente, par exemple à J. Henrich [2004] ou E. Fehr et U. Fischbacher [2003]. Mais, pour résumer, disons que, alors que pendant longtemps la biologie semblait avoir prouvé que l'altruisme était impossible en termes évolutifs, il est maintenant démontré que, dans des modèles plus complexes faisant intervenir par exemple une *viscosité du milieu* (c'est-à-dire une série d'obstacles favorisant la création de groupes en couches multiples, l'inverse d'un substrat uniforme et instantané) ou l'irréversibilité du temps, des comportements (ou mutations) altruistes peuvent non seulement apparaître, mais en outre se développer et envahir un groupe d'égoïstes primaires [Van Baalen et Rand, 1998]. L'application de modèles probabilistes plutôt que déterministes arrive au même résultat, même sans interactions répétées, de phénomènes de réputation ou de conditions d'hétérogénéité du milieu [Sanchez et Cuesta 2005].

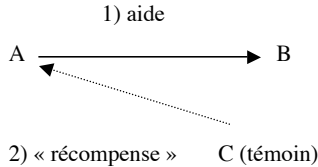
On peut faire le rapprochement avec l'apparition de la vie et de l'ordre dans l'entropie croissante de l'univers qui devient possible loin de l'équilibre en univers visqueux [Prigogine et Stengers, 1979]. De même, dès que l'on décrit un modèle plus proche de la réalité, c'est-à-dire en substrat visqueux (pas de fluidité ni d'instantanéité des échanges, ni de concurrence « parfaite »), dans un monde où la tendance naturelle va vers l'égoïsme primaire (un égoïste peut toujours envahir un monde d'altruistes et se développer à ses dépens), une « mutation altruiste » peut se maintenir et se développer. Empiriquement, il faut bien admettre que la sexualité par exemple n'a pu se développer chez des êtres primitifs que moyennant un minimum de capacités de coopération.

Dès qu'il s'agit de comportements humains, la naissance de la culture et du langage – issus des nécessités de discriminer entre altruistes et égoïstes pour des stratégies de coopération toujours risquées – a ajouté un facteur explosif dans la sélection de comportements de coopération. Dès la naissance de l'humanité, il y eut coévolution nature-culture [Henrich, 2004].

Je vais illustrer ce propos par les théories les plus récentes qui essaient d'expliquer les comportements altruistes non explicables par un égoïsme détourné (parentèle, altruisme réciproque) – par exemple le sacrifice personnel, la volonté de punir ou de récompenser, le don anonyme, etc.

### La réciprocité indirecte

La réciprocité indirecte<sup>5</sup>, ou réciprocité ternaire [Temple, 1997], est, d'un point de vue utilitariste, perçue comme le prolongement un peu plus complexe de la notion d'échange différé. Ce comportement se base sur la notion de réputation : A aide B au vu et au su de C., C est alors enclin à aider A. Il peut s'expliquer parce que C fait confiance à A dans l'éventualité d'une relation de réciprocité directe (explication utilitariste). Ayant pu observer que A est capable de relation réciproque, qu'il a internalisé la norme sociale, C est prêt à « investir » en aidant A, espérant recevoir de A en retour lors d'une prochaine occasion.



Une autre interprétation du comportement de C est qu'il veut récompenser A pour sa générosité. Cette dernière explication est non utilitariste. Pourquoi C voudrait-il récompenser A d'avoir aidé un tiers, B ? Peut-être espère-t-il ainsi renforcer la norme sociale, qui *in fine* lui reviendrait positivement ? Mais, dans ce cas, pourquoi C ferait-il un effort pour récompenser A ? Pourquoi ne laisserait-il pas quelqu'un d'autre le faire à sa place ?

La motivation de C peut donc s'expliquer de deux manières, l'une utilitariste, l'autre non. Mais il est plus intéressant d'analyser les motivations de A. Soit A est altruiste et il n'y a plus rien à

---

5. Voir R. D. Alexander [1987], M. A. Nowak et K. Sigmund [1998], O. Leimar et P. Hammerstein [2001], M. Milinski, D. Semmann *et alii* [2002], M. A. Nowak et K. Sigmund [2005],

démontrer. Soit A est utilitariste et aide B par un calcul intéressé subtil. Que C soit ou non utilitariste importe peu : en se faisant une réputation, A *escompte un bénéfice futur* éventuellement plus important que son investissement initial. Cette stratégie n'a de sens que si, dans le même temps, A *fait un pari de non-utilitarisme de la part des autres agents*. Autrement dit, si A se montre généreux par calcul utilitariste, c'est qu'il ne croit pas en la pureté utilitariste des autres agents. Aussi calculateur soit-il, il ne peut se résigner à croire que les autres sont exclusivement calculateurs.

Donc, à ce point du raisonnement, soit A veut simplement aider B « gratuitement », sans motivation utilitariste, soit il est utilitariste, mais il parie sur le non-utilitarisme des autres agents comme C. Il reste une troisième voie, qui permet de sortir de la tension émanant de cette dernière proposition : c'est que *les deux motivations se conjuguent*. A comme C mêlent des motivations altruistes et utilitaristes dans des proportions variées. Ce dernier modèle est le seul à s'accorder à tous les comportements observés. La nuance est que l'on peut universellement observer des comportements purement égoïstes, alors que l'on n'a pas pu prouver des comportements purement altruistes [Fehr, Fischbacher *et alii*, 2002].

Laissons là pour l'instant ce débat et posons-nous la question des conditions nécessaires pour que la réciprocité indirecte soit possible d'un point de vue utilitariste.

La réciprocité indirecte nécessite un signalement coûteux [Gintis, Smith *et alii*, 2001]. Inutile d'aider « égoïstement » si personne ne s'en aperçoit !

D'autre part, un comportement altruiste n'entraîne pas seulement de la réputation, mais également des jugements moraux et des interactions sociales complexes.

Se préoccuper des générations futures ferait partie de cette recherche de réputation, pour peu que l'on arrive à le faire savoir. Mieux, un être parfaitement utilitariste et intelligent pourrait trouver moins coûteux de faire en sorte d'*avoir l'air* altruiste tout en poursuivant subtilement ses intérêts personnels. Attitude d'autant plus complexe que les interlocuteurs punissent l'hypocrisie, etc. Celui-là tirerait le meilleur bénéfice personnel dans un monde dominé par des interactions sociales.

Mais le signalement coûteux ne suffit pas. Outre le signalement coûteux, la réciprocité indirecte nécessite aussi des *discriminateurs*, c'est-à-dire des individus C capables de discriminer entre individus égoïstes et individus (ou comportements) altruistes [Nowak et Sigmund, 2005]. Inutile en effet de faire savoir qu'on est altruiste si personne ne peut s'en rendre compte. M. A. Nowak et K. Sigmund ont modélisé les conditions dans lesquelles des coopérateurs purs, des discriminateurs<sup>6</sup> et des défecteurs purs (égoïstes) pouvaient évoluer dans un modèle évolutif à plusieurs générations. Ils ont montré qu'en l'absence de discriminateurs, les défecteurs gagnent sur les coopérateurs (le comportement coopérateur disparaît<sup>7</sup>). Autrement dit, sans agents capables d'adapter leur comportement au comportement des autres acteurs, le comportement « altruiste » est voué à disparaître, et avec lui toute possibilité de coopération. Le rôle des discriminateurs se complique par le jeu des leurres potentiels (« barbe verte ») envoyés par les défecteurs « intelligents » évoqués dans le paragraphe précédent.

En situation expérimentale avec des humains, il apparaît que la réputation dépend de l'interprétation que donnent les autres non seulement à nos actes, mais aussi à nos *intentions* [Nowak et Sigmund, 2005]. Il ne suffit pas d'effectuer un acte de générosité, il faut qu'il soit perçu comme tel. De même, un don n'est pas perçu comme don si l'intention n'est pas perçue comme désintéressée [Godbout, 1992 ; Temple et Chabal, 1995]. La force qui oblige à recevoir le don, mais aussi à le reproduire (éventuellement vers un tiers) [MauSS, 1924] ne vient que du caractère reconnu désintéressé du don. Un don reconnu comme intéressé s'apparente à l'échange et n'entre pas dans la dynamique de la réciprocité au sens social du terme<sup>8</sup>.

De plus, comme l'observait déjà Aristote [-350], le *statut* peut également intervenir dans des échanges « inégaux<sup>9</sup> ».

---

6. Agents arrêtant de coopérer avec des agents égoïstes.

7. En l'absence de coopérateurs, les discriminateurs et les égoïstes forment un système bistable dépendant des conditions initiales : soit l'un soit l'autre type finit par dominer. En l'absence de défecteurs, les coopérateurs et les discriminateurs forment un équilibre stable.

8. Pour l'étude de la réciprocité « inverse de l'échange », voir *La Revue du MAUSS* ou D. Temple et M. Chabal [1995].

9. Le statut dont parle Aristote est plus complexe qu'une simple réputation puisqu'il fait intervenir des caractéristiques de naissance (nobles, esclaves...). Aristote considère comme naturel qu'un échange soit pondéré par le statut de chacun des



Il faut en effet distinguer entre défection justifiée ou non [Nowak et Sigmund, 2005]. L'analyse des motivations fait intervenir des notions de *justice* et des informations historiques très complexes. Il est parfaitement justifié de refuser de coopérer avec un être réputé abject. Il faut dès lors savoir si la personne qui vante ou dénigre une autre personne est elle-même fiable, etc.

Si des joueurs ont une perception différente de la réputation d'autrui, les erreurs de perception peuvent miner la coopération [Takahashi et Mashima, 2003]. À l'opposé, dans un contexte où l'information est de plus en plus incertaine, le fait que les individus ne soient pas parfaits peut au contraire favoriser la coopération [Lotem, Fishman *et alii*, 1999]. Les gens ont en effet tendance à sous-estimer la perception d'autrui à leur égard (notamment après un acte d'égoïsme) et à collaborer plus que de raison [Nowak et Sigmund, 2005]. Mais de manière générale, plus d'information sur les autres acteurs mène à plus de coopération [Bolton, Katok *et alii*, 2005].

Concluons par la constatation que le schéma de la réciprocité indirecte est très complexe et met à l'épreuve les capacités de l'*Homo œconomicus* calculateur *et rationnel* cher aux utilitaristes. Si l'homme est calculateur, il doit aussi être extrêmement intelligent, hyper informé et hyper rationnel pour arriver à mettre en œuvre des comportements de réciprocité indirecte.

Il semble bien plus complexe de faire des calculs et des paris en chaîne que d'appliquer « bêtement » un principe de générosité désintéressée. « L'instinct à faire du commerce » observé déjà par Adam Smith chez l'être humain entraînerait une demande cognitive telle qu'elle pourrait être liée au développement exceptionnel de l'intelligence et du langage humains [Nowak et Sigmund, 2005]. Pour D. Temple et M. Chabal, la réciprocité au sens large (intéressée ou non) est à la naissance des valeurs humaines [Temple et Chabal, 1995 ; Temple, 1996-2004]. On en revient aux jugements moraux, à l'importance de la norme sociale et à leurs fondements. Le cœur du débat sur l'utilitarisme réside dans cette question : l'éthique est-elle utilitariste ?

---

protagonistes, récusant de la sorte toute idée d'une valeur intrinsèque (neutre) de l'objet échangé.

### La réciprocité indirecte en chaîne

A aide B qui aide à son tour C. Ce cas particulier de réciprocité indirecte ne peut pas s'expliquer uniquement par la réputation<sup>10</sup>. Ce comportement est souvent observé dans les expériences économiques [Nowak et Sigmund, 2005]. Il est également typique du transfert entre générations, des relations parents-enfants [Chabal, 1996] et de la préoccupation pour le développement durable. Cette forme de réciprocité n'a pas d'explication *exclusivement* utilitariste<sup>11</sup>. Elle pourrait à la limite s'expliquer, d'un point de vue utilitariste, par des biais cognitifs (« mauvais réflexes » : on récompense C en croyant récompenser A<sup>12</sup>) semblables à ceux développés par la théorie de la grande erreur psychologique [Nigel, 2004]. Suivant cette conception, la sélection naturelle n'aurait pas favorisé l'altruisme en tant que tel, mais des mécanismes psychologiques sous-jacents qui favorisent par exemple l'empathie<sup>13</sup> [Turner, 2005]. À l'origine de l'humanité, quand l'homme vivait en petits groupes très apparentés, la sélection aurait favorisé la capacité psychologique à se préoccuper de ses propres enfants ou d'autres parents génétiques de la tribu proche, et c'est cette propension qui aurait par la suite agi en faveur d'une collaboration avec des étrangers non apparentés. Or, on constate que l'homme, avec son grand cerveau, n'est pas moins doué de ce point de vue que les primates qui, eux, reconnaissent les membres extérieurs au groupe. Même si nous sommes capables de coopération à grande échelle, nous reconnaissons très bien nos proches, mais plus de manière comportementale que physique [Wolf, 1970 ; Sepher, 1983 ; Daly et Wilson, 1988]. Le favoritisme de parentèle observé dans beaucoup de sociétés est lui aussi avant tout un trait comportemental [Henrich, 2004].

Une autre explication vient d'une perception anthropologique de la réciprocité, telle qu'exprimée dans des sociétés peu occidentalisées.

---

10. On ne s'occuperait de ses enfants que par crainte du jugement d'autrui.

11. Suivant le point de vue utilitariste, l'amour maternel découlerait uniquement du plaisir éprouvé par la mère à la vue de son enfant, de la pression sociale ou d'un investissement pour ses vieux jours, etc.

12. Voir M. A. Novak et K. Sigmund [2005].

13. C'est le point de vue de la psychologie évolutionniste.

Les anthropologues, à la suite de M. Mauss [1924], considèrent le fait social dans sa globalité : il n'y a pas de fait social (par exemple économique) séparé des autres (émotifs, relationnels). Tout fait social, y compris l'échange, est « total ». En gros, nous donnons à nos enfants l'amour (les dons, les frustrations, l'éducation...) que nous avons reçu de nos parents. Il s'agit de la forme de réciprocité dite « chemin » [Racine, 1986], ou réciprocité ternaire unilatérale [Temple, 1997]. Cette globalité des interactions sociales implique que l'on peut recevoir un sourire de l'un et donner un cadeau à l'autre, refuser un cadeau pour marquer son opposition, etc. Tous comportements humains « irrationnels » du point de vue de la rationalité utilitariste et auxquels même la société de marché est accoutumée.

### La réciprocité dure

Nous avons vu que la réciprocité indirecte met en jeu plusieurs caractéristiques, dont le signalement coûteux et la discrimination, soit deux éléments relevant des capacités de communication. Une capacité supplémentaire essentielle dans ce contexte est la faculté de récompenser ou de punir. Sans l'existence de cette faculté dans une société, la réputation n'aurait aucun sens. Chercher « égoïstement » à se faire une bonne réputation n'a de sens que si l'on arrive à en faire la publicité, si d'autres acteurs sont capables de la décrypter, et enfin s'ils ont la capacité ou la volonté de récompenser ou de punir. C'est ici qu'intervient la réciprocité dure [Fehr, Fischbacher *et alii*, 2002 ; Fehr et Gächter, 2002] incluant la punition [Henrich et Boyd, 2001 ; Fehr et Rockenbach, 2004].

Les gens ont une propension à récompenser, même sans avantage individuel direct ou différé, les personnes perçues comme altruistes et à punir « à perte » les comportements déviants [Fehr *et alii*, 2002 ; Sethi et Somanathan, 2003] :

« Une personne est “réciproque dure” si elle est prête à :

- sacrifier des ressources pour être agréable à ceux qui sont agréables (réciprocité positive dure) ;
- sacrifier des ressources pour punir ceux qui sont désagréables (réciprocité négative dure).

La caractéristique essentielle de la réciprocité dure est la volonté de sacrifier des ressources pour récompenser les comportements bienveillants (*fair*) et punir les comportements malveillants (*unfair*) même si c'est coûteux et que cela ne donne lieu à aucun bénéfice matériel présent ou futur. »

Qu'une action soit perçue comme juste ou injuste dépend des conséquences distributionnelles de l'action par rapport à des actions de référence neutres [Kahneman, Knetsch *et alii*, 1986] *et des motivations ou intentions*<sup>14</sup> [Rabin, 1993 ; Falk *et alii*, 2000].

R. Sethi et E. Somanathan [2001] trouvent par modélisation déterministe<sup>15</sup> que des individus caractérisés par un comportement de réciprocité peuvent non seulement survivre dans un groupe d'individus matérialistes (purement égoïstes), mais également se multiplier et devenir majoritaires. Cela a d'autant plus de chances de se produire, même dans le cadre d'interactions sporadiques (où l'on peut donc exclure tout pari sur un bénéfice futur), qu'il existe des individus adoptant une attitude de *réciprocité dure* [Fehr, Fischbacher *et alii*, 2002], c'est-à-dire qu'ils sont prêts à punir « à perte » des individus égoïstes et à récompenser de manière tout aussi désintéressée des individus altruistes. On observe bien entendu dans la réalité, et partout dans le monde, de telles attitudes.

Dans diverses conditions expérimentales appliquées à divers pays du Nord<sup>16</sup>, on observe une proportion tombant rarement au-dessous de 40 % – et pouvant aller jusqu'à 66 % – de sujets manifestant des comportements de réciprocité positive dure, contre 20 à 60 % de comportements purement égoïstes<sup>17</sup>.

Utiliser ou réagir à la punition coûteuse a, tout comme d'autres jugements moraux, une représentation émotive dans des zones

14. Ce dernier point élimine une autre théorie, l'aversion à l'iniquité, où les motivations ne jouent aucun rôle [*ibid.*, p. 4].

15. Leur modèle comportemental se déroule, comme la plupart des modèles de première génération, en milieu non visqueux ou non aléatoire, donc en dehors des conditions de milieu dont nous avons vu en introduction qu'elles peuvent favoriser la coopération.

16. Autriche, Allemagne, Hongrie, Pays-Bas, Suisse, Russie et USA [Fehr, Fischbacher *et alii*, 2002].

17. Voir F. Bolle [1998], K. McCabe, S. J. Rassenti *et alii* [1998], E. Fehr et A. Falk [1999], K. Abbink, B. Irlenbusch *et alii* [2000], G. Charness [2000], E. Fehr et S. Gächter [2000], K. McCabe, M. L. Rigdon *et alii* [2000], S. Gächter et A. Falk [2001].

déterminées du cerveau [Greene et Haidt, 2002 ; de Quervain *et alii*, 2004 ; Singer *et alii*, 2006]. Ces zones sont différentes suivant que le jugement est utilitariste ou émotionnel, comme en cas de conflit éthique [Greene, Nystrom *et alii*, 2004]. Le rôle de la punition dans la coopération pacifique est par ailleurs observé chez les animaux, chez les poissons coralliens par exemple [Wong, Buston *et alii*, 2007].

Le comportement de réciprocité dure a donc probablement des racines évolutives très profondes se marquant en termes de *compétitivité de groupe*. La réciprocité favorise la coopération et la victoire du groupe sur l'environnement ou sur d'autres groupes. En termes individuels, elle est pourtant défavorable, comme reflété dans toutes les conditions expérimentales de la littérature, puisqu'elle est coûteuse pour les deux protagonistes impliqués.

Remarquons que l'altruisme pur (coopération inconditionnelle avec un partenaire anonyme égoïste) est pratiquement inexistant dans les résultats d'expériences [Falk *et alii*, 2000]. De même, la réciprocité a tendance à disparaître face à un partenaire aléatoire non humain [*ibid.*], ce qui montre son caractère éminemment social.

Alors que, d'un point de vue collectif, la réciprocité simple et l'altruisme favorisent la coopération et donc l'efficacité globale, R. Sethi et E. Somanathan [2001] ont montré qu'elle n'empêche pas des égoïstes de profiter individuellement du système<sup>18</sup>. La réciprocité et l'altruisme pur auraient donc dû disparaître par sélection naturelle, mais cela n'a pas été le cas pour l'homme<sup>19</sup>.

C'est la réciprocité dure qui explique le maintien de caractères génétiques favorisant la sociabilité, donc la réciprocité et l'altruisme, chez l'être humain [Fehr, Fischbacher *et alii*, 2002]. Le groupe ou ses individus peuvent en effet décider de récompenser un pur altruiste ou sa famille, par exemple en lui donnant le statut de chef. Ils peuvent aussi « se débarrasser » d'un irréductible égoïste, comme on l'observe dans toutes les sociétés de manière plus ou moins institutionnalisée et plus ou moins violente (du simple refus d'assistance au bannissement, ce qui pouvait signifier,

---

18. De plus, l'égoïsme est toujours payant et sa probabilité d'apparition ou de réapparition est grande : il s'agit de la tendance naturelle en l'absence de pression, de cadre ou de sélection de type « réciprocité dure ».

19. Il serait intéressant de vérifier si la théorie utilitariste s'applique à l'orang-outang, notre cousin non sociable.

dans les temps préhistoriques, la mort). La force de la réciprocité est telle qu'elle peut pousser un individu perçu comme égoïste à sacrifier sa propre vie (donc ses chances de reproduction génétique) pour « racheter » ses fautes [voir Durkheim, 1897]. Des facultés d'agressivité développées dans le cadre de la prédation, voire d'agressivité vis-à-vis de congénères dans le cadre de la « loi du plus fort », peuvent également, par la pression du groupe, être utilisées à l'avantage de la réciprocité dure, comme dans des guerres ou des conflits. Alors, chez les hommes préhistoriques, les réciproques « durs », motivés par les récompenses et les punitions du groupe, avaient le dessus sur les égoïstes agressifs qui, de plus, s'éliminaient entre eux.

E. Fehr et ses collaborateurs [Fehr, Fischbacher *et alii*, 2002] ont mené à bien diverses expériences avec des étudiants en éliminant progressivement tous les biais explicables par les théories exposées en introduction. Ils ont néanmoins trouvé une part significative de comportements de réciprocité dure qui ne peuvent s'expliquer :

- ni par la théorie de la sélection de parentèle, car les sujets des expériences considérées savaient qu'ils étaient étrangers les uns aux autres,

- ni par la théorie de l'altruisme réciproque (un altruisme conditionnel, suspendu à un retour, en fait une forme d'échange différé), car il n'y avait pas de gain futur lié à la coopération manifestée,

- ni par la théorie de la réciprocité indirecte, car les sujets étaient anonymes et ne se voyaient pas, et ne se connaîtraient jamais, d'où l'impossibilité d'un effet sur la réputation individuelle,

- ni par la théorie des signaux coûteux de traits inobservables pour les mêmes raisons.

Le fait que des individus soient prêts à « punir » ou à « récompenser » à perte non proportionnelle pour eux-mêmes conduit à s'interroger non seulement sur leur utilitarisme (au moins sur leur rationalité), mais aussi sur la survie de leurs gènes. D'autant plus que les individus qui mettent en œuvre ces comportements à risque ne sont pas systématiquement les plus vieux, mais au contraire souvent des jeunes dans la force de l'âge et qui n'ont pas encore eu l'occasion de se reproduire. Songeons aux militaires (volontaires ou non) pendant les grandes guerres ou aux terroristes kamikazes. Tous ces jeunes gens sont-ils morts « pour la bonne cause » seulement

parce qu'ils n'avaient pas le choix ou dans l'espoir des récompenses dues aux valeureux guerriers à leur retour ?

Il semble aujourd'hui acquis que le rôle de la réciprocité dure dans la sélection de comportements altruistes et le renforcement des normes sociales est déterminant [Fehr et Gächter, 2000 ; Fehr, Fischbacher *et alii*, 2002 ; Fehr et Gächter, 2002 ; Fehr et Fischbacher, 2004]. Par exemple, si généralement on ne quitte pas un taxi au beau milieu des embouteillages sans le payer (comportement prédit par la théorie utilitariste), c'est au mieux parce que cela ne se fait pas (intérieurisation de la norme sociale), au pire parce que l'on craint que, même pour quelques euros, le chauffeur de taxi ne soit prêt à planter là sa voiture pour poursuivre son client indélicat (pari sur un comportement de réciprocité dure).

Le rôle de la réciprocité dure dans la sélection des comportements réciproques nous montre que la culture (sur un fond de capacité émotionnelle sans doute génétique) est capable de favoriser cette capacité de la coopération à se maintenir et à résister aux « passagers clandestins ».

C'est avec la capacité de « discriminer » entre amis et ennemis, ou entre coopérateurs et calculateurs, un élément essentiel du développement de l'une des caractéristiques de l'être humain : sa sociabilité.

### **Finitude, milieux hétérogènes, flèche du temps**

Les premières démonstrations évolutionnistes avaient « prouvé » l'inévitable utilitarisme primaire des espèces (ou de leurs gènes) en faisant une erreur fondamentale au regard de la réalité du milieu : elles se basaient sur des modèles déterministes étonnamment proches de ce que l'on retrouve dans les hypothèses fondamentales de la théorie néoclassique : milieu homogène, populations infinies, à croissance illimitée, réversibilité du temps (et des équations), etc.

Des modèles plus récents arrivent au contraire à faire « naître et se maintenir » la coopération dans des conditions plus proches de la réalité observée. Par exemple, une modélisation dans une population finie plutôt qu'infinie permet d'étudier les conditions (taille de la population, taux de mutation) nécessaires pour l'émergence et la

stabilité évolutive de la coopération [Nowak, Asaki *et alii*, 2004]. Introduire une viscosité du milieu (milieu hétérogène, comportant des parois ou zones de repli de groupes de coopérateurs, plutôt que les milieux homogènes dans lesquels les modèles de première génération favorisent l'égoïsme primaire) fait apparaître des résultats différents : dans le dilemme du prisonnier<sup>20</sup> itératif par exemple, si la stratégie d'un individu est de commencer par coopérer au premier tour, puis d'adopter systématiquement l'attitude de son adversaire au deuxième tour (donnant-donnant), le système évolue vers la survie des seuls égoïstes en cas de milieu homogène (le premier geste de coopération ayant engendré un désavantage initial), alors qu'en milieu visqueux cette stratégie arrive à s'imposer et à maintenir l'équilibre avec des stratégies plus égoïstes [Nowak et Sigmund, 1992]. Remarquons que, dans ce modèle, il n'est nullement question d'altruisme mais bien de coopération « intéressée ». Cependant, la possibilité d'émergence de celle-ci pave la route vers des stratégies plus généreuses [Nowak et Sigmund, 1992]. Chez les bactéries, le substrat visqueux (*versus* fluide) joue un rôle favorable au développement de capacités altruistes comme la production (à perte) de toxines contre les bactéries non résistantes<sup>21</sup> [Chao et Levin, 1981]. Plus fondamentalement, la transition entre organismes unicellulaires et organismes pluricellulaires indifférenciés pourrait résulter de formes élémentaires de coopération dans le métabolisme de l'ATP<sup>22</sup> [Pfeiffer, Schuster *et alii*, 2001].

---

20. Les gains se répartissent par exemple comme suit : 3 points pour chacun si les deux protagonistes coopèrent, 1 point pour chacun si aucun ne coopère, 0 point pour le coopérateur si l'autre ne coopère pas, lequel reçoit alors 5 points. Il est admis dans tous ces modèles que, pour que la coopération se maintienne, il faut qu'elle soit au minimum plus intéressante pour la société que l'égoïsme primaire, lequel est en revanche toujours plus intéressant pour un individu égoïste dans un monde altruiste [Axelrod, 1984].

21. La production de toxines se fait au détriment de la vitesse de reproduction et n'est favorable que si la colonie de clones résistants reste géographiquement soudée, par exemple sur du agar-agar. En solution liquide, cette caractéristique est rapidement supplantée par les bactéries qui utilisent toutes leurs ressources à la simple reproduction et colonisent donc plus vite le milieu.

22. Une production lente d'ATP à haut rendement énergétique, la respiration, est respectueuse du « bien commun » et mène au développement durable optimal pour la vie. Cependant, la fermentation, moins rentable car plus extensible, mais beaucoup plus rapide, est la solution la plus compétitive en milieu fluide, alors qu'elle est la moins soutenable.



Le fait de placer les « joueurs » des modèles de dilemme du prisonnier dans un espace (par exemple bidimensionnel), donc de mettre une contrainte environnementale (bien réaliste au regard de la nature), permet de se passer des contraintes de « mémoire » et de faire évoluer des joueurs à la stratégie bien plus simple : toujours coopérer ou toujours agir égoïstement. C'est la situation qui pouvait prévaloir juste après que le premier mutant coopérateur soit apparu dans un monde de sélection naturelle impitoyable. Dans des conditions spatiales et sous certaines conditions de répartition des gains ou de chance de se répandre au détriment du voisin moins performant, les modèles font apparaître et se maintenir la coopération même au détriment des agents purement égoïstes. Suivant les conditions, et quelle que soit la situation initiale, le système entre dans une dynamique continue de type fractal (des zones de coopérateurs et des zones d'utilitaristes s'entrecroisent) qui finit par évoluer vers un équilibre statistique stable en termes de proportion entre coopérateurs et égoïstes purs [Nowak et May, 1992]. Bref, le simple fait de faire tourner des modèles moyennant au moins un biais par rapport au monde parfaitement homogène et réversible dans l'espace et dans le temps des premiers modèles évolutionnistes suffit pour rendre possible et viable l'apparition de la coopération dans l'évolution.

### L'effet des fluctuations

Les modèles faisant intervenir des fluctuations aléatoires (bruit) – sous forme de mutations dans un sens ou dans l'autre, par exemple – donnent des résultats tout à fait étonnants. La seule apparition (ou disparition) aléatoire d'un comportement phénotypique (quel que soit le gène sous-jacent) allant dans le sens de la « punition dure » peut mener à l'installation durable d'altruistes dans un groupe initialement constitué à 100 % d'égoïstes purs [Sanchez et Cuesta, 2005]. Il est bien entendu que, dans les modèles incriminés, le comportement altruiste (réciprocité dure, voire simple punition égoïste<sup>23</sup>) est systématiquement défavorisé par rapport au comportement égoïste. Le modèle proposé par Sanchez et Cuesta a

---

23. Pour le rôle de la punition égoïste dans différents modèles évolutifs, voir notamment M. Nakamaru et Y. Iwasa [2006].

l'esthétique de la simplicité (la simple prise en compte de fluctuations aléatoires, quelle que soit la situation initiale et sans conditions fortes du milieu). Il mène à une étonnante stabilité dans les résultats, une fois les paramètres (taux de mutation, létalité) choisis dans des ordres de grandeur proches de ceux observés dans la nature. La situation d'équilibre arrive à des ordres de grandeur (proportion entre survivants égoïstes et altruistes) semblables à ceux qui prévalent dans les observations humaines<sup>24</sup>. Ce modèle ne fait en outre intervenir aucune notion de sélection de groupe (telle que décrite plus loin) et se base strictement sur la sélection individuelle. De plus, il opère sur des agents sans mémoire ou capacité d'apprentissage. Par exemple, les premiers « réciproqueurs » peuvent apparaître, « punir » et disparaître dès la génération suivante : la répétition de ces « accidents » peut suffire à induire une sélection d'ensemble finissant par faire apparaître et se maintenir une population de réciproqueurs dans un milieu initialement « égoïste élémentaire ». Autrement dit, ce modèle permet d'expliquer l'apparition purement probabiliste des formes les plus élémentaires de coopération dans une population finie nourrie d'énergie extérieure, au même titre et suivant la même dynamique que la vie ou l'ordre ont pu surgir du chaos [Prigogine et Stengers, 1979].

### La sélection de groupe

La théorie de la sélection de groupe [Darwin, 1871 ; Wynne-Edwards, 1962] part du constat qu'il existe des espèces sociales. La sélection de groupe favorise des caractéristiques qui augmentent la force d'un groupe comparé à d'autres groupes [Wolson, 1997]. On suppose que l'homme par exemple était à l'origine organisé en petits groupes. Des mutations seraient intervenues pour rendre des individus coopératifs au sein d'un groupe, ce qui aurait conféré à ce groupe un avantage sélectif par rapport à un autre groupe incapable de coopérer.

Quoique théoriquement possible, la sélection de groupe telle qu'exposée par les premiers auteurs résiste peu à l'analyse empiri-

---

24. Voir J. Henrich *et alii* [2001], E. Fehr et U. Fischbacher [2003], H. Gintis, S. Bowles *et alii* [2003].

que [Williams, 1966]. Le problème de la sélection de groupe est la survie du premier mutant coopérateur. En première analyse, comme celui-ci naît dans un monde égoïste, il n'a aucune chance de survie. En admettant qu'un groupe de coopérateurs arrive à se constituer, celui-ci disparaîtrait à l'arrivée du premier « passager clandestin » égoïste<sup>25</sup>. La naissance d'un groupe de coopérateurs semble alors improbable. De plus, la sélection de groupe suppose qu'il n'y a pas de migrants entre groupes. Or on connaît les difficultés de la consanguinité chez l'homme. Seuls quelques cas plausibles de sélection de groupe ont pu être décrits [Sober et Wilson, 1998].

Après avoir sombré dans l'oubli suite aux premières explications génétiques (sélection de parentèle, altruisme réciproque, etc.), cette notion refait aujourd'hui surface depuis que l'on a pu montrer qu'un mutant altruiste avait des chances de survie dans un monde égoïste et que des modèles de sélection plus élaborés que la simple addition de phénomènes distincts ont été étudiés.

Ce qui a faussé le raisonnement des opposants au principe de la sélection de groupe, c'est d'avoir exclu la possibilité d'une évolution conjointe de la sélection intra et intergroupe [Henrich, 2004]. Une sélection de groupe peut en effet se diviser en deux composantes : une composante intragroupe et une composante intergroupe [Price, 1970, 1972 ; Hamilton, 1975 ; Henrich, 2004]. Si l'on analyse le problème séparément, on observe que la sélection intragroupe opère toujours à l'encontre des altruistes. Mais dès que l'on s'accorde une possibilité de sélection multiniveau, J. Henrich [2004] a montré mathématiquement à la suite de Price (dans le cadre du dilemme du prisonnier répétitif) qu'il était possible de voir se maintenir un gène de coopération, permettant au caractère de se répandre contre les groupes égoïstes et de résister aux passagers clandestins en provenance d'autres groupes. Son modèle devient même explosif dès qu'il y fait entrer la sélection culturelle chez l'être humain (voir plus loin le modèle de coévolution). G. Price et J. Henrich ont démontré de manière purement mathématique et théorique ce que d'autres ont expérimenté dans des modèles évolutifs.

---

25. Rappelons qu'il faut à la fois « barbe verte », discrimination, réciprocity dure et capacité d'agressivité intraspécifique : l'apparition de la coopération dans un monde dominé par l'égoïsme primaire des gènes semble *a priori* aussi improbable que l'apparition de la vie dans un univers d'entropie croissante.

### L'instabilité de la concurrence parfaite

Peu d'auteurs accordent de l'importance à la loi du plus fort elle-même. Les modèles étudiés jusqu'à présent font l'hypothèse *implicite* qu'il existe déjà une coexistence pacifique entre individus, voire des groupes déjà constitués avant même d'envisager une « mutation » altruiste. L'égoïsme est donné au départ. Or, pour qu'il existe un groupe, il faut déjà soit des individus coopérateurs, soit des individus indifférents (c'est-à-dire sans interaction avec les autres individus<sup>26</sup>). En particulier, toute espèce ayant à s'occuper de sa descendance pour sa survie, comme chez les mammifères, a dû nécessairement développer une forme de coopération « de parentèle » pour permettre à ses petits de survivre. Un carnivore parfaitement égoïste a un intérêt immédiat à manger son rejeton, une poule ses œufs, etc. Cette stratégie est évidemment vouée à l'échec en termes évolutionnistes. L'altruisme de parentèle ou la sexualité impliquent des capacités élémentaires de coopération.

Prenons deux organismes simples hétérotrophes génétiquement proches dans un univers infini, par exemple deux brins d'ADN dans un milieu nutritif<sup>27</sup>. L'un est très efficace dans la mesure où il se multiplie de manière intensive, avec un fort rendement de reproduction *par unité de nourriture* avalée<sup>28</sup>, l'autre a une stratégie extensive, gaspilleuse, mais rapide car peu coûteuse, avec un grand rendement de reproduction *par unité de temps*<sup>29</sup>. En début de colonisation, la nourriture est abondante et chacun se multiplie à l'envi. Celui des deux qui a la plus grande vitesse de reproduction va rapidement envahir le milieu et gagner la course évolutionniste à court terme. Ce n'est pas le plus efficace qui va gagner, mais le plus rapide, que l'on qualifiera *arbitrairement* comme le plus égoïste.

---

26. À la rigueur, on pourrait considérer les individus indifférents comme des « égoïstes mous » puisqu'ils ne s'occupent que d'eux-mêmes sans se préoccuper de « concurrence ».

27. Passons sur le fait que, pour qu'il y ait ADN, il faut déjà que la « coopération élémentaire » entre acides aminés soit plus grande que leur « concurrence élémentaire » pour l'appropriation des nutriments (azote, carbone, etc.).

28. On peut déjà considérer que l'utilisation efficiente des richesses est une forme d'altruisme élémentaire puisque cette stratégie permet un optimum de groupe.

29. Remarquons que le choix du dénominateur change la perspective. Le plus rapide peut également être considéré comme « altruiste élémentaire » dans la mesure où il favorise son espèce au détriment de l'intérêt de la concurrence.

Le plus efficace va rapidement se trouver spatialement entouré par le plus rapide, ce qui le mettra en condition de milieu confiné ou « visqueux » et il subira le premier la carence alimentaire et éventuellement la disparition. La sélection naturelle a ici favorisé le moins performant, mais ce n'est qu'un accident de l'histoire.

Le plus rapide étant le plus nombreux, il a en cours de réplication une plus grande probabilité de voir apparaître une mutation.

Imaginons donc que chez ce gagnant apparaît une mutation capable de le faire « entrer en relation » avec un autre individu. Prenons le cas d'une mutation agressive : la capacité de se nourrir de son voisin (ou la production d'une substance l'empêchant de se multiplier, ou toute autre mutation agressive comme la capacité de gagner un jeu comme le dilemme du prisonnier, le jeu de l'ultimatum, le jeu du dictateur, etc.). La littérature souligne que « par définition » le mutant « égoïste » envahirait rapidement le milieu et éliminerait ses congénères. Pour ce faire, il faudrait conjointement une deuxième mutation (donc très peu probable) : celle qui protégerait le mutant contre sa propre agressivité ! En effet, prenons le cas du prédateur. Il s'attaquerait indistinctement à tout organisme à sa portée, à commencer par ses propres répliques. L'organisme mutant « parfait égoïste agressif » mais non protégé contre sa propre forme d'agressivité disparaîtrait *in fine* en milieu visqueux (ses propres descendants étant ses premières victimes), car il perdrait son énergie à se reproduire inutilement sur lui-même alors que ses concurrents se contenteraient de se reproduire à partir de la nourriture disponible sans ce coût supplémentaire. Si sa stratégie était la production d'une substance toxique, il en serait la première victime.

La condition pour l'apparition d'un gène égoïste « dur » est donc une double mutation : une mutation égoïste (agressive) et une mutation d'autodéfense.

Cette double mutation peut être considérée comme une forme élémentaire de coopération de parentèle. Car il y a discrimination dans l'attaque, celle-ci ne s'adressant qu'aux individus indifférents, au bénéfice du « groupe » ainsi constitué.

À partir de ce moment, on peut parler de groupe, mais la coopération a déjà commencé. *L'égoïsme dur n'a pu apparaître que si apparaissait en même temps l'altruisme élémentaire.* Dès l'origine, il est difficile de déterminer avec certitude quel est le mutant *égoïste* ou le mutant *coopérateur*. Ce n'est qu'une question de perspective et

aucun modèle unilatéral parfait n'est capable d'expliquer la réalité observée : égoïsme et coopération sont inséparables à la source.

L'équivalent socio-économique est l'individu (ou l'entreprise) parfaitement égoïste. En l'absence de punition, son intérêt est plutôt le vol que le travail. En cas de concurrence, il peut avoir intérêt à supprimer son concurrent naïf par surprise (donc à coût réduit) pour prendre sa place et occuper une position dominante dont il profitera ensuite à souhait. Sans agressivité directe, il peut se contenter de se jeter le premier sur une ressource et ainsi d'en priver son « concurrent ». Dans ce cas de figure, l'individu parfaitement égoïste ne peut survivre que s'il existe des individus naïfs dont il puisse profiter ou qu'il puisse devancer. Que se passerait-il s'il n'y avait plus d'individus naïfs ? Les individus parfaitement égoïstes devraient s'éliminer entre eux. En effet, un égoïsme sans merci est coûteux à court terme, mais à long terme, c'est la seule stratégie payante face à des concurrents parfaitement égoïstes. Le premier à lancer la compétition profite en principe de sa position initiale dominante pour éliminer ses concurrents s'il n'en a qu'un à la fois. Ce raisonnement ne compte pas sur la réaction de l'autre individu parfaitement égoïste, dont la seule stratégie est de ne pas se laisser faire et d'attaquer le premier.

Le principal concurrent de l'égoïste primaire, c'est l'autre égoïste primaire.

S'il y avait dans l'univers deux égoïstes primaires et des individus indifférents, la seule stratégie payante serait de commencer par éliminer le rival égoïste pour ensuite envahir les individus indifférents. Pour lancer cette compétition, il faudrait consentir un coût qui rendrait aussitôt le survivant des deux rivaux « moins sain » que les individus indifférents. L'égoïsme primaire des agents implique les mêmes caractéristiques que la réciprocité : signalement coûteux (avoir l'air le moins concurrentiel possible, passer inaperçu face à la concurrence), discriminateurs et propension à l'action (« punition » des concurrents, et éventuellement favoritisme des alliés objectifs).

Dans les simulations impliquant des égoïstes et des coopérateurs, il faudrait donc faire entrer comme force primordiale la « punition » infligée par les égoïstes aux autres égoïstes, leurs principaux concurrents. Ce seul facteur, l'autodestruction des « égoïstes durs », donnerait un avantage comparatif aux individus indifférents

ou coopérateurs, *a fortiori* aux coopérateurs durs. C'est effectivement ce que l'on observe dans les modèles faisant intervenir des punisseurs égoïstes [Nakamaru et Iwasa, 2006].

*L'instabilité de la concurrence parfaite* constitue un élément explicatif non négligeable de l'émergence de la coopération.

### La sélection culturelle ou la coévolution groupe-individu

R. Sethi et E. Somanathan [2001] font remarquer une différence entre les caractères transmis verticalement (entre générations) et les caractères transmis horizontalement. Les facteurs culturels ont une forte propension à se répandre horizontalement. Les transmissions culturelles horizontales sont extrêmement rapides comme le montrent les expériences (tout individu adapte rapidement son degré d'égoïsme à celui du groupe) alors que *les transmissions culturelles verticales peuvent mettre plusieurs générations pour envahir une population*<sup>30</sup>.

Il y a donc un fonds culturel et génétique général qui détermine des préférences par défaut (entre égoïsme et réciprocité), avec globalement une tendance plus « réciproque » qu'égoïste, et des adaptations rapides au groupe : les réciproques tirant les égoïstes vers plus de réciprocité et vice versa [*ibid.*]. Une abondante littérature expérimentale et théorique est déjà disponible sur les explications de la coexistence de la réciprocité et de l'égoïsme<sup>31</sup>.

La culture, même dans ses formes les plus élémentaires, peut « doper » l'évolution. On savait que c'est à force d'interventions sur la nature que des êtres humains ont sélectionné des milliers de variétés de plantes et d'animaux domestiques, jusqu'à rationaliser plus récemment cette capacité de manière vertigineuse en utilisant le génie génétique.

Mais cette coévolution nature-culture a eu aussi, dès les origines et « inconsciemment », un impact sur l'évolution de toutes les espèces présentant des prémisses d'organisation sociale. L'universalité

---

30. Cela peut expliquer le décalage culturel des découvertes scientifiques : par exemple, la théorie néoclassique basée sur la compétition darwinienne est toujours diffusée comme culture alors même qu'elle est abandonnée par les chercheurs.

31. Résumée dans Sethi et Somanathan [2001].

de l'interdit de l'inceste est un exemple de coopération génétique imposée par la culture. Les formes élaborées ou non de culture agissent comme la Vie, avec une tendance à se reproduire, à se multiplier et à résister à la force d'entropie ramenant sans cesse la tendance vers des comportements plus « utilitaristes ». Chaque tentative de normalisation utilitariste est dans l'histoire accompagnée de résistances anti-utilitaristes spontanées [voir Polanyi, 1947].

### Conclusion

Après que la philosophie de Descartes a pu éliminer du champ d'investigation scientifique toute explication surnaturelle des phénomènes de groupe, l'argument évolutionniste a permis au siècle dernier d'éliminer toute explication holiste de la coopération. Ne restait que l'utilitarisme comme grande découverte de la science universelle, seule compatible avec les conclusions biologiques de l'époque. Les débats entre psychologues ou sociologues apparaissaient dans ce contexte comme des débats dénués d'intérêt, et les holistes étaient volontiers taxés d'idéalisme par les scientifiques condescendants. Les sciences humaines ont vu naître une frange croissante de « vrais scientifiques » sous la forme d'économètres et autres sociologues du chiffre. En économie, la théorie néoclassique, à la faveur de la chute du Mur de Berlin, devint la seule science économique « sérieuse » comparée aux théories hétérodoxes rapidement soupçonnées d'irréalisme.

La découverte récente de mécanismes permettant la naissance improbable de la coopération dans un contexte de gènes égoïstes remet en scène cent ans de débats sociologiques. Je pense que cette révolution est au moins aussi importante que la révolution née des travaux de Prigogine sur la naissance improbable de la Vie dans des milieux dissipatifs : des instabilités locales peuvent créer de l'ordre dans un contexte où l'entropie augmente [Prigogine et Stengers, 1979].

Dans un contexte où l'évolution naturelle favorise l'égoïsme (les « passagers clandestins »), des formes organisées, coopératives peuvent naître et se reproduire. Pour que cela soit possible, il faut un flux constant d'énergie et d'information (sous forme de complexité). Une fois les éléments en place, comme les capacités



cognitives exceptionnelles de l'être humain sous forme de culture, la création tant d'ordre que de coopération (d'institutions) peut devenir exponentielle : l'homme crée de l'ordre à son gré de manière démesurée sous forme, par exemple, de capital immobilier, de cultures, d'institutions et d'éthiques diverses.

Alors que la science a semblé donner raison à l'utilitarisme pendant plus d'un siècle, il devient aujourd'hui possible, comme nous le disions en introduction à ce texte, de contredire cette hypothèse : l'homme n'est pas *uniquement* égoïste, en plus du fait qu'il n'est pas parfaitement rationnel. Autrement dit, il n'est pas utilitariste et la théorie néoclassique perd l'un de ses fondements les plus irréfutés.

Les récentes découvertes théoriques en biologie évolutionniste, une fois sorties de leur carcan élémentaire, purement génétique et déterministe, pour faire place à une co-influence moins déterministe entre culture et évolution, ont permis de redonner à la sociologie et à l'économie une liberté qu'elles avaient perdue. Un vaste champ d'investigation s'ouvre à nouveau dans la sociologie et l'économie ainsi débridées. La charge de la preuve peut s'inverser [Terestchenko, 2004] : il n'est plus nécessaire de « prouver » qu'une espèce vivante peut être capable d'altruisme. Pour accréditer la théorie utilitariste, il faut désormais prouver que l'homme est *uniquement* égoïste. La preuve est apportée expérimentalement, aujourd'hui, du contraire : si l'homme n'est manifestement pas « uniquement » altruiste, il est néanmoins capable d'une panoplie de comportements et de motivations, dont de l'altruisme et de l'égoïsme, en fonction des contextes dans lesquels il est plongé. Fondamentalement, une part d'égoïsme calculateur reste présente dans la plupart de ses choix, mais des motivations autres, complexes, comme des motivations altruistes authentiques *diverses*, s'y ajoutent le plus souvent, qu'il le veuille ou non.

### Bibliographie

- ABBINK K., B. IRLBUSCH *et alii*, 2000, « The moonlighting game – an experimental study on reciprocity and retribution », *Journal of Economic Behavior and Organization*, n° 42, p. 265-277.
- ALEXANDER R. D., 1987, *The Biology of Moral Systems*, New York, Aldine de Gruyter.

- ARISTOTE, -350, *Éthique à Nicomaque*, extraits et traductions revues dans TEMPLE *et alii*, 1995, *L'Échange chez Aristote*.
- AXELROD R., 1984, *The Evolution of Cooperation*, Basic Books, New York.
- AXELROD R. et W. D. HAMILTON, 1981, « The evolution of cooperation », *Science*, n° 211, p. 1390-1396.
- BOLLE F., 1998, « Rewarding trust : an experimental study », *Theory and Decision*, n° 45, p. 85-100.
- BOLTON G. E., E. KATOK *et alii*, 2005, « Cooperation among strangers with limited information about reputation », *J. Public Econ.*, n° 89, p. 1457-1468.
- BOWLES S., J. K. CHOI *et alii*, 2003, « The co-evolution of individual behaviour and social institutions », *Journal of Theoretical Biology*.
- BOYD R. et P. RICHERSON, 2003, *The Nature of Cultures*, University of Chicago Press, Chicago.
- CHABAL M., 1996, « Quand la réciprocité semble non réciproque... ou la réciprocité cachée », *La Revue du MAUSS semestrielle*, n° 8.
- CHAO L. et B. R. LEVIN, 1981, « Structured habitats and the evolution of anticompetitor toxins in bacteria », *Proceedings of the National Academy of Sciences of the United States of America*, vol. 78, n° 10, p. 6324-6328.
- CHARNESS G., 2000, « Responsibility and effort in an experimental labour market », *Journal of Economic Behavior and Organization*, n° 42, p. 375-384.
- CORNING P., 1997, « Holistic darwinism : synergistic selection and the evolutionary process », *Journal of Social and Evolutionary Systems*, vol. 20, n° 4, p. 636-400.
- DALY M. et M. WILSON, 1988, *Homicide*, Aldine de Gruyter, New York.
- DARWIN C., 1871, *The Descent of Man, and Selection in Relation to Sex*, Murray, Londres.
- DAWKINS R., 1976, *The Selfish Gene*, Oxford University Press, Oxford.
- DE QUERVAIN D., J. F. *et alii*, 2004, « The neural basis of altruistic punishment », *Science*, n° 305, p. 1254-1258.
- DURKHEIM E., 1897, *Le Suicide*, PUF, Paris.
- FALK A., F. E. *et alii*, 2000, *Testing Theories of Fairness – Intentions Matter*, Institute for Empirical Research in Economics, University of Zurich.
- FEHR E. et A. FALK, 1999, « Wage rigidity in a competitive incomplete contract market », *Journal of Political Economy*, n° 107, p. 106-134.
- FEHR E. et U. FISCHBACHER, 2003, « The nature of human altruism », *Nature*, vol. 425, n° 23, p. 785-791.
- 2004, « Third-party punishment and social norms », *Evolution and Human Behavior*, vol. 25, n° 2, p. 63-87.
- FEHR E., U. FISCHBACHER *et alii*, 2002, « Strong reciprocity, human cooperation and the enforcement of social norms », *Human Nature*, n° 13, p. 1-25.
- 2002, « Strong reciprocity, human cooperation and the enforcement of social norms », *Human Nature*, n° 13, p. 1-25.

- FEHR E. et S. GÄCHER, 2000, « Cooperation and punishment in public good experiment », *American Economic Review*, vol. 90, n° 4, p. 980-995.
- 2002, « Altruistic punishment in humans », *Nature*, n° 415, p. 137-140.
- FEHR E. et B. ROCKENBACH, 2004, « Human altruism : economic, neural, and evolutionary perspectives », *Current Opinion in Neurobiology*, vol. 14, n° 6, p. 784-790.
- GÄCHTER S. et A. FALK, 2001, *Reputation and Reciprocity – Consequences for the Labour Relation*, Institute for Empirical Research in Economics.
- GINTIS H., S. BOWLES *et alii*, 2003, « Explaining altruistic behavior in humans », *Evolution and Human Behavior*, vol. 24, n° 3, p. 153-172.
- GINTIS H., E. A. SMITH *et alii*, 2001, « Costly signaling and cooperation », *Journal of Theoretical Biology*, n° 7, p. 1-16.
- GODBOUT J. T., 1992, *L'Esprit du don*, La Découverte/Boréal, Paris et Montréal.
- GOWDY J. et I. SEIDL, 2004, « Economic man and selfish genes : the implications of group selection for economic valuation and policy », *Journal of Socio-Economics*, vol. 33, n° 3, p. 343-358.
- GREENE J. et J. HAIT, 2002, « How, and where, does moral judgment work ? », *Trends in Cognitive Sciences*, vol. 6, n° 12, p. 517-523.
- GREENE J. D., L. E. NYSTROM *et alii*, 2004, « The neural bases of cognitive conflict and control in moral judgment », *Neuron*, vol. 44, n° 2, p. 389-400.
- HAMILTON W. D., 1964, « The genetical evolution of social behaviour », *Journal of Theoretical Biology*, n° 7, p. 1-7.
- 1975, « Innate social aptitudes of man : an approach from evolutionary genetics », *Biosocial Anthropology*, R. Fox, Malaby Press, Londres, p. 133-156.
- HAMMERSTEIN, P., 2003, *Genetic and Cultural Evolution of Cooperation*, MIT Press, Cambridge, MA.
- HENRICH J., 2004, « Cultural group selection, coevolutionary processes and large-scale cooperation », *Journal of Economic Behavior & Organization Evolution and Altruism*, vol. 53, n° 1, p. 3-35.
- HENRICH J. et R. BOYD, 2001, « Why people punish defectors : weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas », *Journal of Theoretical Biology*, n° 208, p. 79-89.
- HENRICH J. *et alii*, 2001, « In search of *Homo economicus* : behavioural experiments in 15 small-scale societies », *American Economic Review*, n° 91, p. 73-78.
- INS, 2006, DBRIS, Bruxelles, SPF Economie, 2006 : *Registre des redevables de l'information statistique*.
- KAHNEMAN D., J. L. KNETSCH *et alii*, 1986, « Fairness as a constraint on profit seeking : entitlement in the market », *The American Economic Review*, vol. 76, n° 4, p. 341-350.

- LEIMAR O. et P. HAMMERSTEIN, 2001, « Evolution of cooperation through indirect reciprocity. Proceedings », *Biological Sciences*, vol. 268, n° 1468, p. 475-753.
- LOTEM A., M. A. FISHMAN *et alii*, 1999, « Evolution of cooperation between individuals », *Nature*, n° 400, p. 226-227.
- MAUSS M., 1924, « Essai sur le don, forme et raison de l'échange dans les sociétés archaïques », *Sociologie et anthropologie*, PUF, Paris.
- MCCABE K., S. J. RASSENTI *et alii*, 1998, « Reciprocity, trust and payoff privacy in extensive form bargaining », *Games and Economic Behavior*, n° 24, p. 10-24.
- MCCABE K., M. L. RIGDON *et alii*, 2000, *Positive Reciprocity and Intentions in Trust Games*, University of Arizona at Tuscon.
- MILINSKI M., D. SEMMANN *et alii*, 2002, « Reputation helps solve the tragedy of the commons », *Nature*, n° 415, p. 424-426.
- NAKAMARU M. et Y. IWASA, 2006, « The coevolution of altruism and punishment : role of the selfish punisher », *Journal of Theoretical Biology*, vol. 240, n° 3, p. 475-488.
- NIGEL B., 2004, *Kindness in a Cruel World : The Evolution of Altruism*, Prometheus Books.
- NOWAK M. A., A. ASAKI *et alii*, 2004, « Emergence of cooperation and evolutionary stability in finite populations », *Nature*, n° 428, p. 646-650.
- NOWAK M. A. et R. M. MAY, 1992, « Evolutionary games and spatial chaos », *Nature*, n° 359, p. 826-829.
- NOWAK M. A. et K. SIGMUND, 1992, « Tit for tat in heterogeneous populations », *Nature*, n° 355, p. 250-253.
- 1998, « The dynamics of indirect reciprocity », *Journal of Theoretical Biology*, vol. 194, n° 4, p. 561-574.
- 2005, « Evolution of indirect reciprocity », *Nature*, n° 437, p. 1291-1298.
- PFEIFFER T., S. SCHUSTER *et alii*, 2001, « Cooperation and competition in the evolution of ATP-producing pathways », *Science*, n° 292, p. 504.
- PNUE, 2002, *L'Avenir de l'environnement mondial 3*, GEO-3-De Boeck, Paris-Bruxelles.
- POLANYI K., 1944, *The Great Transformation* (traduction française : *La Grande Transformation. Aux origines politiques et économiques de notre temps*, Gallimard, Paris.
- PRICE G., 1970, « Selection and covariance », *Nature*, n° 227, p. 520-521.
- 1972, « Extensions of covariance selection mathematics », *Annals of Human Genetics*, n° 35, p. 485-490.
- PRIGOGINE I. et STENGERS I., 1979, *La Nouvelle Alliance*, Gallimard, Paris.
- RABIN M., 1993, « Incorporating fairness into game theory and economics », *American Economic Review*, n° 83, p. 1281-1302.
- RACINE L., 1986, « Les formes élémentaires de la réciprocité », *L'Homme*, 99 (XXVI, 3), p. 97-118.

- SANCHEZ A. et J. A. CUESTA, 2005, « Altruism may arise from individual selection », *Journal of Theoretical Biology*, vol. 235, n° 2, p. 233-240.
- SEPPER J., 1983, *Incest, the Biological View*, Academic Press, New York.
- SETHI R. et E. SOMANATHAN, 2001, « Preference evolution and reciprocity », *Journal of Economic Theory*, vol. 97, n° 2, p. 273-297.
- SINGER T. *et alii*, 2006, « Empathic neural responses are modulated by the perceived fairness of others », *Nature*, n° 439, p. 466-469.
- SOBER E. et D. S. WILSON, 1998, *Unto Others – The Evolution and Psychology of Unselfish Behaviour*, Harvard University Press, Cambridge, MA.
- TAKAHASHI N. et R. MASHIMA, 2003, « The emergence of indirect reciprocity : the standing strategy the answer ? », Hokkaido University, *working paper 29*.
- TEMPLE D., 1997, « Les structures élémentaires de la réciprocité », *La Revue du MAUSS semestrielle*, n° 12.
- 1996-2004, *El nacimiento de la reciprocidad (La naissance de la réciprocité), Teoría de la Reciprocidad, t. II : La economía de reciprocidad*, J. Medina et J. Michaux, La Paz, Bolivie, PADEP-GTZ II, p. 195-202.
- TEMPLE D. et M. CHABAL, 1995, *La Réciprocité et la naissance des valeurs humaines*, L'Harmattan, Paris.
- TERESTCHENKO M., 2004, « Égoïsme ou altruisme ? Laquelle de ces deux hypothèses rend-elle le mieux compte des conduites humaines ? », *La Revue du MAUSS semestrielle*, n° 23, p. 313-333.
- TRIVERS R. L., 1971, « The evolution of reciprocal altruism », *The Quarterly Review of Biology*, n° 46, p. 34-57.
- TURNER D. D., 2005, « Altruism – is it still an anomaly ? », *Trends in Cognitive Sciences*, vol. 9, n° 7, p. 317-318.
- VAN BAALEN M. et D. A. RAND, 1998, « The unit of selection in viscous populations and the evolution of altruism », *Journal of Theoretical Biology*, vol. 193, n° 4, p. 631-648.
- VERJUS L., 2006, *Management Plan*, SPF Economie, Bruxelles.
- WILLIAMS G., 1966, *Adaptation and Natural Selection : A Critique of Some Current Evolutionary Theory*, Princeton University Press, Princeton.
- WOLF A. P., 1970, « Childhood association and sexual attraction : a further test of Westermarck hypothesis », *American Anthropologists*, n° 72, p. 503-515.
- WOLSON E. O., 1997, « Human groups as units of selection », *Science*, n° 276, p. 1816-1817.
- WONG M. Y. L., P. M. BUSTON *et alii*, 2007, « The threat of punishment enforces peaceful cooperation and stabilizes queues in a coral-reef fish », *Proceedings of the Royal Society B : Biological Sciences*, vol. 274, n° 1613, p. 1093-1099.
- WYNNE-EDWARDS V., 1962, *Animal Dispersion in Relation to Social Behaviour*, Hafner Publishing Company, New York.