

Principes de pragmatique formelle du discours

Daniel Vanderveken

Volume 34, numéro 2, automne 2007

URI : <https://id.erudit.org/iderudit/015880ar>

DOI : <https://doi.org/10.7202/015880ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Société de philosophie du Québec

ISSN

0316-2923 (imprimé)

1492-1391 (numérique)

[Découvrir la revue](#)

Résumé de l'article

Pourrait-on enrichir la théorie des actes de langage pour traiter du discours? Wittgenstein et Searle ont signalé des difficultés. Beaucoup de discours n'ont pas de but conversationnel, leur arrière-plan est indéfiniment ouvert, ils contiennent des énonciations dépourvues de pertinence et de félicité, et ainsi de suite. À mes yeux, l'objectif principal de la pragmatique du discours est d'analyser la structure et la dynamique des jeux de langage à but conversationnel. Pareils jeux de langage sont indispensables à tout genre de discours. La logique peut analyser leurs conditions de félicité car leur poursuite obéit à des règles constitutives. Beaucoup d'énonciations ne sont pas littérales ou sérieuses. Les unités de la conversation sont les actes illocutoires tentés, qu'ils soient littéraux, sérieux ou non. Comme Montague, je préconise l'usage de formalismes logiques en pragmatique. J'expliquerai comment il convient de réviser et développer les logiques intensionnelle et illocutoire, la logique des attitudes et de l'action afin de modéliser notre capacité de dialoguer. Je comparerai mon approche à d'autres (Austin, Belnap, Grice, Montague, Searle, Sperber et Wilson, Kamp, Wittgenstein) sur le plan de la méthodologie, des hypothèses et des enjeux.

Citer cet article

Vanderveken, D. (2007). Principes de pragmatique formelle du discours. *Philosophiques*, 34 (2), 229–258. <https://doi.org/10.7202/015880ar>

Principes de pragmatique formelle du discours

DANIEL VANDERVEKEN

(Université du Québec, Trois-Rivières)

www.vanderveken.org

RÉSUMÉ. — Pourrait-on enrichir la théorie des actes de langage pour traiter du discours ? Wittgenstein et Searle ont signalé des difficultés. Beaucoup de discours n'ont pas de but conversationnel, leur arrière-plan est indéfiniment ouvert, ils contiennent des énonciations dépourvues de pertinence et de félicité, et ainsi de suite. À mes yeux, l'objectif principal de la pragmatique du discours est d'analyser la structure et la dynamique des jeux de langage à but conversationnel. Pareils jeux de langage sont indispensables à tout genre de discours. La logique peut analyser leurs conditions de félicité car leur poursuite obéit à des règles constitutives. Beaucoup d'énonciations ne sont pas littérales ou sérieuses. Les unités de la conversation sont les actes illocutoires tentés, qu'ils soient littéraux, sérieux ou non. Comme Montague, je préconise l'usage de formalismes logiques en pragmatique. J'expliquerai comment il convient de réviser et développer les logiques intensionnelle et illocutoire, la logique des attitudes et de l'action afin de modéliser notre capacité de dialoguer. Je comparerai mon approche à d'autres (Austin, Belnap, Grice, Montague, Searle, Sperber et Wilson, Kamp, Wittgenstein) sur le plan de la méthodologie, des hypothèses et des enjeux.

ABSTRACT. — Could we enrich speech-act theory to deal with discourse ? Wittgenstein and Searle pointed out difficulties. Most dialogues lack a conversational purpose, their background is indefinitely open, they contain irrelevant and infelicitous utterances, etc. In my view, the primary aim of discourse pragmatics is to analyze the structure and dynamics of language-games whose type is provided with an internal conversational goal. Such games are indispensable to any kind of discourse. Logic can analyze their felicity-conditions because they are conducted according to systems of constitutive rules. Speakers often speak non-literally or non-seriously. The units of conversation are attempted illocutions whether literal, serious or not. I will show how to construct speaker-meaning from sentence-meaning, conversational background and maxims. Like Montague, I believe that we need the resources of formalisms (*proof, model- and game-theories*) and of logic in pragmatics. I will explain how to further develop *intensional* and *illocutionary* logics, the logic of *attitudes* and of *action* in order to characterize our ability to converse. I will compare my approach to others (Austin, Belnap, Grice, Montague, Searle, Sperber and Wilson, Kamp, Wittgenstein) as regards hypotheses, methodology and issues.

1. Introduction

Dans les sciences du langage, de la pensée et de l'action, on reconnaît à présent le rôle essentiel en toute communication des actes de discours, comme les actes d'énonciation, de référence et de prédication, les actes illocutoires et perlocutoires, ainsi que des attitudes des interlocuteurs comme leurs croyances, désirs et intentions. On a aussi compris l'importance cruciale des jeux de langage que les interlocuteurs entendent poursuivre ensemble, des formes de vie qu'ils

partagent et de leur arrière-plan conversationnel. Selon la théorie des actes de discours, les unités premières de signification et de communication dans l'usage et la compréhension du langage ne sont pas des propositions isolées pourvues de conditions de vérité mais plutôt des actes *illocutoires* pourvus de conditions de félicité (Austin, 1962). *Les actes illocutoires élémentaires* ont une *force* (Frege, 1977) et un *contenu propositionnel*. Les interlocuteurs entendent toujours accomplir et communiquer des illocutions. Jusqu'à présent, la logique, la philosophie et la linguistique ont principalement étudié la capacité des interlocuteurs d'utiliser et de comprendre les seuls énoncés sans prendre beaucoup en considération leur capacité de tenir des discours entiers. Dans le même ordre d'idées, la théorie des actes de langage a surtout étudié les actes illocutoires individuels comme les assertions, les promesses et les requêtes que les interlocuteurs entendent accomplir en utilisant des énoncés à des moments d'énonciation particuliers. Pourtant, les interlocuteurs dialoguent le plus souvent avec d'autres agents. Avant tout, l'usage du langage est une *forme sociale de comportement linguistique*. Il en résulte des conversations orales et des textes écrits.

Peut-on étendre la théorie des actes de langage pour traiter de discours entiers ? Wittgenstein et Searle ont exprimé leur scepticisme. Les interlocuteurs ont souvent des objectifs extralinguistiques, ils sont engagés dans des activités communes, et leur arrière-plan est indéfiniment ouvert. Ils font des énonciations sans félicité ni pertinence sans que cela interrompe pour autant leur dialogue. Ils changent aussi leurs objectifs de façon assez arbitraire. Ainsi l'exigence de pertinence impose relativement peu de contraintes à la structure de maintes conversations. Cependant les langues naturelles sont des langues publiques, et la capacité de dialoguer fait sans nul doute partie de la compétence linguistique. Qui plus est, les protagonistes accomplissent toujours en tout discours (que ce soit un dialogue ou un monologue) des *actes illocutoires capitaux* afin d'atteindre des buts proprement linguistiques correspondant à une direction possible d'ajustement entre les mots et les choses. Ils manifestent des attitudes communes (but expressif), ils décrivent comment les choses sont dans le monde (but descriptif), ils délibèrent comment y agir (but délibératif) et ils font des choses avec leurs mots par déclaration (but déclaratoire). De tels objectifs linguistiques sont purement *conversationnels*. Des interventions discursives comme les échanges de salutations, de bienvenues et de nouvelles, les débats, les entrevues, les consultations, les négociations, les compromis, les prises de position et de décision, les contrats ainsi que les cérémonies d'inauguration, de mariage et de baptême sont des *actes illocutoires conjoints d'un niveau supérieur*, irréductibles à des séquences d'illocutions individuelles instantanées. Plusieurs agents obéissant à des règles constitutives y contribuent chacun à leur tour durant un intervalle de temps. D'un point de vue logique, pareilles interventions possèdent *un type et un thème*, et leur poursuite requiert de la coopération. La *logique du discours*, comme je la conçois, n'entend pas étudier tous les jeux de langage possibles, mais seulement ceux dont l'objectif

propre est linguistique. Il existe un grand nombre de pareils jeux de langage et ils sont indispensables à la poursuite de tout discours. Les interlocuteurs doivent tout d'abord se reconnaître en tant que protagonistes de la conversation et se mettre d'accord sur le genre de conversation en cours. Lorsqu'ils communiquent pour exécuter des tâches complexes, il leur faut tôt ou tard décrire leurs objectifs et délibérer comment procéder pour les atteindre. Le but de cet article est de présenter mon approche théorique en pragmatique du discours et de la comparer aux autres approches en cours quant à la méthodologie, aux hypothèses et à la problématique. Je tiendrai compte des travaux antérieurs en philosophie analytique (sur l'action, les attitudes et les actes de langage), en philosophie sociale (sur la coopération), en logique philosophique (sur le sens, la dénotation, les forces, les attitudes et les actions), en sémantique formelle (en idéographie logique, grammaire universelle et théorie de la représentation du discours), en analyse de la conversation (sur les maximes, l'arrière-plan et les genres conversationnels), en linguistique (sur les marqueurs illocutoires et les verbes d'action) et en psychologie (sur la compréhension mutuelle).

2. Enjeux et objectifs théoriques

À l'instar de Montague (1974), je pense que la pragmatique doit utiliser les ressources des formalismes et de la logique afin d'élaborer une théorie rigoureuse de la *signification* et de l'*usage*. Les langues naturelles peuvent être apprises par des agents humains dont les capacités cognitives sont créatives mais limitées (Chomsky). Les formalismes permettent de construire un meilleur modèle théorique de notre compétence linguistique et du mécanisme de la compréhension. Cependant, il convient de réviser de nombreuses hypothèses de la logique standard des propositions, des attitudes et de l'action afin de bien expliquer les conditions de félicité des illocutions. Les propositions avec les mêmes conditions de vérité ne sont pas les contenus des mêmes attitudes et illocutions, comme elles ne sont pas non plus les sens d'énoncés synonymes. On peut affirmer et croire que Paris est une ville sans pour autant affirmer et croire que c'est une ville et non un érythrocyte. La théorie des actes de langage requiert un critère plus fin d'identité propositionnelle. Elle requiert aussi une explication de l'intentionnalité et de la rationalité des interlocuteurs ainsi que de l'engendrement des différentes espèces d'actes de langage en logique de l'action. Il faut en outre intégrer la logique illocutoire dans l'idéographie de la sémantique formelle afin de clarifier la forme canonique des actes illocutoires exprimés par tous les types syntaxiques d'énoncés.

Bien entendu, les unités premières réelles de la conversation sont les actes illocutoires qui sont tentés plutôt que ceux qui sont exprimés. La signification du locuteur est souvent différente de celle de l'énoncé utilisé. En cas d'ironie, de métaphore ou d'indirection, les locuteurs ne parlent pas littéralement, tout comme ils ne parlent pas sérieusement quand ils jouent dans une pièce de théâtre. Afin de contribuer à la *pragmatique du discours*, la théorie

des actes de langage devrait expliquer pleinement la *signification du locuteur* (tous les actes illocutoires tentés, qu'ils soient littéraires, sérieux ou non) ainsi que la *structure et la dynamique des dialogues* que les interlocuteurs sont aptes à poursuivre en vertu de leur compétence. Pour y parvenir, il faut une véritable théorie des maximes conversationnelles (Grice, 1975) et des faits pertinents à l'arrière-plan conversationnel. Les maximes de Grice sont malheureusement vagues et restreintes ; elles ne s'appliquent qu'aux énonciations assertives dans les discours à but informatif. Il convient de généraliser l'approche de Grice et de mieux expliquer la pertinence que ne l'ont fait Sperber et Wilson (1986) qui négligent à la fois les forces illocutoires et les buts conversationnels. Il convient en outre d'analyser ce que font les interlocuteurs quand ils ne parlent pas sérieusement mais feignent d'accomplir des actes illocutoires, au théâtre comme dans la fiction. Afin de modéliser la dynamique de la poursuite du discours, on devrait enfin exploiter les ressources de la théorie des jeux (von Neumann) et profiter des développements récents en logiques dynamique, dialogique, paracohérente et non monotone ainsi qu'en théories de la représentation du discours et de la décision.

Depuis Montague, de nombreux logiciens et linguistes ont exploité des formalismes comme la théorie de la preuve et des modèles conçus originellement pour les langues artificielles afin d'élaborer une sémantique formelle rigoureuse des langues naturelles. Certains ont indirectement interprété les énoncés ordinaires après les avoir traduits dans une langue objet formelle désambiguïsée. Les idéographies permettent d'exhiber la structure logique profonde commune à toutes les langues naturelles et de définir récursivement la signification du locuteur et la compréhension de l'auditeur. Cependant, les unités premières de la signification et de la communication *ne sont pas* des propositions isolées, comme l'affirme Montague, mais plutôt des actes illocutoires élémentaires comme des assertions, des engagements, des questions, des nominations et des remerciements qui possèdent une *force* en plus d'un *contenu propositionnel*. Les locuteurs qui signifient quelque chose tentent d'accomplir des actes illocutoires. Voilà pourquoi il y a des *marqueurs de force* dans les énoncés élémentaires dont la forme logique est complètement analysée. Le type syntaxique des énoncés élémentaires est leur marqueur de force de base. À chaque fois qu'ils sont utilisés littéralement, les énoncés *déclaratifs* servent à faire des assertions, les énoncés *interrogatifs* à poser des questions, les énoncés *impératifs* à donner des *directives* et les énoncés *optatifs* à exprimer des souhaits. Ainsi, les énoncés avec des clauses synonymes comme « Tu viens m'aider », « Viens-tu m'aider ? », « S'il te plaît, viens m'aider ! » et « Si seulement tu venais m'aider » s'expriment dans les mêmes contextes des actes illocutoires avec le même contenu propositionnel mais avec des forces différentes. De même, des énoncés avec les mêmes marqueurs de force tels que « Viens ! » et « Mange ! » expriment des actes illocutoires avec la même force mais des contenus différents. Les grammairiens ont depuis fort longtemps reconnu les aspects illocutoires de la signification linguistique dans leur classification des types syntaxique

d'énoncés. Grâce à la théorie des actes de langage, on peut dorénavant mieux analyser la signification des marqueurs et verbes illocutoires des langues naturelles. Les énoncés impératifs comme « Allons-y ! », dont le verbe est à la première personne du pluriel, servent à offrir à l'interlocuteur de réaliser une action commune. Des énoncés explicitement performatifs comme « Par la présente je vous promets de l'aide » et « Je vous ordonne de sortir » servent à déclarer plutôt qu'à simplement affirmer que le locuteur effectue au moment de l'énonciation l'acte illocutoire nommé par leur verbe performatif. Comme les déclarations réussies rendent vraies le contenu propositionnel, les énonciations performatives réussies ont la force première de *déclaration* en plus de la force secondaire nommée par leur verbe performatif.

Jusqu'à présent, la logique traditionnelle a surtout étudié les conditions de vérité et les inférences théoriquement valides. Ainsi, la sémantique formelle a principalement analysé des expressions comme les noms propres, les descriptions, les prédicats, les connecteurs et les quantificateurs dont la signification contribue à la détermination des conditions de vérité et elle a tenté de construire la compétence linguistique comme étant la capacité des locuteurs de comprendre les contenus propositionnels. Selon Davidson et Montague, le seul objectif vraiment important de la sémantique est de formuler la *théorie de la vérité*. Pour cette raison, la plupart des applications de la sémantique formelle aux langues naturelles existantes ont été restreintes à l'interprétation de leurs énoncés déclaratifs. Cependant, le langage sert à changer le monde autant qu'à le décrire. Il y a cinq *buts illocutoires* que les locuteurs peuvent atteindre sur les contenus propositionnels qu'ils expriment (Searle, 1979). Le *but assertif* qui leur sert à représenter ce qui se passe dans le monde, le *but engageant* qui leur sert à s'engager eux-mêmes à y faire des actions, le *but directif* qui leur sert à essayer de faire agir les auditeurs, le *but déclaratoire* qui leur sert à changer des choses avec leurs mots, et le *but expressif* qui leur sert à exprimer les attitudes que leur inspirent des faits. Les locuteurs font en parlant des inférences pratiques autant que théoriques. Dans le cas des inférences théoriques ils tirent comme conclusions des actes illocutoires assertifs qui sont vrais ou faux. Quiconque affirme la conjonction de deux propositions affirme chacune d'entre elles. Les conclusions tirées d'inférences pratiques sont des actes illocutoires engageants, directifs ou déclaratoires représentant des actions du locuteur ou de l'auditeur. Quiconque pose une question tente d'obtenir une réponse de l'auditeur. Grâce à la logique illocutoire, la sémantique formelle devient capable d'analyser les marqueurs de force dont la signification contribue à la détermination des conditions de félicité et d'interpréter, sans la moindre réduction *ad hoc* aux énoncés déclaratifs, les autres types d'énoncés impératifs, interrogatifs, performatifs, exclamatoires, subjonctifs et autres, exprimant des actes illocutoires de force non assertive. Dans mon optique, la *compétence linguistique* est inséparable de la *performance*. Elle est avant tout la capacité des interlocuteurs d'accomplir et de comprendre les actes illocutoires.

Par nature, les actes illocutoires sont des *actions intrinsèquement intentionnelles* : leur accomplissement requiert une *tentative* verbale du locuteur. Ils ont de véritables *conditions de succès*. Pour les réussir, les locuteurs doivent exprimer correctement les actes illocutoires tentés en utilisant les mots appropriés et en faisant leur tentative en un contexte adéquat. Un baptême n'est pas réussi quand le prêtre se trompe de nom propre ou de bébé. Les locuteurs lient les contenus propositionnels au monde avec l'intention d'établir une *correspondance* entre les mots et les choses selon une certaine direction d'ajustement. Quiconque fait une assertion entend représenter comment les choses sont dans le monde. Quiconque donne un ordre entend que l'auditeur fasse les choses représentées. Quiconque fait une déclaration entend faire les choses représentées avec ses propres mots. Les énonciations assertives ont la *direction d'ajustement des mots aux choses*, les énonciations engageantes et directives ont la *direction d'ajustement des choses aux mots*, et les déclarations ont la *double direction d'ajustement*. Pareils actes illocutoires ont en outre des *conditions de satisfaction*. Pour qu'un acte illocutoire élémentaire soit satisfait, il ne suffit pas que son contenu propositionnel soit vrai et corresponde à un fait existant dans le monde. La correspondance entre les mots et les choses doit être établie selon la bonne direction d'ajustement. Quiconque obéit à un ordre doit faire ce qui est ordonné à cause de cet ordre. S'il ne le fait pas pour cette raison, il n'obéit pas alors à l'ordre. La notion de *satisfaction* est une généralisation de celle de vérité qui couvre toutes les forces. Ainsi, une *assertion* est *satisfaite* quand elle est *vraie* ; une *promesse* l'est quand elle est *tenue* ; une *requête* quand elle est *accordée* et une *déclaration* quand elle rend *vrai* son contenu propositionnel. La satisfaction implique la *causalité intentionnelle* quand les forces ont la direction d'ajustement des choses aux mots.

Certains actes illocutoires réussis ne sont pas satisfaits. Des assertions peuvent être fausses, des promesses violées, des requêtes refusées et des ordres désobéis. Cependant, le *succès*, la *satisfaction* et la *vérité* sont *logiquement reliés*. La satisfaction d'actes illocutoires élémentaires requiert la vérité de leur contenu propositionnel. Certains actes illocutoires ont des conditions de succès ou de satisfaction plus fortes que d'autres. Quiconque supplie fait une requête. Quand une promesse est tenue, l'assertion de son contenu propositionnel est vraie. De plus, l'accomplissement de nombreux actes illocutoires exige la satisfaction d'autres. Les déclarations réussies contiennent des assertions vraies. Réciproquement, la satisfaction de certaines illocutions requiert l'accomplissement d'autres. On ne peut tenir qu'une promesse que l'on a faite. Quiconque comprend les conditions de félicité raisonne et fait des inférences pratiquement et théoriquement valides. Il comprend que certains actes illocutoires ne peuvent être accomplis et satisfaits sans que d'autres ne le soient également. Ainsi tout locuteur compétent infère d'une prémisses de la forme « Tout homme est mortel et Socrate est un homme » la conclusion assertive « Socrate est mortel » (inférence théorique). De la même façon, il infère d'une prémisses de la forme « Tu devrais aider autrui » la conclusion

directive « Aide les autres ! ». Les conclusions tirées d'inférences théoriques ont la direction d'ajustement des mots aux choses, alors que les conclusions tirées d'inférences pratiques ont la direction d'ajustement des choses aux mots.

Il faut une *théorie réursive unifiée du succès, de la satisfaction et la vérité* pour expliquer tout cela. Dans ma perspective, la théorie sémantique de la vérité défendue par Montague et Davidson pour les langues naturelles fait partie de la théorie plus générale de la satisfaction pour les actes illocutoires de toute force. Les logiques antérieures des actes de discours, comme la logique des commandements de Rescher et la logique des questions de Belnap, et les tentatives d'étendre la logique intensionnelle à l'interprétation des énoncés non déclaratifs (Karttunen, Zaefferer, Stockoff) et d'énoncés dans un discours (Kamp) prétendent à tort que les conditions de succès et de satisfaction des actes illocutoires sont réductibles à leurs conditions de vérité. Elles ignorent les conditions réelles des conditions de félicité des actes de langage. La plupart des approches réduisent, par exemple, chaque *question* à un ensemble de *réponses* possibles à cette question et ensuite ces réponses à leur contenu référentiel ou propositionnel (Harrah).

3. Progrès et méthodologie

Dans *Foundations of Illocutionary Logic* (Searle et Vanderveken 1985), nous avons décomposé les forces en plusieurs composantes, à savoir : leurs buts illocutoires, leurs modes d'atteinte de but illocutoire, leurs conditions sur le contenu propositionnel, leurs conditions préparatoires et de sincérité et leur degré de puissance. Nous avons en outre défini récursivement l'ensemble des forces possibles ainsi que les conditions de succès des actes illocutoires de premier niveau que les locuteurs entendent accomplir à un moment d'énonciation donné. Les cinq *forces primitives* sont les forces les plus simples pourvues d'un but illocutoire propre : elles ont un degré de puissance neutre, et seulement les modes d'atteinte et les conditions qui sont déterminées par leur but. Ce sont : la force d'*assertion* nommée par le verbe performatif « affirmer » et exprimée par le type déclaratif d'énoncé ; la force d'*engagement à une action* nommée par le performatif « s'engager à » ; la force de toute *tentative linguistique destinée à faire agir l'auditeur* qui est exprimée par le type impératif d'énoncé ; la force de *déclaration* nommée par le verbe performatif « déclarer » et exprimée lors d'énonciations performatives, et enfin la force d'*expression* d'attitudes réalisée dans le type exclamatif d'énoncé. Les cinq forces primitives sont universelles. On peut les exprimer en toute langue naturelle possible. Les autres forces illocutoires sont plus complexes ; elles sont obtenues en ajoutant aux forces primitives de nouveaux modes spéciaux d'atteinte de but illocutoire, de nouvelles conditions sur le contenu propositionnel, de nouvelles préparatoires ou de sincérité, ou en changeant le degré de force.

D'un point de vue philosophique, les forces illocutoires actuelles de chaque langue sont des espèces naturelles d'emploi du langage. Elles remplissent des buts linguistiques importants pour les communautés humaines qui parlent

cette langue dans leur environnement historique et naturel. Les composantes spéciales des forces complexes de chaque langue sont donc linguistiquement significatives pour sa communauté linguistique particulière à un certain moment de son histoire. Certaines sont transcendantes, car elles sont indispensables à toute langue humaine capable de remplir ses deux fonctions essentielles d'expression et de communication des pensées des agents. Telles sont les composantes des forces primitives. Telle est aussi la condition préparatoire que le fait représenté par le contenu propositionnel est bon ou mauvais. Car les locuteurs humains sont par nature des personnes avec des préoccupations éthiques. Les forces liées à des aspects historiques particuliers, par exemple la force déclaratoire d'excommunication qui dépend du sacrement de la communion, sont, elles, immanentes.

Les conditions de félicité des actes illocutoires élémentaires sont entièrement déterminées par leur force et leur contenu. Un locuteur *accomplit avec succès* un acte illocutoire de la forme $F(P)$ dont F représente la force et P le contenu propositionnel en un contexte donné d'énonciation quand, en ce contexte, premièrement, le locuteur réussit à atteindre le but illocutoire de cette force sur le contenu avec le mode d'atteinte qui lui est propre, deuxièmement, la proposition P satisfait les conditions sur le contenu propositionnel de F , troisièmement, le locuteur présuppose alors toutes les propositions déterminées par les conditions préparatoires de cette force relativement au contenu et, finalement, il exprime également avec le degré de puissance requis par F toutes les attitudes déterminées par les conditions de sincérité de cette force à propos de P . Ainsi un locuteur fait une *promesse* quand il se met alors lui-même dans l'obligation de faire une action (mode particulier d'atteinte du but illocutoire engageant et condition sur le contenu propositionnel), il présuppose que l'action est bonne pour l'auditeur (condition préparatoire) et il exprime une intention forte de la faire (degré de puissance et condition de sincérité). Un locuteur *rate* un acte illocutoire en un contexte quand il tente sans succès de l'accomplir. À la différence de l'insuccès. Le succès et l'échec contiennent une tentative de l'agent.

Il arrive que les locuteurs présupposent des propositions qui sont fausses et qu'ils expriment des attitudes qu'ils ne possèdent pas. Certaines promesses ne sont pas dans l'intérêt de l'auditeur. Les locuteurs font aussi des promesses qu'ils n'ont pas l'intention de tenir. Pareilles promesses sont défectueuses. Un acte illocutoire est *sans défaut* quand il est accompli avec succès et que toutes ses conditions préparatoires et de sincérité sont remplies. Austin, avec sa notion de *condition de félicité* n'a pas clairement distingué les énonciations réussies qui sont défectueuses des énonciations ratées qui ne sont même pas réussies. En logique illocutoire, un acte illocutoire est *plein de félicité* quand il est à la fois réussi, sans défaut et satisfait. John Searle et moi-même (1985) avons analysé la signification des marqueurs de force et des verbes performatifs de la langue anglaise et avons souligné la nature déclaratoire des énonciations performatives. De nombreux marqueurs de force contiennent des traits syntaxiques qui

s'ajoutent au type syntaxique de l'énoncé et servent à exprimer une force particulière avec des composantes nouvelles. Des énoncés de même type syntaxique peuvent donc exprimer des actes illocutoires de force différente. Les énoncés impératifs comme « Faites-le s'il vous plaît ! » et « Faites-le, que ça vous plaise ou non ! » expriment respectivement une *demande* et un *ordre*. Les énoncés exclamatifs « Comme je suis content de cela ! » et « Comme je suis triste de cela ! » servent à exprimer des attitudes de modes psychologiques différents, respectivement du contentement et de la tristesse. Searle et moi n'avons pas analysé en détail les conditions de satisfaction des actes illocutoires.

À cette fin, j'ai formulé par la suite une *logique prédicative naturelle des propositions* tenant compte de leur double nature : les propositions sont à la fois des *sens d'énoncés* et des *contenus d'actions* et d'*attitudes*. Elles ont une structure de constituants : en exprimant des propositions, nous nous référons aux objets *via* des concepts et leur attribuons des propriétés et des relations. Nous n'avons pas directement à l'esprit les *objets individuels*¹ comme les corps matériels et les personnes. Nous avons plutôt à l'esprit des concepts de ces individus et nous nous y référons indirectement par l'entremise de ces concepts. Nos pensées et nos actions sont donc dirigées vers des individus subsumés sous des concepts plutôt que vers de purs individus. Nous ignorons les valeurs de vérité de nombreuses propositions parce que nous ignorons les dénnotations réelles de leurs constituants en de nombreuses circonstances. On peut se référer à l'épouse du maire sans savoir qui elle est. Pourtant, nous pouvons alors en principe penser à des personnes qui pourraient être cette épouse. Ainsi, dans tous les usages et interprétations du langage, il existe de nombreuses *assignations possibles de dénnotation* aux attributs et aux concepts en plus de l'*assignation standard réelle de dénnotation* qui associe aux constituants propositionnels leur véritable dénnotation en chaque circonstance possible. Les *assignations possibles de dénnotation* que je préconise en sémantique respectent les *postulats de signification*. Seuls des ensembles d'individus subsumés sous des concepts sont des dénnotations possibles de propriétés de premier ordre en chaque circonstance. Les épouses sont des femmes mariées selon toutes les assignations possibles de dénnotation. Selon l'assignation réelle, l'épouse du maire est la femme avec laquelle il est réellement marié quand il est marié. Selon d'autres assignations possibles, ce sont d'autres femmes ou personnes. Lors de la compréhension de propositions, nous savons simplement que leur vérité en chaque circonstance possible est compatible avec certaines assignations possibles de dénnotation à leurs constituants et incompatibles avec les autres. Pour être *vraie*, une proposition doit, bien entendu, être vraie selon l'assignation réelle de dénnotation. La logique prédicative explique pourquoi des propositions vraies dans les mêmes circonstances ont une valeur cognitive différente. Certaines ont différentes structures de constituants. Telles sont les propositions que les mères sont

1. Les objets individuels sont les entités de référence du type logique le plus simple. Voir Strawson (1959).

des femmes et que les érythrocytes sont rouges. Leur expression requiert différents *actes de prédication*. Les autres ne sont pas vraies selon les mêmes assignations possibles de dénotation. Telles sont les propositions que les baleines sont des baleines et que les baleines sont des mammifères. Nous ne les comprenons pas comme étant vraies dans les mêmes conditions. Ainsi nous pouvons affirmer ou croire des vérités nécessaires sans en affirmer ou en croire d'autres. Parmi toutes les vérités nécessaires, peu sont de *pures tautologies* comme la proposition selon laquelle les baleines sont des baleines, qui sont vraies en toute circonstance possible selon toute assignation possible de dénotation. D'un point de vue épistémique, certaines assignations possibles de dénotation sont compatibles avec les croyances d'agents, d'autres ne le sont pas. Si par exemple un agent croit que l'épouse du maire est d'origine étrangère, toutes les assignations possibles de dénotation selon lesquelles elle ne l'est pas sont incompatibles avec ce qu'il croit alors.

Il n'existe pas de correspondance biunivoque entre les verbes performatifs ou les marqueurs de force d'une langue, d'une part, et les forces illocutoires possibles d'autre part. De nombreux verbes performatifs et certains marqueurs de force sont ambigus. On peut aussi bien jurer que quelque chose s'est passé (énonciation assertive) que jurer de faire quelque chose (énonciation engageante). Qui plus est, la structure de surface des marqueurs et des verbes illocutoires n'exhibe pas la forme logique des forces exprimées. Certains types syntaxiques d'énoncés comme les types déclaratif, impératif et exclamatif expriment les forces primitives assertive, directive et expressive, alors que d'autres comme les types conditionnel et interrogatif expriment des forces plus complexes avec des composantes additionnelles. Par exemple, une question est une directive avec une condition spéciale sur le contenu propositionnel, à savoir que l'auditeur donne une réponse correcte à la question. Tout comme on obtient les forces complexes en ajoutant à des forces plus simples des nouvelles composantes ou en changeant leur degré de puissance, on compose en chaque langue des marqueurs de forces complexes en ajoutant au type de l'énoncé des traits syntaxiques nouveaux. Les énoncés impératifs dont le verbe est modifié par l'adverbe « s'il vous plaît » expriment des demandes. Leur adverbe exprime le mode poli d'atteinte du but illocutoire directif : le locuteur donne alors l'option de refus à l'auditeur. Des énoncés comme « Heureusement il pleut » et « Hélas il pleut » expriment des forces assertives complexes ayant respectivement la condition préparatoire et la condition de sincérité exprimée par leur adverbe. Comme Montague l'a souligné, en traduisant les clauses des énoncés ordinaires dans la langue objet de la logique intensionnelle, la sémantique formelle clarifie la forme logique des propositions et parvient à une meilleure explication de leurs conditions de vérité. De même en traduisant les marqueurs de force et les verbes performatifs dans l'idéographie de la logique illocutoire, la sémantique formelle clarifie la forme logique des actes illocutoires et parvient à une meilleure explication de leurs conditions de félicité.

Dans *Meaning & Speech Acts* (Vanderveken, 1990-1991) j'ai élaboré une sémantique formelle générale du succès, de la satisfaction et de la vérité capable d'interpréter, après leur traduction dans l'idéographie, tous les types d'énoncés (déclaratifs ou non) et de formaliser la nature des *inférences* pratiques aussi bien que théoriques. Selon Cocchiarella dans le volume 10 de *Routledge History of Philosophy*, « la notion de *lingua philosophica* en tant que langue idéale contenant à la fois une logique intensionnelle et illocutoire n'est plus simplement un programme mais a d'ores et déjà été réalisée sous de nombreux aspects » (1997 : 72). J'ai enrichi l'ontologie formelle de la sémantique en ajoutant les valeurs de succès (le succès et l'insuccès) aux valeurs de vérité (le vrai et le faux) et en adoptant une analyse prédicative des propositions. Les clauses avec des expressions indexicales expriment différentes propositions en différents contextes. Comme David Kaplan (1989), j'ai procédé à une double indexation sémantique selon laquelle la force et le sens aussi bien que la dénotation peuvent varier d'un contexte d'énonciation à un autre. J'ai ainsi distingué la signification linguistique de la signification en contexte. La signification linguistique d'un énoncé est, dans mon optique, une fonction associant des actes illocutoires aux énoncés en chaque contexte possible. La signification en un contexte d'un énoncé est l'acte illocutoire qu'il y exprime (peu importe que le locuteur l'utilise ou non en ce contexte). De nombreux énoncés expriment différents actes illocutoires en différents contextes, tout comme de nombreux actes illocutoires exprimés en un contexte ont différentes valeurs de succès et de satisfaction en des circonstances différentes. En outre, des énoncés qui ne sont pas synonymes peuvent exprimer le même acte illocutoire en différents contextes. Toute énonciation aujourd'hui de l'énoncé « Il pleut aujourd'hui à Montréal » exprime la même assertion que les énonciations hier de l'énoncé « Il pleuvra demain à Montréal ».

Ainsi, il faut ramifier les notions sémantiques fondamentales d'*analyticité*, de *cohérence* et d'*implication logique* (*entailment*). Toute énonciation littérale de l'énoncé paradoxal de Moore (Wittgenstein, 1958, II, x), « Il va pleuvoir et je ne le crois pas », est *analytiquement ratée* : cet énoncé ne peut être utilisé littéralement avec succès. Car aucun locuteur ne peut sans se contredire accomplir simultanément un acte illocutoire et dénier ses conditions de sincérité. Mais comme nous pouvons mentir, les énonciations littérales de l'énoncé de Moore ne sont pas analytiquement insatisfaites. Cet énoncé exprime une assertion vraie en chaque contexte où le locuteur affirme, sans être sincère, qu'il va pleuvoir. De plus, l'énoncé paradoxal de Moore est *illocutoirement cohérent* : il exprime une assertion performable. Quelqu'un d'autre peut affirmer qu'il va pleuvoir et que je ne le crois pas. Pour ce faire, il doit évidemment utiliser un autre énoncé que celui de Moore.

Comme il y a quatre relations différentes d'inclusion en logique illocutoire entre les conditions de succès et de satisfaction des actes illocutoires, il y a en sémantique générale quatre genres différents d'implication logique (d'*entailment*) entre les énoncés. Certains énoncés *impliquent illocutoirement*

(ou *vériconditionnellement*) d'autres énoncés : ils expriment en chaque contexte un acte illocutoire dont le succès (ou la satisfaction) entraîne le succès (ou la satisfaction) des actes illocutoires exprimés par les autres énoncés dans le même contexte. L'énoncé performatif « Je demande votre aide » implique illocutoirement l'énoncé impératif « Aidez-moi s'il vous plaît ! » et vériconditionnellement l'énoncé conditionnel « Vous pourriez m'aider ». Les notions de cohérence, d'analyticité et d'implication *illocutoire* ont été, par le passé, totalement ignorées en sémantique formelle. Pourtant, elles s'appliquent à tous les types d'énoncés et elles diffèrent des notions de cohérence, d'analyticité et d'implication *vériconditionnelle*. Les différentes notions ramifiées de la sémantique générale ne coïncident pas extensionnellement (voir Vanderveken, 2004). Ainsi les énoncés performatifs comme « Je vous ordonne de le faire » impliquent illocutoirement mais pas vériconditionnellement les énoncés non performatifs comme « Faites-le, que cela vous plaise ou non ! » qui leur correspondent. Comme les déclarations ont la double direction d'ajustement, quiconque déclare avec succès qu'il accomplit un acte illocutoire au moment de l'énonciation accomplit alors *ipso facto* cet acte. Cependant, cet acte illocutoire secondaire n'est pas pour autant satisfait. L'interlocuteur peut désobéir à l'ordre donné.

En assignant comme valeurs sémantiques aux énoncés de chaque contexte des actes illocutoires complets avec leurs conditions de félicité, la sémantique analyse mieux la signification des différents types d'énoncés et les inférences valides qu'on peut faire en les utilisant. Elle établit les liens qui existent entre les énonciations de buts illocutoires différents. D'une part, les énonciations performatives sont du genre le plus fort, car les déclarations réussies sont également satisfaites et sans défaut. Leur succès implique leur félicité. Ce n'est pas sans raison qu'Austin a découvert les forces illocutoires grâce aux énoncés performatifs. Tous les actes illocutoires peuvent être accomplis par déclaration. Par conséquent, les énoncés performatifs impliquent illocutoirement tous les autres types d'énoncés, mais la réciproque n'est pas vraie. On peut accomplir un acte illocutoire sans le faire par déclaration. D'autre part, les énonciations purement expressives sont du genre le plus faible, car tous les actes illocutoires ont des conditions de sincérité. Les énoncés de chaque type impliquent donc illocutoirement les énoncés exclamatifs correspondants. Par exemple, l'énoncé exclamatif « Comme je suis triste qu'il soit mort ! » est impliqué illocutoirement par le déclaratif « Hélas il est mort ».

Nous devons enrichir la *logique intensionnelle* en traitant de l'action, du temps, des modalités et des attitudes afin de bien formuler notre théorie de la félicité. Les actes illocutoires des locuteurs sont des actions intrinsèquement intentionnelles reliées à d'autres actions verbales et non verbales ainsi qu'aux attitudes. C'est en utilisant des signes que les locuteurs tentent d'accomplir leurs actes illocutoires. Certains actes illocutoires ont même des contenus propositionnels qui représentent les actions présentes ou futures des interlocuteurs. Par exemple, les contenus propositionnels des engagements et des déclarations représentent des actions de leur locuteur et celles des directives des actions de

leurs auditeurs. Notre idéographie plus riche donne à la *sémantique formelle* les moyens de mieux montrer après traduction les formes canoniques des actes illocutoires exprimés par tous les types d'énoncés (déclaratifs ou non) des langues naturelles. Indiscutablement, la force et le contenu des énonciations dépendent souvent des énonciations antérieures. Quiconque réplique à l'énonciation interrogative « Quel âge avez-vous ? » en spécifiant son âge fait plus qu'une assertion ; il donne une réponse à la question posée. Nous avons besoin d'une *représentation du discours* (Kamp & Reyle, 1993) afin de résoudre dynamiquement les ambiguïtés, les ellipses, les anaphores, les présuppositions et les composantes de force dans le contexte de l'énonciation. Ainsi, la sémantique doit considérer, en plus du locuteur, du ou des auditeur(s), du moment et du lieu de chaque contexte possible d'énonciation, son arrière-plan conversationnel (les énonciations antérieures, les formes de vie et les attitudes partagées par les interlocuteurs, notamment leurs connaissances mutuelles).

Nous avons besoin d'une conception ramifiée du temps (Prior, 1967) et des modalités historiques (Thomason, 1984) afin rendre compte de l'indéterminisme du monde actuel et de la liberté des agents humains. Selon l'indéterminisme, la façon dont les choses sont à un moment n'est pas entièrement déterminée par la façon dont elles ont été auparavant. En particulier, nos actes de discours et nos autres actions ne sont pas déterminés. Lorsque nous disons et faisons quelque chose, nous aurions pu dire ou faire autre chose, ou ne rien dire ou faire du tout. Selon la conception ramifiée du temps, chaque moment représente un état complet possible du monde actuel à un instant donné. Plusieurs moments incompatibles peuvent succéder à un moment dans le futur de ce monde. Ce moment appartient alors à plusieurs histoires ayant le même passé et le même présent, mais différentes continuations historiques. Chaque histoire représente un des cours possibles de l'histoire du monde actuel. Comme Belnap (2001), je pense que les *circonstances possibles* où les propositions ont une valeur de vérité sont des paires d'un moment du temps et d'une histoire à laquelle ce moment appartient. Les propositions futures peuvent être vraies à un certain moment selon certaines continuations possibles de ce moment et fausses selon d'autres. Certaines histoires ont un moment final. Elles représentent un cours possible du monde qui se termine à ce moment final. Cependant, les moments qui ne sont pas finaux continuent toujours dans une seule direction. Comme Occam l'a souligné, parmi toutes les histoires auxquelles un moment appartient, il y en a une et une seule qui représente la marche du monde après ce moment (voir Prior, 1967). Aucun agent (même Dieu) ne sait à un moment donné quelle histoire sera sa continuation historique réelle, mais chacun est capable de distinguer cette histoire propre des autres et d'y faire référence pour représenter la valeur de satisfaction encore indéterminée des actes et illocutoires, et des attitudes propres aux locuteurs à ce moment. Dans mon optique (Vanderveken, 2006), les actes illocutoires et les attitudes ont une valeur de satisfaction propre dans chaque contexte, même si cette valeur est encore indéterminée quand leur contenu propositionnel est futur. Une prédiction

faite maintenant sera vraie seulement si les choses seront telles que représentées dans la continuation historique réelle de ce monde, quelle que soit cette continuation. De même, une promesse faite au moment présent sera tenue seulement si son agent accomplit la chose promise dans la continuation réelle du monde. Les autres continuations historiques possibles n'importent pas.

Durant la dernière décennie, j'ai enrichi la logique propositionnelle prédicative en traitant de la *généralisation*, des *modalités* logiques et historiques, du *temps* ramifié et des *attitudes* aussi bien volitives que cognitives. J'ai aussi élaboré la *logique de l'action intentionnelle*. À cette fin, j'ai analysé les *tentatives* des agents qui sont des actions d'un genre spécifique : elles sont à la fois personnelles, intentionnelles, libres et réussies. Les agents font personnellement et volontairement leurs propres tentatives. Quiconque tente de faire une tentative la fait. Il aurait pu ne pas la faire. Les agents font leurs tentatives afin d'atteindre un autre objectif. Pour atteindre cet objectif, ils doivent faire une bonne tentative en une bonne circonstance. Quand un agent rate son objectif, sa tentative n'est pas satisfaite. Dans ma logique de l'action (2005), les actions intentionnelles sont premières, comme en philosophie. Les actions inintentionnelles des agents sont en effet *engendrées* par leurs actions intentionnelles et elles auraient aussi pu être tentées. Ainsi, les erreurs et les échecs ne sont pas des actions mais plutôt des événements qui nous arrivent. De plus, aucune action n'est inévitable. L'agent aurait pu ne pas la faire.

Ma théorie formule les lois gouvernant les tentatives des agents et elle explique la nature de leurs actions intentionnelles et de leurs actions de base. Elle rend aussi compte de la rationalité minimale des agents qui ne sont ni parfaitement rationnels ni entièrement irrationnels. Comme nous le verrons, notre rationalité minimale est reliée à la façon dont nous représentons les conditions de félicité. Il nous arrive de croire et de tenter de faire des choses impossibles. Nos possibilités subjectives ne sont donc pas toujours objectives. Cependant, nous ne pouvons ni croire ni tenter de faire des choses impossibles qui soient purement contradictoires. Mon explication des possibilités subjectives est basée sur la définition suivante de la vérité selon un agent : une proposition est *vraie selon un agent dans une circonstance* quand celui-ci a alors à l'esprit tous ses concepts et attributs et que cette proposition est vraie dans cette circonstance selon toutes les assignations possibles de dénotation qui sont alors *compatibles avec ses croyances*. Nous avons à l'esprit des propriétés sans savoir quelles entités les possèdent, tout en sachant bien que seules certaines entités pourraient les posséder. C'est pourquoi certaines, mais pas toutes, les assignations possibles de dénotation sont compatibles avec nos croyances. Nous apprenons *a posteriori* la plupart des propriétés essentielles des objets ; j'entends par là des propriétés qu'ils possèdent réellement en toute circonstance possible comme par exemple que les baleines sont des mammifères. Nous ignorons bien des propriétés essentielles. Rares sont les agents qui connaissent leur code génétique. Des assignations possibles de dénotation compatibles avec nos croyances violent parfois des propriétés essentielles des objets auxquels nous pensons.

Ainsi, il nous arrive d'être incohérents et de croire des propositions nécessairement fausses. Dans le temps, nous croyions que les baleines étaient des poissons. Cependant, toutes les assignations possibles de dénotation à l'œuvre dans l'usage et l'interprétation du langage respectent les postulats de signification. Voilà pourquoi nous sommes et restons à tout le moins minimalement rationnels. Quand une proposition est par définition vraie en toute circonstance possible selon toutes les assignations possibles de dénotation, nous savons *a priori* qu'elle est nécessairement vraie. Ainsi, nous ne pouvons exprimer de pures *tautologies* sans savoir leur vérité. Nous faisons les inférences théoriquement valides dont les prémisses *impliquent fortement* la conclusion. Une proposition en *implique fortement* une autre quand, premièrement, elle contient toutes ses prédications et, deuxièmement, elle *l'implique tautologiquement* en ce sens qu'il est tautologique que si elle est vraie, l'autre l'est également. Ma relation d'*implication forte* exige beaucoup plus que l'*implication stricte* de Lewis : elle est finie, paracohérente, tautologique, analytique et décidable. Elle est à la base des engagements psychologiques et illocutoires des locuteurs basés sur la compréhension des conditions de vérité. Cela permet d'améliorer la théorie de la déduction naturelle. Seules certaines règles de la déduction naturelle comme l'élimination de la disjonction engendrent de l'implication forte. Ainsi, une intention et une promesse de boire de la bière ou du vin contiennent respectivement une intention et une promesse de boire. Par contre, l'introduction de la disjonction n'engendre pas d'implication forte, puisque sa conclusion peut avoir de nouveaux sens. Une intention de boire ne contient pas une intention de boire *ou* de mourir.

Pour qu'un agent *réussisse* à faire des choses, il ne suffit pas qu'il tente de les faire et qu'elles arrivent. Il faut que ces choses soient causées par sa propre tentative. Comment expliquer pleinement la causalité intentionnelle ? Souvent la tentative de l'agent est *la* cause même de l'état de choses tenté. Cependant parfois, il y a une surdétermination causale : les agents font certaines choses pour différentes raisons. La tentative des agents est alors simplement être une *raison pratique* de leur survenance. Il convient d'analyser la notion de *raison pratique* afin d'expliquer la *causalité intentionnelle* et les conditions de satisfaction des actes illocutoires dont la direction d'ajustement va des choses aux mots. Un auditeur *obéit* à un ordre quand cet ordre est une raison pratique pour laquelle il fait les choses ordonnées.

Grâce à la logique de l'action, du temps, des attitudes et des modalités, la logique illocutoire peut mieux analyser les contenus propositionnels et les forces. Son idéographie peut exprimer les actions passées, présentes et futures, les capacités et les attitudes des interlocuteurs. Elle peut aussi caractériser la façon dont les actes illocutoires sont reliés aux autres actes de langage (les actes d'énonciation, les actes propositionnels, les tentatives d'accomplir des actes illocutoires et les actes perlocutoires). Les tentatives visant à accomplir des actes illocutoires sont un nouvel acte de langage fort important dans ma taxonomie. Car elles sont constitutives de la *signification*. À mes yeux, une énonciation est

pourvue de signification lorsque le locuteur tente d'accomplir un acte illocutoire, peu importe qu'il réussisse ou qu'il échoue. Les locuteurs tentent d'accomplir leurs actes illocutoires publics en émettant des signes. Il convient d'expliquer *comment* et *en quelles conditions ils réussissent* et de quelles façons leurs actes illocutoires réussis en *engendrent* d'autres (chaque invitation contient une requête) et ont des effets perlocutoires (l'auditeur est parfois influencé). À la base de toute communication, les agents tentent de mouvoir des membres de leur propre corps et cela *engendre* (Goldman, 1970) de diverses façons leurs actes de langage. Cet *engendrement* dans la communication est d'abord physiquement causal (nous utilisons oralement les énoncés en proférant des sons) ; il est ensuite conventionnel (la signification des énoncés sert à déterminer la nature des actes illocutoires exprimés). L'engendrement est parfois *simple* (les locuteurs mentent lorsqu'ils n'ont pas les attitudes exprimées) ou *par extension* (ils accomplissent parfois indirectement des actes illocutoires non littéraux). Afin d'expliquer l'engendrement des différents genres d'actes de langage, j'intégrerai la logique illocutoire dans le cadre de la logique de l'action.

Ma méthodologie est liée à mon effort de formalisation. Les notions théoriques primitives que j'ai choisies obéissent à des postulats de signification clairs et distincts et se prêtent à des axiomes évidents. Les autres notions importantes sont définissables à partir des primitives par élimination, par explication formelle ou par induction récursive. Les lois de l'idéographie devraient être valides selon la théorie des modèles et, autant que possible, prouvables voire décidables quand les agents les ont intériorisées par l'entremise de leur compétence. Je vise évidemment des axiomatisations aussi complètes que possibles et une théorie permettant un traitement linguistique et computationnel adéquats des langues naturelles.

4. La signification des énoncés et la signification des locuteurs dans la conversation

Comme Frege, je pense que la force, le sens et la dénotation sont les trois composantes de base de la signification des énoncés et des locuteurs. Quels actes illocutoires le locuteur entend-il accomplir dans le contexte d'une énonciation ? La théorie des actes de langage adopte un principe de littéralité : elle identifie la signification du locuteur avec la signification de l'énoncé pour autant que cela soit possible. Dans cette optique, le locuteur tente, en premier lieu, d'accomplir l'acte illocutoire exprimé par l'énoncé dont il fait usage quand pareil acte est alors *performable*. Les langues naturelles offrent de riches moyens linguistiques pour exprimer les forces. De nombreux analystes négligeant ces moyens en viennent à ignorer les forces dans leur sémantique ou à multiplier inutilement les forces littérales. En suivant le principe d'économie d'Occam, Searle et moi, comme Grice, ne multiplions pas sans raison la signification des énoncés ni l'ambiguïté sémantique. Nous faisons autant que possible appel aux maximes conversationnelles et à l'arrière-plan pour expliquer la signification non littérale du locuteur. Ainsi, les énoncés déclaratifs, interrogatifs et

optatifs tels que « Je désire votre aide », « Si seulement vous vouliez m'aider » et « Pouvez-vous m'aider ? » si souvent utilisés pour demander indirectement de l'aide expriment littéralement dans tout contexte, selon ma sémantique générale, respectivement une assertion, une question et un souhait. Comme Grice (1975) l'a souligné, la plupart des actes illocutoires non littéraux sont annulables. Dans d'autres contextes, le locuteur utilisant le même énoncé ne les accomplirait pas. Il en va ainsi pour les demandes récurrentes indirectes que l'on fait avec les énoncés précédents, lesquelles sont des implicatures conversationnelles généralisées dans la terminologie de Grice.

La plupart des énoncés peuvent être utilisés littéralement. Grâce à sa nouvelle logique propositionnelle et de l'action, ma sémantique générale dispose désormais d'une idéographie plus riche avec des capacités expressives plus grandes. Elle peut dorénavant mieux traduire les énoncés des langues naturelles et assigner aux actes illocutoires exprimés une forme canonique plus précise clarifiant leurs conditions de félicité. Dans mon optique, lors d'énonciations non littérales, les locuteurs expriment les actes illocutoires littéraux mais ils entendent en accomplir d'autres. Autrement, il y aurait tellement de sens et de forces qu'une véritable théorie de la signification linguistique serait impossible. Les locuteurs utilisant des énoncés illocutoires incohérents comme « Je ne suis pas moi-même aujourd'hui » savent, en vertu de leur compétence, que l'acte illocutoire littéral est imperformable. Ils ne signifient pas ce qu'ils disent. La sémantique générale qui traite du seul usage littéral n'est pas en mesure d'analyser les énonciations non littérales. Mais elle peut être intégrée dans une pragmatique systématique et unifiée plus générale capable de construire effectivement la signification non littérale des locuteurs à partir de la signification des énoncés utilisés et de l'arrière-plan conversationnel.

En vertu de sa forme logique, chaque acte illocutoire tenté engage le locuteur à de nombreux autres. Un acte illocutoire *engage fortement le locuteur* à un autre quand il ne peut alors accomplir cet acte sans *ipso facto* accomplir l'autre. Toutes les prédictions contiennent des assertions, mais la réciproque n'est pas vraie. Les assertions sur le passé ne sont pas des prédictions. Quiconque entend accomplir un acte illocutoire entend *ipso facto* accomplir les autres actes illocutoires ayant moins de conditions de succès. Mais les locuteurs sont aussi parfois engagés à des actes illocutoires plus forts. Celui qui réplique à une assertion en disant le contraire fait plus qu'affirmer la négation de son contenu. Il contredit alors le locuteur précédent. Ma sémantique formelle générale peut dorénavant expliquer par l'engendrement simple de pareils engagements à des actes illocutoires plus forts. Dans les contextes où sont remplies des conditions sur le contenu propositionnel, des conditions préparatoires ou de sincérité spéciales, l'acte illocutoire littéral réussi engendre alors des actes illocutoires avec une force supérieure. Dans les autres contextes, il n'y a pas de tel engendrement illocutoire simple.

Comme Searle et moi l'avons signalé, les locuteurs sont aussi faiblement engagés à des actes illocutoires qu'ils n'accomplissent pas ouvertement. Celui

qui accepte d'aider tout le monde est engagé à accepter de vous aider même s'il n'a fait aucune référence à vous. Nul ne peut littéralement dire « Je refuse de vous aider et j'accepte d'aider n'importe qui. » Les tentatives peuvent échouer. Mais quiconque tente d'accomplir un acte illocutoire est faiblement engagé à cet acte. Comment expliquer à quels actes de langage le locuteur est faiblement engagé ? Il existe une relation réflexive et symétrique de *compatibilité agentive* en logique de l'action. Deux moments sont *compatibles eu égard à un agent* quand toutes les actions qu'il a accomplies à ces deux moments sont simultanément performables. Cette relation permet de définir rigoureusement l'engagement illocutoire en théorie du succès. Un locuteur est *faiblement engagé* à un acte illocutoire, dans un certain contexte, quand il peut accomplir cet acte illocutoire à tout moment qui est compatible, en ce qui le concerne, avec le moment de l'énonciation de ce contexte. Toutes les lois gouvernant les engagements illocutoires faibles découlent de cette définition. Voilà une simplification majeure de *Foundations* où la notion d'engagement faible était primitive.

La signification du locuteur est souvent différente de la signification de l'énoncé dans le dialogue. Non seulement les interlocuteurs parlent souvent de façon non littérale ou non sérieuse, mais ils partagent des formes de vie et entendent en outre, par leurs interactions verbales, atteindre des objectifs communs et accomplir ensemble des actions. Comme Wittgenstein l'a souligné (1958), les énoncés sont des instruments qui possèdent des rôles et des fonctions à l'intérieur de jeux de langage. « Ici, le terme de '*jeu de langage*' est un moyen de souligner le fait que *l'usage* du langage fait partie d'une activité, d'une forme de vie » (Wittgenstein, 1958 : 23). Notre question est à présent : peut-on généraliser la théorie des actes de langage de façon à traiter de la conversation ?

Avant tout, la poursuite d'une conversation est gouvernée par un principe général de coopération (Grice, 1975). Les interlocuteurs qui entendent contribuer à une conversation respectent certaines maximes conversationnelles comme « Dites le Vrai ! », « Croyez ce que vous dites ! ». Selon Grice, ce n'est pas sans raisons que les interlocuteurs respectent pareilles maximes. Cela est dû à leur rationalité ; il est en effet rationnel de respecter ces maximes pour coopérer et poursuivre effectivement une conversation. Comme je l'ai déjà dit, ce sont les actes illocutoires tentés qui importent dans le discours. Beaucoup d'énonciations ne sont pas littérales ou sérieuses. Quiconque exprime dans un contexte une assertion qui est manifestement fausse, défectueuse ou insincère, étant donné l'arrière-plan conversationnel, n'entend pas alors faire cette assertion mais un autre acte illocutoire avec d'autres conditions de félicité compatibles avec l'arrière-plan. Dans l'approche de Searle et Grice, afin de comprendre les actes illocutoires non littéraux, l'auditeur doit d'abord déterminer l'acte illocutoire littéral et comprendre que le locuteur ne peut alors tout simplement vouloir accomplir cet acte de langage s'il respecte les maximes conversationnelles, étant donné la connaissance factuelle partagée de l'arrière-plan du contexte de l'énonciation. Afin d'expliquer la signification réelle des locuteurs, la *pragmatique* doit ajouter à la sémantique, conçue

comme théorie de la signification des énoncés, une théorie des maximes conversationnelles et une analyse des faits pertinents de l'arrière-plan conversationnel. Dans mon approche, les auditeurs *infèrent* les actes illocutoires non littéraux tentés à partir de la signification dans le contexte de l'énonciation des énoncés utilisés, des conditions de félicité des actes illocutoires exprimés et pertinents et de l'hypothèse du respect des maximes conversationnelles étant donné l'arrière-plan. J'ai généralisé (Vanderveken, 1997) la *maxime conversationnelle de qualité* de Grice (expliquée comme « Veillez à ce que votre acte illocutoire tenté soit plein de félicité ! ») ainsi que celle de *quantité* (expliquée comme « Faites-le aussi fort qu'il le faut pour parvenir à vos fins ! »). Comme je l'ai dit plus haut, les conditions de félicité d'un acte illocutoire sont la somme de ses conditions de succès, d'accomplissement sans défaut et de satisfaction. Un acte illocutoire est *plus fort qu'*un autre quand il possède plus de conditions de félicité. Les deux maximes généralisées s'appliquent dorénavant à tous les genres d'énonciations. La *sous-maxime de qualité pour les directives* est la suivante : « Veillez à ce que votre directive soit réussie, sans défaut et suivie par l'auditeur ! » La *sous-maxime de quantité* est : « Veillez à ce que votre directive soit aussi forte qu'il le faut ! » Les maximes de qualité et de quantité, telles que Grice les a formulées, sont juste des applications de mes maximes dans le cas particulier des assertions de discours à but informatif. Afin de respecter les maximes, les locuteurs doivent bien sélectionner la force et le contenu des actes tentés. Ainsi, ils obéissent aux principes suivants : « Ne donnez pas d'ordre mais faites plutôt une requête à l'auditeur lorsque tout repose sur son bon vouloir ! » (sinon votre directive sera défectueuse). « Ne lui demandez pas des choses qu'il ne fera pas ! » (sinon votre directive ne sera pas satisfaite). « Ordonnez à l'auditeur quand vous êtes en position d'autorité et que vous voulez vraiment qu'il agisse ! » (une demande serait alors trop faible). N'ordonnez ni plus ni moins que ce que vous voulez ! » (autrement le contenu de votre directive ne sera pas aussi fort qu'il le faut).

Il existe des formes normales d'inférences menant aux actes illocutoires non littéraux, tentés. Leurs prémisses représentent un nombre fini de faits de l'arrière-plan dont l'existence est présupposée par le locuteur et sur lesquels il entend attirer l'attention des auditeurs. Par conséquent, une approche théorique est en principe compatible avec l'ouverture de l'arrière-plan. Les faits pertinents sont des contenus d'attitudes et reliés aux conditions de félicité. Les auditeurs infèrent les actes illocutoires non littéraux que tentent les locuteurs en exploitant et utilisant les différentes maximes conversationnelles. Comme Grice l'a remarqué, des locuteurs peuvent tranquillement et inostensiblement violer une maxime sans pour autant tenter d'accomplir un acte illocutoire non littéral. Parfois, il y a un conflit entre deux maximes. Le locuteur voudrait bien boire toute la bouteille (maxime de quantité), mais il sait que l'auditeur n'est pas disposé à la lui donner (maxime de qualité). En pareil cas, il vaut mieux simplement faire une demande plus faible, demander un verre plutôt que la bouteille. Il arrive en outre aux locuteurs d'accomplir

sans le vouloir des actes dépourvus de félicité. Il leur arrive aussi de mentir et de faire, par exemple, des promesses qu'ils ne sont pas capables ou n'ont pas l'intention de tenir afin d'induire en erreur les auditeurs. En pareils cas, les locuteurs ne veulent pas que les auditeurs soient au courant de la violation de la maxime. Pour qu'il y ait exploitation d'une maxime au sens de Grice, il faut que le locuteur veuille que la violation de cette maxime soit manifeste pour les interlocuteurs. Ma notion d'exploitation est une généralisation de celle de Grice. Dans mon optique, un locuteur *exploite* une maxime conversationnelle quand il entend que l'auditeur reconnaisse des faits de l'arrière-plan qui l'empêcheraient de respecter cette maxime si son acte illocutoire principal était l'acte littéral. Dans le contexte le locuteur est bien coopératif et il peut respecter la maxime en question sans enfreindre une autre. C'est donc qu'il entend alors accomplir principalement un autre acte illocutoire que le littéral et communiquer cette intention. Pareille explication de l'exploitation des maximes permet de caractériser des figures de style. L'ironie et le sarcasme sont des cas particuliers d'exploitation de la maxime de qualité alors que l'euphémisme et l'hyperbole sont des exploitations de la maxime de quantité. Pour comprendre pareilles figures de style, l'auditeur doit reconnaître les faits exploités de l'arrière-plan.

Dans le cas de l'*ironie*, la signification du locuteur est *opposée* à la signification de l'énoncé utilisé. L'arrière-plan contient alors des faits exploités incompatibles avec les conditions de félicité de l'acte littéral qui forcent le locuteur à tenter d'accomplir un acte illocutoire non littéral ayant des conditions opposées. L'ironie du locuteur est dirigée vers les composantes de la force et du contenu littéral qui déterminent les conditions manifestement violées de félicité. En disant « Je promets de vous blesser », un locuteur menace ironiquement l'auditeur lorsque l'action envisagée est clairement néfaste pour ce dernier. Son ironie concerne alors une condition préparatoire de l'acte littéral. Une promesse doit être bonne pour l'auditeur. En disant « Je vous remercie pour votre aide », un locuteur se plaint ironiquement que le locuteur ne lui soit pas venu en aide, lorsqu'il est manifeste qu'il n'a reçu aucune aide et qu'il n'est pas du tout reconnaissant mais au contraire fort insatisfait de cela. L'ironie, dans ce cas, est liée à la fois aux conditions de sincérité et de satisfaction littérales.

Dans le cas des *actes de langage indirects*, la signification du locuteur est *une extension* de la signification de l'énoncé utilisé. Le locuteur s'appuie alors sur le fait que l'acte illocutoire littéral n'est manifestement pas assez fort pour atteindre tous ses objectifs linguistiques. La plupart du temps, le locuteur tente manifestement d'atteindre un autre but illocutoire que le littéral sur le contenu propositionnel ou bien il tente d'atteindre le but illocutoire littéral avec un mode spécial non exprimé. Souvent, des conditions préparatoires ou de sincérité non littérales sont manifestement remplies dans l'arrière-plan. Ainsi, on peut promettre indirectement de l'aide en disant « Je vous aiderai ». Quand l'indirection est dirigée vers le contenu propositionnel, les conditions de satisfaction de l'acte littéral sont des conditions de félicité de l'acte indirect. Le locuteur affirme parfois que ces conditions sont remplies. Parfois il

demande à l'auditeur si elles le sont. Dans le premier cas, l'assertion littérale est alors manifestement vraie, dans le second cas, la question littérale a ou pourrait fort bien avoir une réponse positive. Ainsi, nous offrons et même souvent nous promettons indirectement de l'aide en disant « Je pourrais vous aider » (conditions préparatoires de l'acte indirect), « J'ai l'intention de vous aider », (conditions de sincérité), « Je devrais vous aider », « Devrais-je vous aider ? » (mode de réalisation du but engageant propre à une promesse). Lors de tels usages idiomatiques, le contenu propositionnel des actes illocutoires indirects fait partie du contenu littéral. Un acte illocutoire indirect est *catégorique* à chaque fois que ses conditions de félicité non littérales font partie de la connaissance mutuelle d'arrière-plan. Parfois, les intentions du locuteur dépendent de la réponse de l'auditeur. En pareils cas, l'acte de langage indirect est *conditionnel*. Nous *offrons* indirectement notre aide à l'auditeur quand nous lui demandons s'il est disposé à recevoir cette aide. De telles offres sont des promesses indirectes qui sont conditionnelles à l'acceptation de l'auditeur.

Mon analyse rend compte des indirections dirigées vers la force aussi bien que vers le contenu. Elle explique les actes illocutoires indirects accomplis grâce à des énonciations non assertives. Contrairement à Sperber et Wilson (1975) dont la notion de pertinence reste vague et informelle, je pense que nous calculons effectivement les actes illocutoires non littéraux tentés. Ils sont minimalement différents des actes de langage littéraux. En général, les actes illocutoires ironiques diffèrent seulement des littéraux par le fait qu'ils ont les composantes opposées aux composantes littérales manifestement incompatibles avec l'arrière-plan. Les actes de langage indirects sont obtenus en ajoutant aux composantes littérales de nouvelles composantes non exprimées déterminant les conditions de félicité manifestement remplies dans l'arrière-plan. Les locuteurs respectent donc les maximes exploitées en faisant de l'ironie et de l'indirection. Les actes illocutoires ironiques ont des conditions de félicité compatibles avec l'arrière-plan (comme l'exige la maxime de quantité). Dans les cas d'indirection, le locuteur entend accomplir, en plus de l'acte littéral, l'acte indirect. Ces deux actes illocutoires sont ensemble aussi forts qu'il le faut pour que le locuteur atteigne tous ses objectifs. Il respecte donc bien alors la maxime de quantité. Les autres figures de style et maximes peuvent être définies selon la même approche. Notons qu'une pragmatique intégrée des actes illocutoires non littéraux est nécessaire pour établir lier les sémantiques synchronique et diachronique en sémiotique. En effet, des modèles récurrents de signification non littérale dans des formes de vie récurrentes d'arrière-plan ont tendance à être lexicalisés après un certain temps. L'on peut ainsi concevoir une théorie du changement de la signification expliquant l'apparition de nouvelles significations littérales (par exemple de métaphores, d'ironies et d'indirections mortes) dans l'histoire de chaque langue. Soit dit en passant, les maximes conversationnelles de qualité et de quantité concernent la nature même des actes illocutoires. Elles ne sont pas relatives à une culture humaine particulière mais sont plutôt des universaux pragmatiques. Selon Grice, elles dérivent

de la rationalité des locuteurs. Les actes illocutoires sont des moyens naturels d'atteindre des buts linguistiques. Or il existe une certaine raison pratique dans la sélection des moyens appropriés à chaque fin. Des agents rationnels devraient décider d'utiliser en chaque circonstance les meilleurs moyens efficaces disponibles. De même, des locuteurs rationnels devraient tenter d'y accomplir des actes illocutoires à propos. Ainsi, il est raisonnable de respecter la maxime de qualité. Il en va de même pour la maxime de quantité. Tout agent rationnel devrait utiliser des moyens aussi efficaces qu'il le faut dans ses actions verbales aussi bien que non verbales. Or à chaque but linguistique correspond un acte illocutoire unique servant à atteindre ce but pleinement et de la façon la plus effective. Un locuteur qui tenterait d'accomplir un acte illocutoire plus faible ou plus fort n'agirait pas avec la plus grande efficacité pour parvenir à son but.

Searle (1982) a souligné d'autres difficultés importantes pour la théorie du discours : les interlocuteurs qui dialoguent n'ont souvent aucun objectif conversationnel ; ils changent librement le genre et le sujet du dialogue ; leurs énonciations ratées, sans félicité et totalement impertinentes n'empêchent parfois pas leur dialogue de continuer. De plus, l'arrière-plan conversationnel est indéfiniment ouvert et par conséquent impossible à décrire exhaustivement. Enfin, leur intentionnalité discursive est irréductiblement collective. J'ai répliqué à Searle (Vanderveken, 1997, 2001), en faisant les considérations suivantes. Bien sûr, les locuteurs sont engagés dans des formes de vie sociales, et leurs objectifs sont fort souvent extralinguistiques. Ils communiquent alors afin de coordonner intelligemment leurs actions non verbales. Je suis d'accord avec Wittgenstein (1958) et Searle (1982), qu'il n'existe pas de véritable théorie possible de *tous les jeux de langage possibles*. Il y a d'innombrables genres de formes de vie et d'objectifs que nous pouvons partager en tant qu'être humain. Conséquemment, il existe « d'innombrables genres » de jeux de langage que nous pouvons pratiquer. Comme Wittgenstein le dit, « [...] cette multiplicité n'est pas quelque chose de fixe, de donnée une fois pour toutes ; mais de nouveaux types de langages, de nouveaux jeux de langage, nous pouvons le dire, en viennent à exister, alors que d'autres deviennent obsolètes et tombent dans l'oubli » (Wittgenstein, 1958 : 23).

Pourtant, en tout dialogue, les locuteurs entendent accomplir des actes illocutoires pertinents afin de contribuer à des échanges verbaux comme les présentations, les justifications, les débats, les compromis et les ententes qui sont des actes de langage de niveau supérieur à but conversationnel. La logique du discours peut analyser la structure et la dynamique de tels jeux de langage dont la conduite est régie par des systèmes de règles constitutives. Ma taxonomie postule *quatre buts conversationnels* correspondant aux quatre directions possibles d'ajustement entre les mots et les choses. Les discours avec la direction d'ajustement des mots aux choses (comme les rapports, les entrevues, les nouvelles et les interrogations) ont un *but descriptif* : ils servent à décrire comment les choses sont dans le monde. Les discours avec la direction d'ajustement

des choses aux mots (comme les négociations, les marchandages, les paris, les tentatives d'arrangement et les sermons) ont un *but délibératif* : délibérer comment agir dans le monde. Les discours avec la double direction d'ajustement (les inaugurations, les permis, les baptêmes, les investitures et les classifications) ont un *but déclaratoire* : ils servent à faire des choses par déclaration. Finalement, les discours sans direction d'ajustement (les salutations, les accueils, les éloges et les protestations) ont un seul *but expressif* : ils servent seulement à exprimer les attitudes communes des interlocuteurs. Les locuteurs compétents sont tous capables d'atteindre les buts conversationnels à cause de leur intentionnalité. Il y a une correspondance biunivoque entre les buts conversationnels et les directions possibles d'ajustement entre les mots et les choses, car les véritables protagonistes de chaque dialogue sont les interlocuteurs qui y contribuent successivement. Dans un contexte isolé d'énonciation, le locuteur est actif alors que l'auditeur est passif. Aucun locuteur ne peut engager quelqu'un d'autre à agir par son énonciation personnelle. Il peut seulement s'engager lui-même ou donner une directive à un autre agent. C'est pourquoi le langage distingue naturellement deux buts illocutoires avec la direction d'ajustement des choses aux mots. Quand l'énonciation est engageante, la responsabilité de changer le monde repose sur le locuteur ; quand elle est directive, elle repose sur l'auditeur. Cependant, la situation est entièrement différente lors de la poursuite d'une conversation. Tout auditeur est un locuteur potentiel qui peut parler à son tour et faire sa propre contribution. Ainsi, lors d'une délibération, les auditeurs peuvent répliquer aux locuteurs qui leur donnent des directives et font des offres. Ils peuvent par la suite accepter, refuser ou faire des contre-offres. Souvent, les engagements des locuteurs sont conditionnels aux engagements futurs des interlocuteurs. Par conséquent, il y a un seul but conversationnel alors qu'il y a deux buts illocutoires différents avec la direction d'ajustement des choses aux mots. Les délibérations servent autant à engager les interlocuteurs qu'à leur donner des directives ; ils servent à déterminer des actions réciproques à faire dans le monde.

Toutes les forces ayant le même but illocutoire ne jouent pas le même rôle dans l'usage du langage. Il vaut parfois mieux conseiller qu'ordonner. De même, tous les types de dialogue ayant le même but conversationnel n'ont pas la même fonction. Parfois, il vaut mieux expliquer et justifier sa position plutôt que de la présenter unilatéralement. Ma typologie décompose les types discursifs en d'autres composantes que leurs buts conversationnels. Beaucoup ont un mode caractéristique d'atteinte du but conversationnel, qui requiert l'usage de certains moyens ou une façon particulière de procéder. La plupart des modes imposent une certaine séquence d'actes illocutoires majeurs. Dans les contrats écrits, les parties doivent s'engager à des actions réciproques futures et s'entendre officiellement en donnant leur signature. De nombreux types de discours imposent des conditions à leur thème. Une entrevue d'embauche doit décrire les qualités professionnelles de l'interviewé. Les conditions thématiques déterminent à la fois la force et le contenu propositionnel des actes illocutoires

majeurs. Les délibérations d'un jury doivent aboutir à un verdict. De nombreux types de discours exigent que des conditions préparatoires soient remplies dans l'arrière-plan. Par exemple, lors d'une consultation médicale, on présuppose que la personne consultée a des aptitudes en médecine et que celui qui consulte veut des conseils concernant sa santé. Un discours est défectueux quand ses conditions préparatoires sont violées. Finalement, beaucoup de types discursifs exigent que les protagonistes expriment des attitudes communes. Lors d'échanges de salutations, les interlocuteurs reconnaissent courtoisement la présence des autres à leur approche. Dans ma typologie deux jeux de langage ont le même type discursif quand ils ont les mêmes buts conversationnels, les mêmes modes d'atteinte de ces buts, les mêmes conditions thématiques et les mêmes conditions d'arrière-plan et de sincérité.

J'ai formulé une définition récursive de l'ensemble de tous les types possibles de discours à but conversationnel et j'ai systématiquement analysé leurs conditions de félicité. Les *quatre types primitifs de discours* sont les types les plus simples correspondant aux quatre buts conversationnels : la *description*, la *délibération*, la *déclaration* et l'*expression* ; ils sont communs à tous les jeux de langage respectivement descriptifs, délibératifs, déclaratoires, et expressifs. Les autres *types de discours plus complexes sont obtenus en ajoutant de nouvelles composantes aux types primitifs*. Ainsi le type de la *négociation* possède le mode particulier d'atteinte du but délibératif qu'est la concertation : les protagonistes doivent se concerter lors de leur délibération. Le type du *marchandage* a la condition thématique additionnelle que les protagonistes négocient l'achat et la vente de certaines choses. Le type des *pourparlers de paix* a la condition d'arrière-plan particulière que les négociateurs représentent les partis belligérants et qu'ils sont autorisés à conclure la paix. Finalement, les *tentatives de règlement à l'amiable* ont la condition de sincérité particulière que les négociateurs expriment alors leur volonté d'arriver à un arrangement sans animosité.

Chaque composante de type discursif détermine une condition de succès particulière correspondant à une *règle constitutive* à suivre pour tenir ce type de discours. Les protagonistes *réussissent à poursuivre un discours d'un type particulier* durant un intervalle de temps quand, premièrement, leur thème satisfait aux conditions thématiques de leur type, deuxièmement, ils atteignent le but conversationnel sur le thème avec le mode d'atteinte requis, troisièmement, ils présupposent que les conditions d'arrière-plan requises sont remplies et, finalement, ils expriment les attitudes communes requises par les conditions de sincérité. Ainsi, les locuteurs *marchandent* quand ils délibèrent sur un achat et une vente (but discursif et conditions thématiques), ils négocient en effectuant des offres, des contre-offres, des accords et des refus de commerce (mode d'atteinte du but délibératif), ils tiennent pour acquis qu'ils sont acheteurs et vendeurs potentiels (conditions d'arrière-plan), et ils expriment leur volonté commune de faire des affaires (conditions de sincérité). Quand un type de discours possède plus de composantes qu'un autre, il a des conditions de succès plus fortes. On ne peut poursuivre ce type de discours sans poursuivre un discours de l'autre

type. Ainsi les séances de marchandage, les pourparlers de paix et les tentatives de règlement à l'amiable sont des négociations, et les négociations des délibérations.

Remarquez que les conditions de succès des discours ne sont pas la somme des conditions de succès de leurs actes illocutoires successifs. Comme Roulet (1990) l'a bien souligné, un dialogue ne peut pas être décomposé directement en la suite des actes illocutoires individuels instantanés tentés lors de ses énonciations successives. Avant tout, un dialogue est une suite d'actes de discours de deuxième niveau comme des salutations, présentations, prises de position, explications, illustrations, répliques, répétitions, commentaires, résumés, prises de décision et conclusions qui durent un certain intervalle de temps. Pareilles interventions discursives sont des unités de discours d'articulation supérieure plus complexes que les simples actes illocutoires auxiliaires correspondant aux énonciations successives d'énoncés. En cela, Wittgenstein avait bien raison de dire que la *poursuite d'un discours ressemble à la pratique d'un jeu*. Prenons l'exemple d'une partie d'échecs. Rappelons que Saussure (1916) a comparé la langue au jeu d'échecs. Une partie d'échecs est plus que la suite des coups successifs des deux adversaires. Elle consiste plutôt en une suite d'échanges de plusieurs coups comme le développement initial du jeu de chaque joueur, leurs mouvements de pièces pour se concentrer et disposer d'une force supérieure sur l'échiquier, les attaques et les contre-attaques pour gagner l'avantage et les tentatives d'échec et mate. De même un discours complet tel qu'un débat théorique consiste en différents échanges linguistiques comme la présentation des participants, l'introduction à l'état de la question, les thèses de chacun, leurs explications, justifications, clarifications et illustrations, les objections, leurs critiques et tentatives de réfutation d'autres thèses, les résumés et les conclusions. D'un point de vue théorique, il faut distinguer en tout discours pareilles interventions des actes illocutoires auxiliaires plus simples tentés lors de chaque énonciation d'énoncés. Comme Searle, je pense que ces interventions sont des *actes illocutoires de deuxième niveau* dont la conduite exige une suite de plusieurs actes illocutoires instantanés. Les locuteurs font seuls ou avec d'autres plusieurs énonciations en chaque intervention.

Comme les linguistes de Genève, je distingue donc en tout discours des unités de niveaux différents. Les interventions y ont une fonction propre. À une certaine étape conversationnelle, les locuteurs comprennent qu'il leur faut développer leurs positions, qu'ils ont besoin d'argumenter, d'expliquer et de clarifier leurs idées, de se justifier ou de répliquer à des énonciations précédentes. Ainsi de nombreux types de discours imposent un ordre à la suite de leurs interventions, de même que de nombreuses interventions imposent un ordre à la suite de leurs actes illocutoires instantanés. Dans une théorie scientifique, il y a des interventions à but descriptif (la formulation des hypothèses, les justifications), à but déclaratoire (les axiomatisations et démonstrations) et à but délibératif (les expériences et tentatives de confirmations empiriques). Dans une démonstration, les locuteurs posent les prémisses avant de conclure.

Dans mon optique, la poursuite d'un type de discours est bien un *processus* dont le développement est pourvu d'une certaine *structure*. Les interlocuteurs accomplissent chacun à leur tour des actes illocutoires individuels instantanés afin de contribuer à des interventions qui jouent un rôle dans la conversation en cours.

Ainsi l'intentionnalité interlocutoire à l'œuvre dans la poursuite d'un discours est-elle supérieure à la somme des intentionnalités à l'œuvre dans les actes illocutoires individuels qui en font partie. La poursuite des discours à but conversationnel requiert seulement la réussite et parfois aussi la félicité des *actes illocutoires capitaux* (Foster, 1984) de premier et de second niveau qui sont centraux dans ces discours. La forme et le moment d'accomplissement des actes illocutoires capitaux dépendent du *type conversationnel* et du *thème* du discours. Tout d'abord, chaque discours doit contenir la suite d'interventions capitales que requiert le mode de réalisation du but conversationnel propre à son type. Beaucoup de types discursifs exigent une présentation initiale de leurs protagonistes. Lors d'une entrevue d'embauche il faut que des questions soient posées à l'interviewé et qu'il tente d'y répondre. Pareilles interventions constitutives sont elles-mêmes pourvues d'un type et d'un thème ; ce sont des *sous-discours* qui peuvent être tenus séparément. Il arrive que la suite des questions posées lors d'une entrevue soit publiée séparément dans un journal. Ensuite, chaque intervention capitale doit contenir la suite des actes illocutoires individuels qui sont centraux en cette intervention. Pour réussir une intervention il faut y accomplir des actes illocutoires individuels capitaux ayant la direction d'ajustement propre au but conversationnel. Il y a des assertions capitales en toute description comme il y a des actes engageants et directs capitaux en toute délibération. En outre les forces et les contenus des actes illocutoires individuels capitaux doivent satisfaire les conditions thématiques du type discursif de l'intervention relativement à son thème. Lors d'un marchandage, les protagonistes peuvent faire de nombreuses remarques superflues qui ne sont pas pertinentes. Mais ils doivent à tout le moins faire des offres d'achat ou de vente et essayer de parvenir à un accord commercial. Parfois, un acte illocutoire capital met fin au discours. Un accord de vente est une façon plutôt agréable de conclure un marchandage. Mais un désaccord peut aussi y mettre fin. Les délibérations d'un jury lors d'un procès doivent se terminer par un verdict. Les actes illocutoires d'un discours n'ont pas tous la même importance. Seuls les actes illocutoires capitaux importent vraiment. Leur accomplissement est essentiel. Les actes illocutoires superflus peuvent être ratés et sans pertinence. Ils n'empêchent pas les discours à but conversationnel de continuer. Les interlocuteurs qui entendent tenir pareils discours doivent cependant garder leur intention conversationnelle commune durant leur poursuite. La maxime de pertinence impose donc à présent de fortes contraintes à la structure et au développement de ces discours. Leurs protagonistes doivent en effet coopérer et suivre les règles constitutives en effectuant au bon moment des contributions capitales appropriées. Ils ne peuvent alors changer d'objectifs linguistiques sans arrêter ou faire échouer la conversation en cours.

Les discours avec une direction non vide d'ajustement ont des *conditions de satisfaction*. Comme on peut s'y attendre, leur *satisfaction* dépend de la satisfaction de leurs actes illocutoires capitaux. Les descriptions sont *exactes* quand leurs assertions capitales sont *vraies*. Les délibérations sont *respectées* quand leurs interlocuteurs *tiennent* leurs engagements capitaux et *suivent* leurs directives capitales. Enfin les discours à but déclaratoire sont *satisfaits* quand leurs interlocuteurs *réussissent* leurs déclarations capitales. Lors d'un discours les interlocuteurs raisonnent et ils peuvent changer d'attitudes et réviser leurs actes illocutoires antérieurs. Ce qui compte est leur position finale. Des interlocuteurs différents peuvent ne pas s'entendre et même se contredire. Mais souvent ils argumentent et essaient de se convaincre. Il arrive que des interlocuteurs changent de position en admettant des critiques. Parfois ils procèdent à un vote pour prendre une décision commune. Quand les protagonistes parviennent à un accord final, leur dialogue s'achève de façon heureuse. Ils s'entendent alors sur ce qui se passe dans le monde ou sur ce qu'il faut y faire. Cependant, la réussite d'un discours ne requiert pas l'accord final. Bien des marchandages se terminent sans achat. En cas de désaccord, le discours contient des actes illocutoires capitaux simultanément insatisfaisables. La théorie du succès requiert en général moins que la félicité et la bonne performance. Bien des actes illocutoires réussis lors d'une énonciation sont mal faits, faits à des moments inappropriés, défectueux, insincères ou insatisfaits. Il en va de même pour les discours entiers. Il arrive qu'une consultation d'expert soit mal faite. Les participants n'ont pas parlé des choses importantes et ils se sont basés sur des données erronées. Il se peut même que l'expert soit un imposteur. Beaucoup de discours réussis et bien faits contiennent des actes illocutoires capitaux insatisfaits. De bonnes consultations d'experts peuvent aboutir à de mauvaises expertises. Il arrive qu'une bonne annonce de publicité ne fasse pas augmenter les ventes.

Jusqu'à présent, les théoriciens de la conversation ont négligé les buts conversationnels. Beaucoup ont distingué les bons et les mauvais discours de certains genres. Mais peu ont reconnu que les dialogues sont des actes illocutoires conjoints d'ordre supérieur pourvus de conditions de félicité. Ceux qui ont analysé les conversations ordinaires (Schegloff) ont empiriquement analysé des modèles récurrents d'interaction verbale comme les règles gouvernant *le tour de parole* entre interlocuteurs. Certains linguistes (van Eemeren, Jucker, Atkinson et Drew) ont analysé des dialogues à but conversationnel d'un type particulier comme les argumentations, les procès en cour de justice, les entrevues d'embauche, les journaux télévisés et les leçons à l'école. Les philosophes des sciences ont étudié la nature du discours scientifique et des philosophes de la religion celle du discours religieux. Les logiciens ont étudié la nature des démonstrations mathématiques (la théorie de la preuve). Pareilles investigations sont restreintes mais fort utiles. D'autres ont adopté une approche plus générale ; ils ont analysé divers types de discours. Les linguistes de Genève (Eddy Roulet) utilisent le modèle hiérarchique des échanges. Hans Kamp a

construit une théorie formelle de la représentation des discours permettant de caractériser les thèmes et résoudre dynamiquement les anaphores et les pré-suppositions. Sperber et Wilson ont étudié la pertinence. Mais la plupart ont ignoré les types discursifs. J'entends renforcer l'approche théorique sur la conversation en intégrant les résultats obtenus en syntaxe, sémantique et pragmatique formelles dans une pragmatique plus générale du discours utilisant des formalismes logiques pour caractériser la structure et la dynamique des jeux de langage à but conversationnel. Voir à ce sujet mon prochain livre, *Speech Acts in Dialogue*. Une pragmatique du discours trop dépendante de la terminologie du langage ordinaire et dépourvue de vocabulaire technique et de formalisme est à éviter.

Quelles ressources le langage met-il à notre disposition afin de dialoguer ? Il existe de nombreux *verbes conversationnels*, notamment des *performatifs* (les *expositifs* de Austin) ainsi que des connecteurs discursifs comme « par conséquent ». K. MacQueen et moi avons procédé à une analyse lexicale raisonnée des verbes conversationnels de langue anglaise nommant des types discursifs sur la base de ma typologie. Il n'y a pas de correspondance biunivoque entre les types discursifs et les verbes conversationnels, tout comme il n'y a pas de correspondance biunivoque entre les forces illocutoires et les verbes performatifs. Des verbes comme « marier » et « parier » nomment à la fois un type de discours et un acte illocutoire individuel qui y joue un rôle capital. Des verbes conversationnels comme « répondre » et « conclure » nomment des interventions discursives sans but conversationnel propre. Une réponse et une conclusion peuvent avoir n'importe quel but. Certains verbes conversationnels comme « débattre » et « argumenter » sont ambigus. Les argumentations en faveur ou à l'encontre d'une thèse sont descriptives, celles en faveur ou à l'encontre d'une ligne de conduite sont délibératives. Lorsqu'un verbe conversationnel est ambigu, les types de discours qui y correspondent partagent certains traits. Ceux qui argumentent tentent de se convaincre. La volonté commune de persuader les autres protagonistes est un mode particulier de réalisation du but conversationnel propre à toute argumentation.

Afin de contribuer aux fondements de la pragmatique formelle du discours, il convient d'analyser la nature propre aux attitudes et actions conjointes en philosophie de l'esprit et de l'action, lesquelles ne sont pas réductibles, comme Searle l'a montré, à des sommes d'attitudes et d'actions individuelles. Les dialogues à but conversationnel sont des actes illocutoires conjoints de niveau supérieur dont la poursuite requiert de l'*intentionnalité constitutive conjointe* et de la *coopération* entre les interlocuteurs. Comment les protagonistes en viennent-ils à partager leur intention conversationnelle et à coordonner leurs contributions ? Il leur arrive de négocier le type conversationnel. Selon quels principes choisissent-ils leurs actes illocutoires et la façon de les accomplir ? Comment l'arrière-plan conversationnel et les énonciations antérieures interviennent-ils ? Vu notre rationalité, nous révisons et révoquons des attitudes et actes illocutoires antérieurs lorsqu'ils s'avèrent défectueux ou impossibles

à satisfaire. Les logiques paraconsistente et non monotone (da Costa, Prakke & Vreeswijk) sont utiles pour indiquer comment nous raisonnons quand des informations antérieures s'avèrent incomplètes, fausses ou même incohérentes. Pouvons-nous les utiliser ? Il convient de réfléchir à cette question.

J'entends en outre exploiter des développements récents en *logique dynamique* et *dialogique* (Lorenz, Hintikka, van Benthem, Rahman) afin de caractériser la dynamique propre aux dialogues à but conversationnel. Les dialogues de la logique dialogique traditionnelle visent à vérifier si une hypothèse initiale est ou non valide. Le premier locuteur est censé défendre son hypothèse, le second la contester. Il convient d'utiliser la logique illocutoire afin de caractériser avec exactitude les différents coups permis aux interlocuteurs en de tels dialogues. Toutes les défenses ne sont pas des assertions catégoriques. Certaines sont des assertions conditionnelles. En outre, les attaques sont des directives. Avec M. Paquette j'utiliserai la théorie des jeux (von Neumann) afin d'élaborer une *théorie qualitative de la décision* définissant une *valeur de choix* d'une façon comparable à la théorie bayésienne en termes d'ensemble hiérarchisé de préférences et de structure de probabilité (Doyle) plutôt que de fonctions de probabilité. Il s'agit de définir le succès (l'existence d'une solution) comme un équilibre de Nash pour les dialogues simples à but conversationnel. J'envisage aussi d'utiliser la théorie de la représentation du discours afin de mieux analyser le *thème conversationnel* ainsi que la *dépendance vis-à-vis de l'arrière-plan* des actes illocutoires tentés lors de chaque contribution. En considérant les énonciations dans le cadre des échanges, la pragmatique du discours offre une nouvelle perspective à la théorie de la signification qui devient plus dynamique et collective. Non seulement la théorie peut dorénavant mieux déterminer la force et le contenu des énonciations de tout genre en tenant compte des énonciations antérieures et des faits de l'arrière-plan. Mais les interlocuteurs peuvent aussi clarifier la signification d'énonciations antérieures et même leur donner une nouvelle interprétation. Enfin la théorie peut déterminer les actes illocutoires conjoints auxquels les interlocuteurs entendent contribuer en faisant leurs énonciations individuelles instantanées.

Bibliographie

- Austin, J. L. *How to Do Things with Words*, Clarendon Press, 1962.
- Belnap, N. et al. *Facing the Future Agents and Choices in our Indeterminist World*, Oxford University Press, 2001.
- Belnap, N. & Steel T. *The Logic of Questions and Answers*, Yale University Press, 1976.
- Cherniak, C. *Minimal Rationality*, MIT Press, Bradford Books, 1986.
- Cocchiarella, N. « Formally Oriented Work in the Philosophy of Language », in J. Canfield, (dir.) *Philosophy of Meaning, Knowledge and Value in the 20th Century*, Routledge, 1997, p. 39-45.
- Van Eemeren, F. H. et R. Grootendorst. *A Systematic Theory of Argumentation*, Cambridge University Press, 2004.
- Fotion, N. « Master Speech Acts », *Philosophical Quarterly*, 21, 1984, 232-243.
- Frege, G. *Logical Investigations*, Yale University Press, 1977.

- Goldman, A. I. *A Theory of Human Action*, Princeton University Press, 1970.
- Grice, H. P. *Logic and Conversation*, in P. Cole & J. Morgan (ed.) *Syntax and Semantics*, vol. 3, New York Academic Press, 1976, p. 41-58.
- Harrah, D. « The Logic of Questions » in D. Gabbay & F. Guenther (eds.) *Handbook of Philosophical Logic*, vol. 2, Reidel, 1984.
- Kamp, H. & U. Reyle. *From Discourse to Logic*, Kluwer, 1993.
- Kaplan, D. *Demonstratives*. in J. Almog et al, *Themes from Kaplan*, Oxford University Press, 1989, p. 481-563.
- Montague, R. *Formal Philosophy*, Yale University Press, 1974.
- Prior, A. N. *Past, Present, Future*, Oxford, Clarendon Press, 1967.
- Resher, N. *The Logic of Commands* Routledge & Kegan Paul, 1966.
- Roulet, E. 1992. « On the Structure of Conversation in Negotiation », in J. R. Searle et al. *On Searle and Conversation*, Benjamins, 1992, p. 91-99.
- Saussure (de), F. *Cours de linguistique générale*, Payot, 1995.
- Searle, J. R. 1992. « Conversation », in J. R. Searle et al. *On Searle and Conversation*, Benjamins, 1992, p. 7-29.
- Searle, J. R. & D. Vanderveken. *Foundations of Illocutionary Logic*, Cambridge University Press, 1985.
- Sperber, D. & D. Wilson. *Relevance*, Harvard University Press, 1986.
- Thomason, R. *Combinations of Tense and Modality*, in D. Gabbay & F. Guenther (ed.) *Handbook of Philosophical Logic*, vol. 2, 1984, p. 135-165.
- Vanderveken, D. *Meaning and Speech Acts*, vol. 1-2, Cambridge University Press, 1990-1991.
- . 1997. « La logique illocutoire et l'analyse du discours » in D. Luzzati et al. (dir.) *Le dialogique*, Peter Lang, p. 59-94
- . 1997. « Formal Pragmatics of Non Literal Meaning », *Linguistische Berichte*, 8, p. 324-341.
- . 2001. « Illocutionary Logic and Discourse Typology », *Revue internationale de philosophie*, n° 216, p. 243-255.
- . 2004. « Success, Satisfaction and Truth in the Logic of Speech Acts and Formal Semantics », in Davis, S. & B. Gillan (eds.), *Semantics A Reader*, Oxford University Press, p. 710-734.
- . 2005. « Attempt, Success and Action Generation : A Logical Study of Action Generation », in D. Vanderveken (dir.), *Logic, Thought & Action*, Springer, p. 315-342.
- . 2006. « Croyances, certitudes et rationalité minimale des agents humains », *Psychologie de l'interaction*, n° 21 et 22, p. 13-44.
- . 2006. « Fondements de la logique des attitudes » dans le numéro spécial *Language and Thought* de *Manuscrito*, vol. 29, n° 3, p. 351-98.
- Wittgenstein, L. *Philosophical Investigations*, Blackwell, 1958.
- Von Neumann & Morgenstern. *Theory of Games and Economic Behavior*, Princeton, University Press, 1944.

Les premières articulations du fonctionnement intentionnel : le projet d'un *Raumbuch* chez Husserl entre 1892 et 1894

Robert Brisart

Volume 34, numéro 2, automne 2007

URI : <https://id.erudit.org/iderudit/016988ar>

DOI : <https://doi.org/10.7202/016988ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Société de philosophie du Québec

ISSN

0316-2923 (imprimé)

1492-1391 (numérique)

[Découvrir la revue](#)

Citer cet article

Brisart, R. (2007). Les premières articulations du fonctionnement intentionnel : le projet d'un *Raumbuch* chez Husserl entre 1892 et 1894. *Philosophiques*, 34 (2), 259–272. <https://doi.org/10.7202/016988ar>

Les premières articulations du fonctionnement intentionnel : le projet d'un *Raumbuch* chez Husserl entre 1892 et 1894

ROBERT BRISART

Facultés universitaires Saint-Louis à Bruxelles

Grâce à l'édition réalisée par Strohmeyer et publiée dans la seconde partie du volume XXI des *Husserliana*, on sait qu'en 1892 Husserl entreprit plusieurs séries de recherches consacrées à la philosophie de la géométrie qui l'amènèrent, dès octobre 1893, à nourrir le projet d'un ouvrage sur l'espace dont très vite l'ampleur de la problématique lui parut telle qu'il entama aussi, dès cette date, un journal consacré à la programmation et à l'état d'avancement de ce qui devait donc devenir un *Raumbuch*, lequel néanmoins ne verra jamais le jour. Pourquoi donc, au lendemain de la *Philosophie de l'arithmétique*, alors que, comme annoncé, il était censé en rédiger le second tome, Husserl dévia-t-il ses réflexions vers la géométrie et s'attaqua plus exactement au problème de l'origine de la notion formelle d'espace ? Sans aucun doute, la réponse à cette question exige-t-elle de libérer les quelques textes de Husserl préparatoires au *Raumbuch* du cadre par trop étroit de questions seulement relatives à la pure théorie de la géométrie où les a confinés Strohmeyer dans l'introduction au demeurant richement documentée de son édition. À l'inverse, il nous semble que, pour qu'en soient mieux perçus les enjeux, ces textes qui, pour l'essentiel, traitent autant de l'espace de l'intuition que de l'espace de la géométrie, afin précisément d'en élucider le rapport, gagneraient à être rattachés à l'ensemble des manuscrits de recherches en psychologie des années 1893-1894 et que viendra ponctuer l'article sur les *Études psychologiques pour la logique élémentaire*, lesquels ont tous été édités dans le volume XXII des *Husserliana*. Ainsi s'agirait-il de montrer que la réponse à notre question liminaire peut éclairer l'itinéraire qui, de sa première œuvre, conduisit progressivement Husserl au dégagement et à l'articulation de la double vie intentionnelle de la conscience, d'une part, celle de l'intuitivité, et, d'autre part, celle de l'apophantativité. Ce travail de longue haleine entrepris donc de très bonne heure n'a pas seulement fait qu'initier l'une des tâches séminales de la phénoménologie des *Recherches logiques* de 1901, mais il a inauguré aussi tout le programme de l'œuvre à venir, tant il vrai qu'en filigrane de la théorie de la connaissance sans cesse redéployée par Husserl jusqu'au soir de sa vie, il n'a jamais été question d'autre chose que de réécrire une esthétique et une logique transcendantales pour cerner au plus près les relations généalogiques et téléologiques qu'entretiennent entre elles les deux modalités intuitives et significationnelles de l'intentionnalité.

Bien qu'il n'en soit que très indirectement question dans sa *Philosophie de l'arithmétique*¹, Husserl ne s'est pas pour autant désintéressé de la géométrie et donc de la question de l'espace à l'époque de sa première publication. Au semestre d'hiver 1889-1890, il fit en effet un cours intitulé « Questions choisies de philosophie des mathématiques » dont une longue partie fut, à la demande de ses étudiants, consacrée à un « Aperçu historique sur le développement de la géométrie », et dont le texte est reproduit dans le même volume XXI des *Husserliana*². Dans ce cours, Husserl, qui manifestement connaît de très près tous les linéaments du développement récent de la géométrie, rappelle les trois périodes qui ont marqué l'histoire de l'axiome XI des *Éléments* d'Euclide, dit encore le cinquième postulat, soit l'axiome des parallèles qui, dans sa version traditionnelle, demande que si deux lignes droites sont coupées par une troisième de telle manière que les angles intérieurs situés d'un même côté ont une somme inférieure à deux droits, elles se rencontrent du côté considéré quand on les prolonge suffisamment.

La première période fut d'abord celle où l'on pensa que cette proposition était démontrable a priori, c'est-à-dire qu'elle était nécessairement déductible des concepts qu'elle contient. Néanmoins, parce qu'en sa formulation même le cinquième postulat ressemblait davantage à un théorème plutôt qu'à un axiome³, il apparut tout aussi vite qu'il ne pouvait avoir le caractère d'une proposition immédiatement analytique, mais paraissait plutôt devoir être déductible des autres postulats d'Euclide. Et c'est d'ailleurs à cette démonstration que, depuis l'Antiquité, s'étaient employées des générations de mathématiciens, mais sans succès, au point que d'Alembert lui-même n'hésiterait pas à considérer l'irrésolution du problème posé par le postulat des parallèles comme « le véritable scandale des *Éléments* de géométrie »⁴. L'objection demeurerait donc que si ce postulat était intuitivement compréhensible, il lui manquait d'être logiquement évident et qu'en tout cas sa dérivation à partir d'éléments fondés sur l'intuition ou l'imagination ne pouvait suppléer aux exigences de sa dérivation purement logique à partir des autres axiomes.

La seconde période de l'histoire du cinquième postulat fut essentiellement marquée par les travaux de Lobatchevski, de Bolyai et de Gauss, qui, bien loin d'apporter la démonstration depuis si longtemps attendue, menaient au contraire, selon les propres dires de ce dernier, à rendre douteuse la vérité de la géométrie euclidienne traditionnelle. Gauss, par exemple, montrait tout d'abord l'indépendance du postulat des parallèles pour établir ensuite que l'on pouvait très bien s'en passer ou, comme dans la géométrie imaginaire de Lobatchevski, choisir un postulat contradictoire et le combiner avec les autres

1. Cf. Hua XII, p. 32 sq à propos de la synthèse spatiale de Lange et Baumann. *Philosophie de l'arithmétique*, trad. fr. J. English, Paris, 1972, p. 43 sq. (cité PA dans la suite).

2. Hua XXI, pp. 312-347.

3. Cf. Hua XXI, p. 318.

4. Cf. Hua XXI, p. 320.

axiomes d'Euclide pour déduire de nombreux théorèmes en contradiction avec les propositions euclidiennes. La preuve commençait ainsi d'être faite qu'une géométrie non euclidienne pouvait constituer une alternative logique à la géométrie euclidienne classique et qu'en outre, selon Gauss, elle pouvait tout aussi bien servir à décrire l'espace comme si celui-ci ne devait pas être forcément euclidien.

Mais, comme le rappelle Husserl dans ses leçons de 1889-1890, ce n'est qu'au cours d'une troisième période essentiellement marquée par les travaux de Riemann-Helmholtz que l'idée d'une géométrie non euclidienne va considérablement s'élargir jusqu'au point de se présenter désormais non plus seulement comme une alternative à la géométrie euclidienne, et tout aussi cohérente qu'elle, mais comme une théorie unifiée du plan et de l'espace à l'intérieur de laquelle la conception euclidienne n'apparaîtra plus que comme un cas particulier de multiplicité parmi d'autres. Le seul ressort de ce changement considérable fut d'abandonner l'ancienne méthode axiomatique pour procéder à une arithmétisation du plan et de l'espace grâce à laquelle leur géométrie devenait déterminable en fonction seulement de la mesure de leur courbure.

Ainsi, en matière de planimétrie à laquelle Husserl consacre un développement important dans ses leçons, trois cas deviennent envisageables parmi lesquels le plan euclidien bidimensionnel avec toutes ses propriétés classiques n'apparaît plus désormais que comme le cas extrêmement particulier d'un plan dont la courbure est de constante égale à 0. Dans un tel cas, le point le plus court entre deux points est toujours effectivement une droite par rapport à laquelle ne passe qu'une et une seule parallèle par un point extérieur à elle. Mais si l'on considère que la courbure du même plan peut être aussi de constante positive, nous obtenons alors une géométrie elliptique de ce plan dont il découle que les lignes de la plus courte trajectoire se coupent dans tous les cas et qu'il n'existe donc pas de parallèles en ce plan. Alors que si la courbure du plan est de constante négative, nous en obtenons enfin une géométrie hyperbolique où par un point donné par rapport à une droite fixée il devient possible de tracer un faisceau de parallèles. Tous ces résultats déduits de façon purement analytique montrent donc que la planimétrie euclidienne ne représente, en fonction de la courbure du plan, qu'un niveau de calcul parmi d'autres laissant apparaître la polymorphie de multiplicités à géométries variables. En outre, il est bien évident, remarque Husserl, que si l'on considère l'espace comme « l'arrière-fond en lequel sont contenues les surfaces en tant que figures et si nous le comparons à ces figures, alors apparaissent des analogies caractéristiques »⁵. Autrement dit, nous pouvons nous doter d'une expression analytique qui permet de parler d'une mesure de courbure de l'espace et considérer ainsi l'espace comme une multiplicité tridimensionnelle de mesure de courbure nulle, c'est-à-dire comme une multiplicité de type euclidien. Mais que nous ayons précisément exprimé comme égale à 0 la mesure

5. Hua XXI, p. 331.

de la courbure d'une telle multiplicité en fait aussitôt un cas particulier parmi d'autres déterminations analytiques de mesure de courbures constantes et, par-tant, de mesure de courbures variables. Si l'on tente dès lors de traduire la notion d'espace à partir de la généralisation de toutes les différentes façons de le déterminer analytiquement en fonction de la mesure de sa courbure, on peut parler « de multiplicités à n dimensions dont les éléments ou les points sont déterminés par n coordonnées »⁶.

Ce qui de toute évidence dut retenir l'attention de Husserl dans ce développement récent de la géométrie, et plus encore dans la réévaluation des géométries non euclidiennes faite par Riemann, c'est sans aucun doute l'introduction du concept de multiplicité pour traiter géométriquement de l'espace. Lorsque, à l'extrême fin de ses leçons de 1889-1890, Husserl dit que, sur le plan purement analytique, la théorie riemannienne de l'espace possède « une valeur durable »⁷ (*bleibenden Wert*), ce qu'il maintiendra notamment dans une lettre à Brentano du 29-12-1892 où il répète qu'elle n'est pas sans avoir « un contenu plein de valeur »⁸ (*Wertvolle Gehalt*), on devine assez facilement — même s'il n'en dit pas grand chose — que cette valeur tient essentiellement au mode de traitement riemannien des multiplicités avec lesquelles justement Husserl se débat à l'époque dans le cadre qui lui est propre, à savoir celui d'une réflexion sur les multiplicités en arithmétique et surtout sur la logique de leur extension dans le but de fournir une détermination rigoureuse à tout le champ de l'*arithmetica universalis*. À cet égard, toute l'inventivité de Riemann, qui ne pouvait qu'intéresser Husserl dans ses propres recherches en arithmétique, tient essentiellement à sa *Mannigfaltigkeitslehre*, laquelle consiste tout d'abord à montrer qu'à l'inverse d'un ensemble, une multiplicité ne détermine pas les éléments qu'elle rassemble en fonction de leur matière, c'est-à-dire des propriétés qui leur sont communes, mais qu'elle détermine seulement ses éléments par des formes déterminées de connexion purement analytiques dont l'expression est algébrique. En outre, ce que Riemann a également indiqué, et qui ne pouvait qu'apparaître comme du plus haut intérêt à Husserl alors confronté au problème de l'extension arithmétique, c'est que le passage d'un genre de multiplicité à un autre très différent est obtenu lui aussi de façon purement formelle moyennant, par exemple, l'expression analytique des variations du degré de courbure de l'espace ou du plan en géométrie.

Outre le fait de libérer ainsi le concept géométrique d'espace de tout recours intuitif à la spatialité, l'arithmétisation de la géométrie, dont Riemann fut l'un des pionniers, eut pour conséquence directe de rendre les objets ou plus précisément les champs d'objets géométriques totalement corrélatifs à des procédures d'engendrement réglées par des opérations déterminées. En d'autres mots, les champs d'objets, c'est-à-dire les multiplicités, pouvaient dès

6. Hua XXI, p. 333.

7. Hua XXI, p. 347.

8. Hua Dok III, 1, p. 11.

lors apparaître comme autant de régions purement formelles seulement déterminées par des règles analytiques. Il en découlait encore qu'un espace spécifiquement géométrique n'était plus qu'une multiplicité dont la dimension perdait toute signification spatiale pour ne garder que la signification algébrique d'un certain nombre de variables ou de signes numériques sur lesquels les opérations ne pouvaient plus être assimilées à des constructions spatiales, mais seulement à des fonctions algorithmiques. Nul doute qu'aux yeux de Husserl cette arithmétisation de la géométrie non seulement revenait à en faire un bon échantillon de l'*arithmetica universalis*, mais de surcroît pouvait aussi servir d'étalon pour déterminer ce à quoi elle devait dès lors ressembler. Ce que reconnaîtra très explicitement Husserl, notamment dans une lettre à Carl Stumpf datée de février 1890, en la définissant comme « un système de signes et d'opérations sur ces signes »⁹, c'est-à-dire, ajoutera-t-il, « en harmonie avec Helmholtz »¹⁰, sous-entendu donc aussi de façon assez proche du formalisme analytique de Riemann lui-même.

S'il y a donc aux yeux de Husserl quelques bonnes raisons pour reconnaître une valeur mathématique à la géométrie de Riemann, il n'en reste pas moins que l'autre conclusion du cours de 1889-1890 est que d'un point de vue philosophique elle reste « sans valeur »¹¹. Autrement dit, comme le répétera Husserl dans la même lettre à Brentano de décembre 1892 où il reconnaissait à la géométrie de Riemann/Helmholtz un « contenu plein de valeur », le défaut qui leur reste néanmoins imputable est de « n'être parvenus à la moindre clarté philosophique sur les fondements de leur science »¹². On ne s'étonnera pas de l'ambivalence manifeste de l'appréciation de Husserl, puisqu'elle est en quelque sorte indexée sur la propre ambivalence de la *Philosophie de l'arithmétique* eu égard à laquelle doit être bien sûr recontextualisée l'approche husserlienne de la géométrie entre 1890 et 1892.

On le sait, le premier ouvrage de Husserl est marqué d'une rupture interne qui effectivement finit par l'orienter vers une conception purement formaliste de l'arithmétique. Étant d'abord parti de l'idée, qui le dirigeait encore dans son *Habilitationsschrift* de 1887, selon laquelle c'est le concept de numération qui constitue le fondement de l'arithmétique, Husserl fait volte-face au chapitre X de l'ouvrage pour remarquer que, si l'arithmétique est essentiellement faite d'opérations sur des nombres, alors il ne se peut que ceux-ci soient pris au sens propre, c'est-à-dire comme des concepts pour des quantités dont nous serions dès lors chaque fois contraints d'avoir une représentation. Ce sont donc les opérations et, avec elles, l'extension de l'arithmétique qui amènent Husserl à reconsidérer le matériau sur lequel elle procède en admettant désormais qu'il ne s'agit que de signes dont nous n'avons qu'une représentation impropre, tant

9. Hua XXI, p. 246.

10. Hua XXI, p. 247.

11. Hua XXI, p. 347.

12. Hua Dok III, 1, p. 11.

il y a impropriété à parler en ce cas d'une représentation, c'est-à-dire d'une intuition quelconque.

Les conséquences de cette bifurcation sont au moins triples. Tout d'abord, du fait que l'arithmétique relève désormais d'un processus exclusivement significatif, la compréhension de la formation d'un système de nombres semble donc ne requérir plus rien d'autre qu'une sémiotique destinée à montrer comment l'arithmétique tout entière se constitue par algorithmes, c'est-à-dire par procédures purement techniques de calcul permettant la production de certaines catégories de nombres à l'intérieur de certaines chaînes opératoires. Ensuite, du fait que l'extension des nombres et, avec elle, le développement progressif de l'*arithmetica universalis* paraissent désormais ne résulter que de l'extension des techniques algorithmiques engendrant de nouvelles formations de signes chaque fois formellement dérivées des formations plus élémentaires, comme l'a écrit Husserl à Stumpf dans la lettre de 1890 déjà citée, l'*arithmetica universalis* peut dès lors être comprise « non pas comme une science, mais comme une partie de la logique formelle, qu'[il] définir[a] [lui]-même comme un art des signes¹³ ». Troisièmement, enfin, la critique du formalisme qui était au départ adressée notamment à Helmholtz tant que Husserl considérait que le fondement de l'arithmétique ramenait au concept de numération devient désormais sans objet, puisque le recours aux processus significatifs revient à adopter une position formaliste et nominaliste qui notamment transparaît très clairement dans cet extrait de 1889-1890 :

Le concept de numération n'autorise aucune extension du tout ; ce qui est en extension et autorise l'extension, c'est seulement la technique arithmétique. C'est elle qui connaît et elle qui forme les signes négatifs, imaginaires, irrationnels et fractionnaires qui servent à rendre le calcul plus complet et qui de ce point de vue ont une signification logique importante. Mais ce qui leur manque, c'est tout contenu conceptuel par delà l'algorithmique¹⁴.

Toutefois l'ambivalence que nous pointions à propos de la *Philosophie de l'arithmétique* ne consiste pas en ce revirement en faveur du signe et au détriment du concept ou de toute signification par delà la technique arithmétique. Elle tient bien davantage au fait que la *Philosophie de l'arithmétique* ne s'épuise pas dans la sémiotique de l'*arithmetica universalis*, ni donc n'en reste à son formalisme, mais que, en tant que philosophie justement, elle ne renonce pas à cette question que l'on aurait pu croire définitivement abandonnée et que pourtant Husserl pose ouvertement au chapitre XI de l'ouvrage : « sur quoi se fonde et s'appuie la symbolisation ?¹⁵ ». Si l'invention husserlienne de la phénoménologie a quelque part une origine, c'est bel et bien ici. En voici l'idée : pour être complète une théorie de la science, et quelle que soit la science en

13. Hua XXI, p. 248.

14. Hua XXI, pp. 42-43.

15. Hua XII, p. 220 ; PA, p. 241.

question, ne saurait se clore sur elle-même par une exposition de ses concepts fondamentaux ou de ses principes. De ceux-ci, il reste encore à avérer l'évidence que nous en avons par réflexion sur les actes où ces concepts et ces principes trouvent un étayage ou un fondement. En l'occurrence, si les représentations signitives ou symboliques échouent à procurer l'objet en une donation primaire, et c'est bien en ce sens qu'elles sont des représentations impropres, elles n'en sont pas moins des représentations objectivantes dans la mesure où, par la médiation du signe, c'est bien un objet et, en arithmétique, plus précisément une quantité ou une multiplicité dénombrable qui est signifiée et, par là, visée à vide. La question proprement philosophique est dès lors de savoir comment nous pouvons nous rapporter à quelque chose comme une multiplicité, d'abord la plus simple, c'est-à-dire matérielle ou sensible, pour poser sur cette base la question de la modalité selon laquelle peuvent nous être données des multiplicités formelles comme celles auxquelles nous avons affaire en arithmétique. Ce travail spécifiquement philosophique est, comme le dit encore Husserl dans la *Philosophie de l'arithmétique*, « l'étude de fond de la formation et de la signification des représentations symboliques sur le domaine du nombre¹⁶ ».

Il ne nous importe pas ici d'analyser la façon dont Husserl a entrepris ce travail en 1890, mais, pour en revenir à notre propos, d'avancer que c'est assurément l'absence d'un tel travail dans la théorie géométrique de Riemann-Helmholtz, ou du moins sa confusion avec un travail de construction d'un tout autre genre, qui justifie qu'à la même époque Husserl dénie à cette théorie la moindre valeur philosophique. Dans le cours de 1889-1890, la critique adressée à Riemann au nom même de la philosophie, mais qui, dans le chef de Husserl, s'étend aussi aux théories de Gauss et de Helmholtz, revient à leur contester le bien-fondé qu'il y a à prétendre fonder la géométrie sur une théorie de la courbure. Les recherches de Riemann, rappelle Husserl, visent à fournir au moyen d'une théorie générale les signes caractéristiques par lesquels une multiplicité plane se distingue d'une multiplicité courbe. Ce qui suffit à concevoir les multiplicités planes comme un cas tout à fait particulier parmi une infinité de cas possibles, puisque la planéité renvoie à une valeur déterminée, parmi d'autres, de la mesure de la courbe, celle précisément où cette valeur est égale à 0. Contre quoi Husserl fait valoir que n'importe quelle valeur de courbure d'un plan ou d'un espace n'est déterminable que relativement à la planéité du plan ou de l'espace. Autrement dit, il existe une priorité logique de la multiplicité plane par rapport à toute multiplicité courbe et par conséquent une priorité logique aussi de la géométrie euclidienne par rapport aux géométries non euclidiennes.

C'est une erreur de croire, écrit Husserl, pouvoir déduire le concept de multiplicité plane à partir de celui de multiplicité courbe. Une telle tentative n'aurait

16. Hua XII, p. 217; PA, p. 238.

de sens et de signification que si la seconde avait une priorité logique sur la première. Mais ceci est si peu le cas que c'est le concept de multiplicité courbe qui présuppose celui de multiplicité plane¹⁷.

Ainsi, explique Husserl, les mathématiciens, qui aujourd'hui ne se forment pas à la priorité logique de la planéité, se comportent comme s'ils voulaient déduire le concept d'unité du concept de nombre sans considérer que, chaque nombre se composant d'unités, chaque concept de nombre présuppose le concept d'unité. C'est exactement le même type de cercle logique que de prétendre fournir le fondement de la géométrie par une théorie de la courbure. Husserl termine enfin le cours de 1889-1890 par la présupposition que « l'on peut effectivement former le concept de multiplicité à n dimensions et que l'on peut donc concevoir des concepts d'espace supérieurs au nôtre. Mais, même de ce point de vue, ajoute-t-il, la recherche de Riemann est lacunaire »¹⁸. N'est-ce pas finalement dans la mesure où ce qui n'est encore qu'une construction, et ne peut être que cela, est pensé à tort par Riemann comme une fondation, se faisant donc passer pour ce qu'elle n'est pas, à savoir une philosophie de la géométrie.

Toutefois, on s'en aperçoit également, la critique de la valeur philosophique de la géométrie riemannienne qui est menée par Husserl en 1890 reste encore, pour ainsi dire, bien en deçà des exigences fondationnelles qui sont celles de la philosophie elle-même. En effet, l'argumentaire philosophique de Husserl contre Riemann s'y appuie uniquement sur la priorité logique de la géométrie euclidienne et donc sur son rôle fondateur par rapport à toute construction de géométries non euclidiennes. Mais qu'est-ce qui, indépendamment de cette priorité logique par rapport à toute construction métagéométrique autorise à penser que le concept de multiplicité tridimensionnelle de type euclidien s'impose en soi comme le concept formel que nous avons prioritairement de l'espace ? Reposée en ces termes, la question fondationnelle de la géométrie ne peut plus se suffire de la logique mais en appelle à une généalogie de cette logique. C'est, peut-on dire, à tenter de réaliser cette tâche d'élucidation philosophique proprement dite que se consacrera Husserl deux ans plus tard, en 1892-1893, dans les analyses menées en vue de son *Raumbuch*.

Selon la planification établie par Husserl dans le journal qu'il tenait sur l'état d'avancement du *Raumbuch*¹⁹, celui-ci avait prévu trois groupes de recherches sur l'espace : psychologiques, logiques et métaphysiques, en donnant manifestement plus de poids au premier groupe, puisque, avec les recherches psychologiques, il y allait de la représentation pré-scientifique de l'espace dont il s'agissait essentiellement de décrire comment c'est de son idéalisation que procède la représentation scientifique en géométrie. Avant donc que d'aborder des recherches logiques notamment consacrées au rapport entre espace euclidien

17. Hua XXI, p. 345.

18. Hua XXI, p. 347.

19. Hua XXI, pp. 404-406.

et non euclidien, puis des recherches métaphysiques essentiellement consacrées, elles, au rapport entre les multiplicités géométriques et l'espace réel ou effectif, la question était tout d'abord, pour Husserl, d'examiner les divers contenus psychologiques que nous pouvons avoir de l'espace.

Procéder de la sorte, c'est donc bien en venir à la généalogie de la logique géométrique pour tenter de rétablir les liens génétiques existant entre la représentation intuitive et les concepts purement formels qui caractérisent le champ de l'idéalisation scientifique en général. Aussi cette question, qui avait déjà animé les premières recherches de Husserl en arithmétique et dont il avait très vite pressenti le besoin de l'étendre à tout le domaine de la logique dès après sa critique de Schröder en 1891, l'occasion lui était maintenant donnée de la redéployer sur le terrain de la géométrie, c'est-à-dire en ce domaine des mathématiques où le problème prend une acuité exemplaire de savoir comment, à partir de ces contenus abstraits que sont les moments spatiaux de la représentation intuitive pré-scientifique, on passe à cette représentation purement conceptuelle qu'est l'espace géométrique. En somme, la question était de comprendre comment à partir de certains abstracta matériels s'engendrent les abstracta formels de la géométrie : l'espace comme pure multiplicité et ses composants formels tels que points, droites, plans, etc. C'est donc bien la problématique centrale de la *Philosophie de l'arithmétique* qui est ici remise sur le métier, mais à présent dans le but d'éclairer ce qui ne l'y était que très imparfaitement encore, à savoir les étapes successives que la conscience intentionnelle doit franchir pour qu'à partir des couches les plus élémentaires de la représentation intuitive puissent se former des représentations conceptuelles où se noue le rapport à des objets purement formels.

Exprimée de la sorte, la problématique de fond du *Raumbuch* — celle en tout cas sur laquelle Husserl a visiblement le plus travaillé en laissant quasiment pour compte les recherches de logique et de métaphysique qui devaient lui faire suite — se rattache clairement, on le voit, au groupe des analyses psychologiques entreprises par ailleurs entre 1892 et 1894, sans que l'on ne voie plus très bien si ces analyses furent menée pour étayer une philosophie de la géométrie ou si celle-ci fut menée pour exemplifier une psychologie essentiellement consacrée à l'articulation de la vie intentionnelle en sa bipolarité, intuitive et conceptuelle. Sans aucun doute est-ce la seconde interprétation qui est la bonne, puisque si le *Raumbuch* ne verra jamais le jour, c'est notamment en raison de l'étroitesse de sa problématique spécifique par rapport à celle qui, dès 1894, engageait la psychologie husserlienne sur la voie d'une vaste réflexion où se trouvait engagée la question des statuts de l'objet intentionnel. Mais il n'en reste pas moins que l'article publié par Husserl en 1894 sous le titre *Études psychologiques pour la logique élémentaire*, lequel vient en quelque sorte ponctuer la série d'analyses réalisées par Husserl à cette époque, se clôt en relevant que, si ces analyses ont certes un intérêt purement psychologique, elles ont surtout une importance capitale pour « l'élucidation de l'origine de la représentation de l'espace » où justement « on tombe sur les

fonctions intuitives et re-présentatives qui ici agissent partout ensemble et sans la compréhension desquelles on tâtonne désespéré dans l'obscurité²⁰».

Or, disait Husserl dans le premier texte de 1892 préparatoire au *Raumbuch*, élucider la différence entre représentation intuitive et représentation conceptuelle nécessite d'en revenir de nouveau à la différence entre représentation propre et représentation impropre et aussi à la différence entre représentation abstraite et représentation concrète²¹. Ce sont précisément ces deux couples de questions qui font la substance de l'article de 1894, comme si celui-ci avait été entièrement conçu en réponse à l'une des premières questions prescrites par la philosophie de la géométrie. Cet article est donc fait de deux parties : la première, la plus courte, traite de la distinction entre contenu abstrait et contenu concret de la représentation ; la seconde, la plus longue et la plus importante aussi, porte sur « l'intuition » (*Anschauung*) et la « représentation » (*Repräsentation*).

Mais à la détermination de quels paramètres pouvaient donc bien servir les deux axes (concret-abstrait, intuitif-signitif) successivement envisagés en de telles analyses ? Pour le voir, contentons-nous ici d'esquisser à grands traits quelques-unes des thèses emblématiques de ces recherches sur l'origine de la géométrie qui bien sûr n'ont eu de sens dans l'ordre des préoccupations de Husserl que dans la mesure où, contrairement à Frege par exemple, jamais il n'a considéré les sources de la géométrie comme seulement intuitives et donc comme étrangères à la logique. Tout au contraire, pour Husserl, la géométrie est une science a priori²² dont les fondements sont constitués de concepts purement formels qui possèdent leur propre système de lois formelles, lui-même transposable — ainsi que Hilbert en fera un peu plus tard la preuve dans ses *Grundlagen der Geometrie* de 1899 — en un ensemble limité et exhaustif de groupes d'axiomes pouvant enfin donner lieu aux multiples inférences logiques ou aux théorèmes de la géométrie tout entière.

Toutefois, le caractère apriorique de la géométrie ne saurait justifier que l'on doive traiter l'espace géométrique de façon purement déductive comme s'il ne relevait que d'une pure construction intellectuelle. Certes, l'espace géométrique est la résultante d'une activité logique, et c'est pourquoi il est un espace seulement pensable et non pas un espace représentable lié donc à l'intuition, ou alors seulement à titre tout à fait occasionnel et inessentiel d'illustration figurative des propositions géométriques²³. Mais il reste que cette activité logique n'est elle-même qu'une idéalisation conceptuelle de l'expérience intuitive de l'espace²⁴. En d'autres mots, ce sont des configurations et des relations

20. Hua XXII, pp. 190-191 ; *Articles sur la logique*, trad. fr. J. English, Paris, 1975, pp. 162-163.

21. Hua XXI, p. 262.

22. Cf. Hua XXI, pp. 296-300.

23. Cf. Hua XXI, p. 271.

24. Hua XXI, p. 286 sq.

spatiales intuitionnées qui, pour ainsi dire, servent de substrats à l'idéalisation conceptuelle en géométrie.

Or de quelle nature est ce substrat intuitif ? Nous avons l'intuition d'étendues finies, de surfaces déterminées, de lignes qui les séparent en les coupant, et même d'une certaine continuité dans le champ de perception : ce sont là autant de moments abstraits partiels des divers contenus d'intuition. Par modification d'intérêt, l'intuition peut se concentrer sur ces moments abstraits en passant de l'un à l'autre de manière à provoquer un déroulement d'intuitions successives qui d'une portion d'espace finie renvoie à une autre, et ainsi de suite jusqu'à ce que, au bout du compte, soit reconnu l'espace global (*Gesamtraum*)²⁵. Comment cela se produit-il ? Les analyses psychologiques portant sur les déroulements d'intuitions réalisées dans un manuscrit de 1893 intitulé « Intuition et représentation, intention et remplissement »²⁶ ont précisément servi à le montrer, un contenu primaire ou partiel est chaque fois appréhendé comme renvoyant à un autre qui fait partie de son halo, de telle façon que chaque contenu partiel intuitionné pour lui-même est investi d'une intention qui trouve son remplissement dans le passage à l'intuition d'un autre contenu partiel, et ainsi de suite jusqu'à ce qu'au bout du compte ce soit l'unité objective de la chose qui est reconnue. Pour qu'ait lieu ce genre de déroulement continu d'intuitions, au cours duquel chaque contenu intuitif partiel se double d'une charge signitive renvoyant à un autre contenu partiel selon un processus temporel qui sans cesse fait devenir autre le contenu intuitionné, aucune médiation conceptuelle n'est requise. Par contre cette médiation doit bien l'être pour que soit reconnue l'unité objective vers laquelle est finalement tendu comme vers son but tout le déroulement d'intuitions. Car où cette unité est-elle reconnue explicitement ? Non pas dans l'intuition proprement dite, répond Husserl, mais dans les jugements qui l'accompagnent et où, sous une forme conceptuelle, un contenu global se trouve exprimé comme représentant l'unité objective à laquelle sont rattachés, comme autant de ses déterminations, les divers contenus partiels rencontrés dans la suite des intuitions partielles. C'est pourquoi « l'unité objective est une unité par jugement, non pas par simple intuition, mais sur fondement d'intuition²⁷ ». Telle que reprise dans les textes du *Raumbuch*, l'idée est donc que la totalité formée par le fusionnement continu des différentes parties intuitionnées, à savoir ici la totalité spatiale, est une unité objective qui, elle, n'est pas intuitionnée comme telle, mais simplement intentionnée. C'est pourquoi, explique Husserl, il y a lieu de « pratiquer une certaine *reservatio mentis* lorsque nous parlons de l'espace intuitif », car tel qu'il est reconnu par la conscience commune ou pré-scientifique, « l'espace comme totalité n'est pas intuition, mais déjà une unité

25. Hua XXI, p. 276.

26. Cf. Hua XXII, 275-283 ; *Sur les objets intentionnels*, trad. fr. J. English, Paris, 1993, pp. 222-232, (cité OI).

27. Hua XXII, p. 282 ; OI, p. 231.

logique qui dépend de la représentation conceptuelle et se produit par jugement sur le donné intuitif²⁸ ».

Si Husserl maintient donc fermement l'idée d'une fondation de l'espace géométrique dans l'espace de la représentation, elle ne lui semble pas pour autant attenter au caractère a priori ou pur de la géométrie. Cette fondation, en effet, n'est pas directe, mais est stratifiée en plusieurs couches de vécus intentionnels dont la restitution du mode de transformation montre comment la représentation elle-même passe du registre intuitif au registre conceptuel et, de façon corrélatrice, comment les objets auxquels la conscience a successivement affaire passent du registre matériel au registre idéal et à l'intérieur même de celui-ci au registre purement formel. Derechef, pour l'expliquer, rien ne requiert le recours à des opérations mentales de seconde main au sens classique de l'abstraction, car ces transformations disent elles-mêmes la dynamique ou la vie de la conscience qui, dans le déroulement de son rapport intuitif et originaire aux objets, se voit téléologiquement conduite à sa propre modification en un mode d'exercice intégralement nouveau de l'activité intentionnelle, lequel relève désormais de la représentation par concept ou par signification, dont l'unité idéale, réitérable en une multitude de jugements, traduit le rapport à un autre mode de prestation de l'objet où celui-ci se donne précisément telle une idéalité.

Pour Husserl, c'est ce destin de la conscience que sert à exemplifier la genèse de la géométrie, plus précisément encore dans le cas où l'idéalité objective revêt le caractère du formel pur. En guise d'intuitions de l'espace n'existent en fait que celles qui portent sur les moments spatiaux non autonomes à l'intérieur d'intuitions plus larges, avec cette particularité pourtant que chaque chose possédant son étendue, le moment spatial n'est pas seulement un abstractum matériel, mais aussi un substrat pour toute unité d'intuition, donc un « fondement abstrait²⁹ ». Mais tout aussi peu qu'il est une intuition, l'espace en tant que tel, c'est-à-dire comme totalité infinie, n'est pas du même ordre que ses parties finies qui, elles, sont matériellement visibles ; il est pour la conscience commune une « figure idéale » (*ideelles Gebilde*) seulement pensable par représentation conceptuelle³⁰. En l'espèce, faudrait-il ajouter, parce que le conceptuel est ici d'une généralité qui ne laisse plus rien filtrer de la matérialité première des moments spatiaux, il exprime le rapport intentionnel de la pensée à une objectivité qui n'est autre qu'une forme générale du domaine des objets matériels en général. Bref, c'est à la représentation pré-scientifique elle-même qu'il appartient de se rapporter à un espace dont l'existence idéale et purement formelle le détache des intuitions premières du spatial. Si, par suite, la géométrie peut se voir conférer le titre de science purement a priori, c'est seulement dans la mesure où elle n'est que « l'élaboration logique »³¹ de cette

28. Hua XXI, p. 283.

29. Hua XXI, p. 276.

30. Hua XXI, pp. 283-284.

31. Hua XXI, p. 271.

objectivité idéale de l'espace dont elle tire ses propres définitions élémentaires et cherche à dégager analytiquement les lois formelles, moyennant l'établissement d'un ensemble exhaustif de groupes d'axiomes.

Concluons. Dans un texte intitulé « Symbole de la géométrie » datant de décembre 1894, c'est-à-dire de l'époque où le projet du *Raumbuch* semble avoir été abandonné, puisque ce texte fut le dernier que Husserl lui consacra, celui-ci écrit que « un exposé scientifique de la géométrie doit rigoureusement opposer le symbole et la chose en question (*Sache*) ; il doit par des axiomes rigoureux et en se fondant sur eux par un processus purement formel, exclure les erreurs et montrer dans ses limites le processus intuitif³² ». Écartant ainsi le symbolique de la géométrie, car le symbolique ne traduit jamais que l'intrusion en elle de l'intuition, celle précisément des représentations figuratives qui même là où elles sont possibles ne constituent jamais une preuve formelle, Husserl entend bien accentuer ainsi l'autonomie logique de la géométrie et rappeler, comme il en avait déjà défendu l'idée dans un texte de 1893, qu'elle est une « science purement apriorique³³ ». Mais Husserl n'entendait certainement pas remettre pour autant en question que cette logique a une genèse et qu'elle remonte à notre expérience de l'espace. Aussi appartient-il à l'analyse psychologique — à laquelle Husserl ne va pas tarder de préférer le nom de phénoménologie — de pratiquer une eidétique de la conscience d'où puissent être dégagées les transformations successives de l'exercice intentionnel à travers lesquels sont avérées les différentes modalités de la donation objective, depuis le niveau le plus primitif de la réalité matérielle jusqu'au niveau plus élevé des idéalités. Ce travail d'abord entrepris sur le terrain de l'arithmétique, Husserl l'aura bien mieux approfondi encore sur celui de la géométrie, en montrant comment cette science marquée au sceau du pur a priori ne doit son origine qu'à l'abstraction comprise au sens radicalement nouveau d'une réflexion idéalisante sur les actes par lesquels se forment les concepts généraux et où se noue la relation significationnelle à un type d'objectivité regroupant ce que l'on peut appeler les abstracta formels. Mais en raison de cette nécessaire rétro-référence aux actes de significations générales et donc au mode de donation des objets généraux correspondants, les idéalizations de la géométrie, comme le dit Husserl, « n'ont rien d'arbitraire³⁴ ». C'est pourquoi, en matière de géométrie, Husserl continue de récuser toute conception purement analytique qui, comme chez Gauss, Helmholtz ou Riemann, ne vise qu'à déterminer de façon métrique n'importe quel espace considéré comme une multiplicité arbitraire à n dimensions. Pour Husserl, au contraire, c'est l'espace tridimensionnel euclidien qui s'impose d'abord à l'idéalisation géométrique, et ce n'est qu'à partir de lui que peuvent être échafaudées ensuite des métagéométries opérant sur des espaces de dimensions supérieures. Par cette restriction fondationnelle

32. Hua XXI, p. 295.

33. Hua XXI, p. 296.

34. Hua XXI, p. 308.

de la géométrie à l'a priori de l'espace euclidien, Husserl n'entend bien sûr pas s'opposer aux différents développements analytiques qui, à partir de lui, peuvent mener à la construction d'espaces non euclidiens ; mais il entend à tout le moins maintenir les liens génétiques par lesquels la géométrie trouve ses conditions de possibilité dans la représentation pré-scientifique de l'espace en comprenant par là non pas tant la représentation intuitive mais, sur la base de son déroulement, la représentation conceptuelle où, dans le développement du fonctionnement intentionnel, l'espace s'impose toujours aux jugements comme tridimensionnel.

La génération spontanée et le problème de la reproduction des espèces avant et après Descartes

Justin E. H. Smith

Volume 34, numéro 2, automne 2007

URI : <https://id.erudit.org/iderudit/016990ar>

DOI : <https://doi.org/10.7202/016990ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Société de philosophie du Québec

ISSN

0316-2923 (imprimé)

1492-1391 (numérique)

[Découvrir la revue](#)

Citer cet article

Smith, J. E. H. (2007). La génération spontanée et le problème de la reproduction des espèces avant et après Descartes. *Philosophiques*, 34 (2), 273–294. <https://doi.org/10.7202/016990ar>

Résumé de l'article

Dans cet article je mets en évidence quelques problèmes conceptuels importants posés par le prétendu phénomène de la génération spontanée, en montrant comment ils étaient liés historiquement à la question théorique des origines et de l'ontologie des espèces biologiques. Au XVI^e et XVII^e siècle tout particulièrement, la possibilité que des formes organiques soient générées dans la matière inorganique supposait la possibilité que le hasard gouverne non seulement l'apparition d'une anguille ou d'une souris, mais qu'il gouverne l'apparition originelle de leurs espèces mêmes. En outre, dans la conception de la reproduction sexuelle que le mécanisme parvient à répandre, toute génération, des êtres humains aussi bien que des anguilles, menace de ne plus s'expliquer autrement que par ce que Descartes appelle « les causes mineures ». Ainsi, comme je tenterai de le prouver, les problèmes théoriques que la génération spontanée, telle que le début de la modernité la concevait, posaient à la compréhension de l'ontologie des espèces, n'étaient pas essentiellement différents de ceux soulevés par l'explication mécaniste de la reproduction sexuelle, et si nous n'accordons pas à ce fait l'attention nécessaire, nous perdons de vue, je pense, un facteur important dans le rejet ultime de la génération spontanée.

La génération spontanée et le problème de la reproduction des espèces avant et après Descartes

JUSTIN E. H. SMITH
Université Concordia

RÉSUMÉ. — Dans cet article je mets en évidence quelques problèmes conceptuels importants posés par le prétendu phénomène de la génération spontanée, en montrant comment ils étaient liés historiquement à la question théorique des origines et de l'ontologie des espèces biologiques. Au XVI^e et XVII^e siècle tout particulièrement, la possibilité que des formes organiques soient générées dans la matière inorganique supposait la possibilité que le hasard gouverne non seulement l'apparition d'une anguille ou d'une souris, mais qu'il gouverne l'apparition originelle de leurs espèces mêmes. En outre, dans la conception de la reproduction sexuelle que le mécanisme parvient à répandre, toute génération, des êtres humains aussi bien que des anguilles, menace de ne plus s'expliquer autrement que par ce que Descartes appelle « les causes mineures ». Ainsi, comme je tenterai de le prouver, les problèmes théoriques que la génération spontanée, telle que le début de la modernité la concevait, posaient à la compréhension de l'ontologie des espèces, n'étaient pas essentiellement différents de ceux soulevés par l'explication mécaniste de la reproduction sexuelle, et si nous n'accordons pas à ce fait l'attention nécessaire, nous perdons de vue, je pense, un facteur important dans le rejet ultime de la génération spontanée.

ABSTRACT. — In this article I shall draw out some of the important conceptual problems posed by the purported phenomenon of spontaneous generation, showing how these problems were historically connected with the theoretical question of the origins and nature of biological species, and above all with the problem of their boundaries. In the 16th- and 17th- centuries in particular, the possibility of organic forms arising from inorganic matter carried with it the possibility that chance governs not only the emergence of an individual eel or mouse, but indeed governs the original emergence of the mouse- and eel-kinds. Moreover, on the newly ascendant mechanist understanding of sexual reproduction, all generation, whether of eels or of horses and men, now threatened to be exhaustively accounted for in terms of what Descartes called 'minor causes'. Thus, I argue, the sort of problems that spontaneous generation was perceived to bring about for the early modern understanding of the ontology of species were not in principle any different from the problems posed by the mechanist account of sexual reproduction, and if we fail to note this, I show, we overlook an important factor in the eventual demise of spontaneous-generation theory.

Introduction

Dans cet article, je porte mon attention sur quelques-uns des importants problèmes conceptuels posés par le prétendu phénomène de la génération spontanée, en montrant comment ceux-ci sont liés à la question théorique des origines et de l'ontologie des espèces biologiques, et en premier lieu au problème de leurs frontières. Je défends l'idée importante que la distinction conceptuelle

de l'histoire de la théorie de la génération ne s'est jamais trouvée entre les animaux sexuellement et spontanément générés, mais plutôt entre les animaux générés par la reproduction cyclique de la forme, d'un côté, et par ce que Descartes appelle les « causes mineures », d'un autre côté. Pour un hylomorphiste tel qu'Aristote, toute reproduction est formelle, qu'elle soit sexuelle ou spontanée ; pour un mécaniste tel que Descartes, toute reproduction, y compris la reproduction sexuelle, est mécanique.

Dans la première section, je fais un rapide survol du problème de l'ontologie des espèces et de la génération spontanée chez Aristote, comme il a été traité dans la littérature secondaire depuis l'attribution bien connue, par David Hull, du fixisme des espèces [*species fixism*] au philosophe grec. Dans la deuxième section, je considère les candidats plus probables des origines de ce fixisme, à savoir les écrivains d'histoire naturelle du début de la période moderne qui étaient préoccupés par la classification des types morphologiques à l'intérieur des systèmes taxonomiques. Dans la troisième section, je discute de la manière dont la nouvelle ontologie émergente des espèces au XVI^e siècle a accueilli la théorie très répandue de la génération spontanée qui est apparue durant le Moyen Âge et qui était largement aristotélicienne. Selon cette théorie, on ne refuse pas l'imposition d'une certaine forme par un agent, mais l'on considère plutôt que ce sont les cieux qui jouent le rôle du principe formel. Cependant, cette théorie d'un flux astral générateur de forme a été largement remplacée, au XVI^e siècle, par une théorie de la génération spontanée vue comme sous-produit de la putréfaction. Dans la quatrième section, je discute de l'arrière-plan sceptique et libertin de certains aspects de la génération spontanée débattus au début de la période moderne. Dans la cinquième section, les choses se corsent, dans la mesure où nous étudierons les considérations par lesquelles une vision prépondérante de la génération sexuelle au début de la modernité, comme celle de Descartes, a réduit celle-ci à la génération spontanée par d'autres moyens, qui sont la création de nouveaux organismes par le biais de parties coagulées aussi bien que par l'imposition d'une forme par un agent. Dans cette optique, j'argumente que le statut ontologique des espèces est un problème particulier pour la théorie mécaniste générale — et particulièrement celle de Descartes — qui ne concerne pas seulement la génération spontanée, mais aussi la génération *simpliciter*. Le mécanisme omet entièrement toute explication permettant de comprendre comment et pourquoi le même engendre le même, bien que, contrairement à la biologie d'Aristote, tous les théoriciens du début de la modernité, mis à part les libertins, soient tous, au moins implicitement, des défenseurs du fixisme des espèces.

Cet article s'appuie dans une large mesure sur le travail de Scott Atran concernant les fondations cognitives de la taxonomie biologique occidentale, tout en critiquant aussi sa position sur certains points¹. Plus précisément, il

1. Scott Atran, *Cognitive Foundations of Natural History : Toward an Anthropology of Science* (Cambridge, Cambridge University Press, 1991).

sera démontré qu'au cours de la période moderne des changements significatifs dans le concept des espèces sont survenus pas seulement, ou premièrement, comme conséquence des nouvelles pratiques et exigences de classification, mais aussi comme résultat de courants philosophiques et théologiques qui n'étaient pas de prime abord préoccupés par les pratiques scientifiques naturelles.

1. La génération spontanée et l'ontologie des espèces chez Aristote

Sur la base d'une fausse compréhension de la conception aristotélicienne des espèces, David Hull identifie un pseudo problème maintenant bien connu concernant la place des créatures générées spontanément au sein de la métaphysique aristotélicienne de la biologie. Son inquiétude se rapporte au fait que la génération spontanée constituerait une rupture dans la théorie fixiste des espèces d'Aristote. Il pose ainsi le problème :

Aristote qualifie habituellement de spontané un événement ou une chose dans la mesure où il apparaît qu'il ou elle se produit toujours ou la plupart du temps, mais dans le cas cité (HA, 551a, 13) et dans de nombreux autres, les changements surviennent régulièrement même quand le matériau approprié est présent et que les conditions sont réunies [...]. Si des animaux de moindre statut peuvent participer de l'éternel et du divin sans l'aide de causes efficientes, formelles et finales coïncidentes, pourquoi cela ne serait-il pas possible pour toutes les espèces ? Et s'il n'y a aucun besoin pour les causes efficientes, formelles et finales de coïncider entre elles dans toutes les espèces, alors rien n'empêche d'accepter la théorie évolutionniste².

Même si cette inquiétude n'a pas sa place, la position de Hull doit néanmoins être appréciée en raison du fait que son pseudo problème a nourri grandement la réflexion concernant la question plus vaste de la relation historique entre les ontologies de type biologique, d'une part, et les théories de la génération spontanée, d'autre part.

Aristote a longtemps été considéré comme un partisan de ce qui a été appelé le « concept typologique des espèces », qui soutient (i) la fixité éternelle des espèces, et (ii) la doctrine que « tout individu devient nécessairement ce qu'il doit être en vertu de propriétés spécifiques à son espèce, permettant de saisir les propriétés qui définissent l'essence du genre d'être de l'individu et qui font aussi de celui-ci un être particulier³ ». Selon Ernst Mayr :

[le typologisme] cherche à assigner la variabilité de la nature à un nombre fixe de types de base considérés selon différents niveaux. Il s'agit de postuler que tous les membres d'un même taxon reflètent une même nature essentielle, ou, en

2. David Hull, « *The Conflict between Spontaneous Generation and Aristotle's Metaphysics* », *Proceedings of the Inter-American Congress of Philosophy VII* (1967, 8, 245-250, 247).

3. Atran, *Cognitive Foundations of Natural History*, 138 ; citing W. V. O. Quine, « Three Grades of Modal Involvement », in *The Ways of Paradox* (New York, Random House, 1966, 173-174).

d'autres mots, qu'ils se conforment au même type [...]. Conséquemment, la variation est considérée par le typologiste comme triviale et inappropriée⁴.

Mais bien que d'une certaine manière Aristote considère les espèces comme « fixées », il ne le fait pas en vertu d'une croyance en des universaux sous-jacents requérant une telle fixité, et dans la mesure où il croit que les individus deviennent ce qu'ils sont en raison de propriétés d'espèce spécifiques, il ne considère pas que cela survient par nécessité. Dans les dernières décennies, les études aristotéliennes ont largement cherché à surpasser les attributions précédentes et simplistes d'essentialisme chez Aristote, en tant que croyance en des universaux fixés et éternels. Par exemple, D. M. Balme avance que l'essentialisme d'Aristote,

bien qu'apparemment soutenu par plusieurs citations de la *Logique* et des *Métaphysiques* [...], est directement contredit par certaines de ses théories les plus matures et savamment argumentées en biologie [...]. Aristote soutient que l'animal se développe d'abord par la filiation parentale, incluant même des détails non essentiels, alors que la forme commune de l'espèce est uniquement une généralité qui « accompagne » cette filiation⁵.

Selon Balme, Aristote traite les espèces en tant qu'elles sont simplement obtenues par généralisation. Même s'il est vrai qu'un membre d'une espèce peut permettre d'expliquer les caractéristiques des individus, cela ne résulte pas du fait que les espèces sont une cause efficiente de la formation des individus, mais du fait que dans certaines circonstances les individus sont avantagés par ces caractéristiques (*ibid.*).

Quelle est la preuve textuelle sur laquelle Hull fonde son argument ? Aristote donne une formulation classique de son concept de genre [*kind*] dans le *De Anima*. Les créatures, dit-il, ou au moins celles dont « le mode de génération est spontanée »,

produisent d'autres de leur genre, les animaux produisant les animaux, les plantes produisant les plantes, afin qu'ils puissent être unis, autant que leur propre nature le leur permet, dans l'éternité et le divin. C'est là l'idéal que toute créature cherche à atteindre, et qui détermine le comportement, pour autant que celui-ci soit naturel... Mais puisque les choses mortelles ne peuvent perdurer continuellement dans l'éternité et le divin (car rien de ce qui périclisse ne peut préserver son identité ni demeurer numériquement un), elles participent à l'éternité et à la divinité de la façon qui leur est propre, et avec un succès inégal, atteignant l'immortalité non par elles-mêmes, mais indirectement par leurs progénitures, lesquelles, à travers les individus distincts, forment une unité spécifique (*De Anima* 415a, 27-b9).

4. Ernst Mayr, *Principles of Systematic Zoology* (New York, McGraw-Hill, 1969, 66).

5. D. M. Balme, « Aristotle's Biology was not Essentialist », in Gotthelf and Lennox (eds.), *Philosophical Issues in Aristotle's Biology* (Cambridge, Cambridge University Press, 1986 291-301, 291).

La phrase « avec un succès inégal » est mentionnée ici. À vrai dire, ce qui préoccupe Aristote, c'est l'analyse fonctionnelle et morphologique des genres dans leur environnement écologique. Bien sûr, il existe pour lui une théorie des natures sous-jacente, mais au bout du compte, la méthode d'Aristote est de déterminer l'appartenance à un genre par le biais de l'aspect morphologique. Comme le dit Balme, Aristote fait d'abord des systématiques en zoologie, et d'abord de la morphologie en systématique. Selon Aristote, les natures se développent sous des conditions naturelles spécifiques, bien que rien n'assure de manière certaine qu'un organisme individuel deviendra nécessairement ce qu'il doit être simplement en vertu de la force intrinsèque de sa nature d'espèce spécifique. Il est vrai que pour lui tous les organismes ont comme but de devenir des copies parfaites de leurs pères, qui, eux, contribuent au principe actif de la reproduction à travers le véhicule qu'est le sperme, mais ce but est toujours empêché par le fait que l'enfant requiert aussi la contribution passive du sang menstruel de la mère afin de se développer en une substance particulière et déterminée, et que son développement est aussi influencé par des facteurs environnementaux.

Donc, la reproduction parfaite n'arrive jamais. Les causes environnementales et matérielles jouent toujours un rôle dans l'embryogénèse. Elles peuvent être appelées « préventives » : elles produisent un certain effet en demeurant comme un obstacle à d'autres causes actives — ici, le sperme du mâle — parce que si elles étaient absentes les causes actives auraient un tout autre effet. Quand les causes préventives sont prédominantes, l'enfant est suffisamment « déformé » pour devenir une femelle (GA, 766a, 20-25). À un niveau plus élevé de difformité, « lorsque les mouvements du mâle se relâchent et que le matériel venant de la femelle n'est pas maîtrisé, ce qui reste, c'est ce qu'il y a de plus « général », c'est-à-dire « l'animal » tout simplement » (GA, 769b, 11-13). La ressemblance avec d'autres espèces animales est « seulement une ressemblance » et « en aucun cas elles ne sont ce qu'elles sont présumées être » (GA, 769b, 18-20). Pour Aristote, une « monstruosité est réellement une sorte de difformité (*kai gar to teras anaperia tis estin*) » (GA, 769b, 30-1), catégoriquement non différente d'un cas mineur de déviation face à l'idéal absolu de la reproduction du père. Un monstre survient quand la nature formelle dans la reproduction est incapable de contrôler la nature matérielle (GA, 770b, 17-18). Cela est un corollaire « non naturel » à la nature des animaux qui cherchent à se reproduire, mais il n'y a absolument rien de non naturel quant à l'apparition de monstres dans l'accouplement de deux animaux.

Pour Aristote, un monstre n'est pas une rupture abrupte dans l'ordre de la nature, il ne s'agit que d'un cas de *steresis* ou de déviation — et, en effet, tous les organismes dévient face à l'idéal de la reproduction de quelque manière, en tant qu'ils ne sont jamais la copie identique de leur père. C'est en raison de cet arrière-plan qu'Aristote affirme que les femmes sont déformées — elles dévient de l'idéal reproductif, en effet quelque peu plus que l'homme, mais cela n'a rien d'exceptionnel. Pour Aristote, la monstruosité est,

dans une large mesure, qualifiable de norme. Nous pouvons dire que pour lui la tétatologie est la façon de comprendre la diversité des individus dans une espèce : le fait que tous les individus ne soient jamais la reproduction exacte de leur père ; et que les pères accomplissent l'éternité « avec un succès inégal ». Une approximation très pauvre de l'éternité consisterait dans le cas où les influences environnementales ou matérielle subies par une créature ne l'amèneraient pas seulement à ne pas porter les traits particuliers de son père, mais aussi ceux qui sont typiques à l'espèce à laquelle les parents appartiennent. L'espèce des parents est, dans tous les cas, seulement « un universel représentant une... possibilité dans le cours futur du développement optimal de l'individu⁶ ». Cet universel n'a pas, par lui-même, d'existence réelle mais est plutôt la substance dérivée ou secondaire à travers laquelle les substances primaires particulières peuvent être connues.

Comme Scott Atran l'explique, les espèces surviennent naturellement comme une « nécessité empirique » — qui fait partie du mélange ontologique [*the ontological fold*] de la nature — et sont cependant conditionnelles à une constellation idéale de circonstances matérielles pouvant ne jamais être rencontrées⁷. Ainsi, il n'y a rien qui garantisse que le même engendrera constamment — encore moins éternellement ! — le même, bien que, pour Aristote, la reproduction sexuelle s'élève à une sorte d'approximation de l'éternité pour les substances corporelles qui sont mortelles et sublunaires. En effet, si nous considérons la position plus générale d'Aristote concernant la relation entre les substances particulières premières et les substances secondaires, il est remarquable qu'il puisse continuer à être considéré comme un partisan d'une théorie fixiste rigide des espèces. Pour lui, la substance secondaire telle que l'humanité ou la bovinité possède seulement une réalité dérivée de l'existence actuelle d'humain ou de vache. Que cette conception métaphysique constitue ou non un renversement de la théorie platonicienne des formes, il est clair que, pour Aristote, il n'y a pas d'éternel, de standard fixé avec lequel un individu peut être comparé afin de justifier sa qualification pour l'appartenance à une espèce inchangée. L'éternité selon la vision aristotélicienne de la reproduction sexuelle ne peut exister que dans les limites possibles d'une substance corporelle finie. C'est-à-dire, assez peu. La nature ne fait qu'approximer l'éternité, et les circonstances environnementales pourraient bien causer cela ; dans le cas de certaines espèces, elles le font de façon assez médiocre.

En quoi cette compréhension révisée des espèces chez Aristote concerne-t-il notre compréhension de sa vision de la génération spontanée ? Jim Lennox a fait vivement ressortir la connexion entre ces deux problèmes. Selon lui, la différence fondamentale entre les deux modes de génération se rapporte au fait qu'il n'y a pas de causalité téléologique en cause dans la variété spontanée. La génération sexuelle est une réplique formelle qui engendre des explications

6. Atran, *Cognitive Foundations of Natural History*, 193.

7. *Ibid.*, 138.

téléologiques, lesquelles sont des « explications de processus par la moyenne de leurs résultats⁸ ». Dans le cas des espèces générées sexuellement, leur « noblesse » est évaluée selon « la chaleur interne vitale caractéristique de l'espèce⁹ ». À l'opposé, dans la génération spontanée, la quantité et la force exacte du *pneuma*, et conséquemment de la chaleur vitale qu'une enceinte « capturera » sera fonction du lieu où il prendra forme. Comme Aristote l'explique dans *De la génération des animaux* :

Les animaux et les plantes sont issus de la terre et du liquide parce qu'il y a de l'eau dans la terre et de l'air dans l'eau, et dans tout air il y a de la chaleur vitale ; en un sens toute chose est ainsi pleine d'âme. De la sorte, les choses vivantes se forment rapidement quand cet air et cette chaleur vitale sont enfermés dans quelque chose. Quand ils sont très comprimés, le liquide corporel est chauffé, et il se produit comme une bulle mousseuse. La différence selon laquelle nous pourrions évaluer si le genre est plus ou moins honorable est déterminée par l'organisation du principe vital dans l'enceinte. Et tant les lieux de développement que le matériel enfermé sont des causes de cette organisation (GA, 762a, 18-27).

Cette même vision est répétée dans des détails correspondants à de nombreux endroits de l'*Historia Animalium*¹⁰. Le modèle de développement est fonction de la quantité de *pneuma* relative au matériau terreux et liquide du mélange enfermé. Ce mélange, qui peut changer selon les variables climatiques et géographiques, déterminera que l'enceinte devienne un oursin de mer ou une huître. La chaleur vitale qui est productrice des membres de l'espèce n'est donc pas dérivée d'une unité de forme avec ce membre de l'espèce. En d'autres mots, « il n'y a pas de description du déroulement d'un tel processus que l'on identifie comme production d'un certain type d'organisme¹¹ ». Lennox conclut que les processus spontanés ne devraient pas être mis en contraste avec ceux qui sont ordonnés, mais plutôt avec ceux qui impliquent une reproduction formelle. En bref, ce n'est pas qu'une bernacle n'a pas de forme. C'est simplement qu'elle n'a pas de père.

Dans la génération sexuelle, nous avons le phénomène unique de la reproduction formelle, dans lequel une idée universelle abstraite est approchée dans une génération du même par le même, mais ici encore la forme ne guide pas le processus ni ne garantit un résultat. Au contraire, la forme ou la substance secondaire peut être dite existante seulement dans le cas effectif où des individus semblables peuvent continuer de générer leurs semblables. Dans la génération spontanée, la génération régulière et ordonnée de ce qui apparaît comme membre d'un genre — le genre huître, le genre bernacle, etc. — est

8. James G. Lennox, « Teleology, Chance, and Aristotle's Theory of Spontaneous Generation », in James G. Lennox, *Aristotle's Philosophy of Biology: Studies in the Origins of Life Science* (Cambridge, Cambridge University Press, 2001, 229-249, p. 231).

9. *Ibid.*, 233, avec une référence à GA, II, 1, 732a, 24, 733b, 23.

10. E.g. HA, V, 539a, 18-26, 547b, 18-22.

11. Lennox, « Teleology, Chance, and Aristotle's Theory of Spontaneous Generation », 233.

possible par la constance relative des circonstances environnementales. Rien ne garantit que les circonstances environnementales vont demeurer les mêmes, et donc rien ne garantit qu'il y aura toujours des bernacles. Mais ce que Lennox sous-estime peut-être, en contraste avec Atran, c'est aussi que rien ne garantit que les circonstances locales permettront d'obtenir la certitude que les spatules (type d'oiseaux ayant un bec en forme de spatule) continueront de se reproduire, et donc, rien n'assure que la reproduction formelle continuera. Les spatules s'approchent, dans la mesure de leurs capacités, de l'éternité, et ce, à travers la reproduction formelle ; ainsi, leur habileté dépend autant des circonstances environnementales que dans le cas de la génération spontanée des bernacles.

Donc, la génération spontanée ne crée pas de problèmes réels pour la philosophie de la biologie d'Aristote, dans la mesure où il n'y a pas de raison de croire qu'il souscrivait à une sorte d'essentialisme typologique dans lequel la génération spontanée constituerait une menace. L'inquiétude de Hull ne met pas en lumière un point très intéressant de la connexion entre la génération spontanée et l'ontologie des espèces chez Aristote. En fait, loin de constituer une rupture dans sa biologie, le fait qu'il soit si convaincu de la réalité de la génération spontanée devrait nous porter à voir l'inadéquation fondamentale de notre considération précédente d'Aristote comme un fixiste rigide.

2. Les origines du fixisme

Selon Atran, la première étape vers une classification systématique globale des genres biologique est survenue au XVI^e siècle, quand des botanistes tels qu'Andrea Cesalpino ont cherché à fixer les espèces en tant qu'entité se perpétuant éternellement. Comme conséquence de cet effort pour développer un système compréhensif et universel de classification au début de la période moderne,

[il était, selon Atran,] nécessaire de fixer un critère pour déterminer les espèces dans l'optique de découvertes futures. Sans ce critère il ne pourrait y avoir de principe de justification pour unir sur une base commune des types provenant de différents climats à l'intérieur du même genre. Un tel critère devait donc établir que les caractères morphologiques perçus habituellement comme constants sont effectivement ceux qui devraient être constants en accord avec le plan éternel de Dieu¹².

En 1571, le botaniste italien écrit que :

l'éternité peut seulement venir de l'éternel : depuis que le travail propre à l'âme végétative est d'engendrer son semblable, lequel est fait pour l'éternité des espèces, il est nécessaire que sa substance ne soit pas corruptible. La raison de l'éternel ne réside pas dans les existences corruptibles prises individuellement ou dans leur totalité¹³.

12. Atran, *Cognitive Foundations of Natural History*, 142.

13. Andrea Cesalpino, *Quaestionum peripateticorum libri quinque* II, vii (Venice, Juntas, 1571) ; cité dans Atran 1991, 139.

En d'autres mots, les espèces doivent être comprises indépendamment du contexte écologique et non sujettes à une transmutation résultant de quelque changement écologique ; l'essence des espèces doit être conçue comme fidèlement transmissible d'une génération à la suivante par le biais de la reproduction sexuelle. L'essence d'une créature de l'espèce est déterminée exhaustivement par les parents. « [P]our que le même engendre partout le même », comme le dit Cesalpino, « en accord avec la nature¹⁴ ».

De manière intéressante, Cesalpino utilise Aristote pour appuyer son idée selon laquelle « le soleil et l'homme génèrent l'homme¹⁵ ». Cependant, le botaniste italien conçoit son désaccord avec le philosophe grec comme un amendement plutôt qu'un rejet. Pour Cesalpino, c'est l'homme qui se génère, et il est donc superflu de mentionner le soleil dans cette connection. Si le soleil joue un rôle dans la génération, c'est comme auxiliaire dans l'acte de reproduction, et non pas comme parent. Les éléments essentiels pour la continuité de la reproduction de l'espèce ont été placés dans les premiers parents de l'espèce à la Création, et même si le soleil est requis pour la *successio* post-Création des générations, cela ne diminue pas le fait que la continuité de l'existence de l'espèce est garantie par la *prima creatio*¹⁶. Toutefois, Cesalpino ne peut pas confronter le point de divergence fondamentale avec Aristote, à savoir, que celui-ci ne croit pas en une création, tandis que lui, en tant que chrétien, doit absolument la reconnaître. La conséquence de la divergence de Cesalpino avec Aristote serait une conception des espèces selon laquelle les essences des espèces sont fixées à la Création et où les générations subséquentes se développent par nécessité à partir de leurs premiers parents, sans égard aux altérations des circonstances environnementales. Près de deux siècles plus tard, Linné le formulera de manière encore plus ambiguë : « *Constans, immo constantissima, naturae lex est, quod similes procreentur a similibus, nec imbellem feroces progenerent aquilae columbam*¹⁷. »

Atran avance que cette nouvelle conception des essences prenant racine dans la reproduction sexuelle a été développée en conjonction avec les travaux des botanistes du début de la modernité. Alors que pour la biologie traditionnelle, y compris celle d'Aristote, l'intérêt premier avait toujours été de déterminer et de regrouper les espèces par rapport à l'aspect morphologique et à la tendance écologique, l'histoire naturelle après Cesalpino, et plus clairement encore après Linné, en est venue graduellement à se concentrer sur la détermination des affinités entre espèces génétiquement liées. Pour Cesalpino, l'essence est transmise à travers une reproduction de l'engendrement

14. Andrea Cesalpino, *De Plantis libri XVI* (Florence, Mariscot, 26), cité dans Atran, 1991, 142.

15. Pour un traitement intéressant de cette doctrine aristotélicienne, voir Andrea Falcon, *Aristotle and the Science of Nature : Unity without Uniformity* (Cambridge, 2005).

16. *Ibid.*, 92f. « [F]rustra apponitur Sol : si hominis calor includit Solis calorem. Separatim igitur haec intelligenda sunt, ut hominem homo gignat, & Sol sini hominis auxilio. »

17. Linnaeus, *Systema Naturae*, 1757, 26.

du même par le même, et c'est ce processus qui assure l'éternité des espèces plutôt qu'une approximation imparfaite de l'éternité, comme c'était le cas pour Aristote.

Selon Atran, la métaphysique des espèces des plantes de Cesalpino est une conséquence de certaines exigences du projet de classification systématique. Les espèces de plantes en sont venues à paraître fixées, et leurs essences à être aussi certaines et inchangeables dans la transmission d'une génération à la suivante, et ce, dans une visée qui n'avait plus rien à voir avec le contexte écologique de l'espèce.

En résumé, nous pouvons dire que la position décrite ci-dessus, qui a émergé au début de la période moderne comme résultat des nouvelles exigences de la classification scientifique, a été, par la suite, projetée injustement jusqu'à Aristote lui-même, qui était, de l'aveu général, l'aïeul reconnu de Cesalpino et autres, mais qui n'était pas, malgré tout cela, en accord sur tous les points avec les naturalistes du début de la modernité, y compris la question de l'ontologie des espèces.

3. Le flux astral comme agent formateur et le caractère non spontané de la génération non sexuelle

Il peut être valable de réviser la thèse d'Atran quant à l'importance accordée aux botanistes du XVI^e siècle dans l'émergence du concept moderne d'espèce, afin de voir que ce ne sont pas seulement de nouvelles méthodes de classification scientifique qui ont mené à une nouvelle ontologie des espèces, mais aussi l'interprétation de l'Écriture. Pour ne citer qu'un exemple, dans son commentaire sur la Genèse, Martin Luther écrit que, en formant Ève, Dieu a créé pour Adam un « associé pour la génération et la conservation de l'espèce » (*socia generationis et conservationis speciei*). De cette manière, Dieu s'est assuré que les futurs humains seront générés à la manière des bêtes, qui est opposée à celle des êtres forgés à partir de la boue : « Dieu n'a pas voulu pour les descendants d'Adam qu'ils naissent de la même manière qu'Adam, c'est-à-dire à partir de la terre. Il voulait plutôt que l'homme possède une génération semblable à celle des autres bêtes (*voluit, ut haberet generationem, qualem aliae bestiae habent*). » À partir de ce moment, le critère de l'appartenance à une espèce, pour les hommes et les bêtes, est clair : « Peu importe la chose vivante générée, elle doit l'être du mâle et de la femelle, de telle façon qu'elle vienne au monde par la femelle (*sic ut per foeminam edatur in lucem*)¹⁸. » Dans cette vision, le premier homme est généré à partir de la terre, et, les termes d'Aristote sont littéralement repris, avec Dieu qui joue le rôle du père. Dans une certaine mesure, Dieu communique une certaine forme à un amas de matière réceptive et appropriée dont Il possède

18. Martin Luther, *Vorlesungen über 1. Mose*, in *D. Martin Luthers Werke*. Kritische Gesamtausgabe, vol. 42. (Weimar, Hermann Böhlaus, 1911).

une idée intellectuelle¹⁹. Par la suite, le mode de génération change puisque Dieu n'a plus besoin de se voir attribuer chaque acte de génération ; il peut confier le rôle de géniteur aux descendants mâles d'Adam par un travail conjoint avec les femelles humaines.

Le fait que dans la tradition chrétienne le premier homme est né de la terre est peut-être très significatif. À l'évidence, cela ne devrait pas compter comme une instanciation de la génération *spontanée*, car, après tout, Adam est créé de manière supranaturelle par un Dieu possédant une idée précise de ce qu'il veut produire. En d'autres mots, la création d'Adam admet une explication téléologique. Si ce n'est pas en rapport avec Aristote, mais plutôt en rapport avec les médiévaux tardifs, la Renaissance et les idées du début de la modernité concernant les créatures nées de la terre établissent une distinction imprécise entre le mode de génération de ces créatures et celle d'Adam, car les deux sont vus comme générés de conséquences célestes ou d'un flux astral agissant sur la matière terrestre. Bien sûr, le Dieu chrétien est dit « céleste » uniquement dans un sens métaphorique, bien que le rôle de Dieu dans la première génération de l'humain et des espèces animales porte une ressemblance remarquable avec les idées médiévales et du début de la modernité concernant la « fertilisation » céleste de la matière terrestre. Les histoires de la génération spontanée des XIX^e et XX^e siècles laissent généralement de côté le rôle du soleil et des autres corps célestes, bien que ceux-ci soient presque toujours invoqués par ceux qui cherchent à définir ce mode de génération, même de manière critique, à l'époque où cela était encore une option. Par exemple, une note plutôt tardive dans un travail de référence de 1728 nous indique que :

La génération ÉQUIVOQUE est une méthode de production des animaux et des plantes, différente de la voie habituelle du coït entre mâle et femelle, mais je ne sais pas quel est ce pouvoir plastique, ou cette vertu dans le soleil [...]. Ainsi, les insectes, les asticots, les mouches, les araignées, les grenouilles, etc., ont habituellement été considérés comme produits par la génération équivoque, c.-à-d. par la chaleur du soleil réchauffant, agitant et imprégnant la poussière, la terre, la boue et les parties putréfiées des animaux. Cette méthode de génération, que nous appelons aussi spontanée, était communément revendiquée et supposée par les philosophes anciens : mais les Modernes, à partir de plus nombreuses et de meilleures observations, l'ont rejetée de façon unanime et maintiennent que tous

19. L'idée que la race originelle des hommes nés de la terre précède celle des hommes nés de la reproduction sexuelle apparaît déjà dans des sources grecques, y compris *L'homme d'État* de Platon. Celui-ci semble y suggérer, suivant les lignes que nous proposons, que pour être né de la terre il ne s'agit pas seulement d'émerger par chance du réarrangement de la matière, mais plutôt d'être créé et développé directement par Dieu, sans l'intermédiaire de la reproduction humaine : « Dans ce cycle d'existence il n'y a pas une telle chose que la procréation d'animaux par d'autres, mais ils naissent de la terre, et nos ancêtres, qui sont venus à l'existence immédiatement après la fin du dernier cycle et au début de celui-ci, ont préservé la recollection [...]. Cette vie bénie et spontanée n'appartient pas à cela, mais à l'état précédent, dans lequel Dieu était le gouverneur de tout le monde » (*L'homme d'État*, 271a-b). Il existe aussi des références au *gegeneis* dans la *République* (415d), et dans le *Sophiste* (248c).

les animaux, et les végétaux aussi, sont produits de façon univoque, par les parents chez les animaux et par les membres d'une même espèce ou dénomination chez les végétaux²⁰.

Dans *De la génération des animaux*, Aristote lui-même mentionne, sans plus d'explications, que le *pneuma* dans la semence animale est « analogue » au matériel des étoiles (GA, II, 3, 736b, 29). Comme un commentateur le remarque, bien qu'Aristote ne veuille pas dire que le *pneuma* provient de l'éther, nous retrouvons néanmoins cette idée dispersée dans la tradition aristotélicienne²¹. Avicenne décrit le *pneuma* comme une « *virtus informativa* » et pas seulement comme analogue à la vertu des corps célestes. À l'évidence, ces deux vertus seraient du même genre²². Tout au long de la philosophie scolastique, il existe une présomption commune selon laquelle le pouvoir formateur à l'œuvre dans la nature en général est une manifestation de la faculté formative traditionnellement conçue pour gouverner la croissance et le développement biologique. Ce pouvoir formateur est filtré de la sphère céleste vers la sphère terrestre, et ainsi nous pouvons dire que dans tous les cas de croissance et de développement naturel, y compris l'émergence des fossiles et des cristaux aussi bien que les organismes générés « spontanément », il existe un agent dont le rôle est analogue à celui du père dans la génération sexuelle : les cieux.

Nous observons une telle conception de l'émergence des formes dans la matière terrestre au sein de la vision scolastique traditionnelle de l'origine des fossiles. Le dominicain Antoine Goudin argumente qu'il existe des causes tant efficientes que finales travaillant dans la production terrestre des roches ressemblant à des animaux ou à des parties d'animaux. Leur cause efficiente est une sorte de cuisson amenée par les exhalations venant des profondeurs de la terre qui forment la strate où les fossiles se retrouvent dans une espèce de fournaise. Leur cause finale, quant à elle, est :

une certaine force que la terre possède elle-même de manière variable, suivant les différents endroits où les corps mélangés sont formés. Cette force, similaire au sein maternel duquel les animaux sortent, joue assurément un grand rôle dans la formation de ces corps : c'est pourquoi, selon Aristote et saint Thomas, la terre

20. Ephraim Chambers, *Cyclopaedia, or, An universal dictionary of arts and sciences*, contient les définitions des termes et les visions des choses signifiées ainsi dans plusieurs arts, tant libéral que mécanique, et dans plusieurs sciences, humaine et divine : les figures, les genres, les propriétés, les productions, les préparations, et les usages ; des choses naturelles et artificielles : la montée, le progrès et l'état des choses ecclésiastiques, civiles, militaires et commerciales ; avec plusieurs systèmes, sectes, opinions, etc. ; parmi les philosophes, les divins, les mathématiciens, les physiciens, les antiquaires, les critiques, etc. ; le tout vu comme un cours sur les savoirs anciens et modernes (London, 1728).

21. Voir C. S. F. Burnett, « The Planets and the Development of the Embryo », in *The Human Embryo: Aristotle and the Arabic and European Traditions*. G. R. Dunstan (dir.), Exeter, University of Exeter Press, 1990, 113-122.

22. Voir B. Nardi, « La teoria dell'anima e la generazione delle forme secondo Pietro d'Abano », in *Saggi sull'Aristotelismo Padovano dal secolo xiv al xvi*, Florence, 1958, 1-17.

et l'eau fournissent la matière à tout ce qui provient des entrailles de la terre, tout comme le ferait une mère, alors que les cieux et les étoiles remplissent le rôle du père, qui donne la forme²³.

Un principe formateur « mâle » exerce son influence sur la matière « maternelle » de la terre et, de la sorte, donne naissance à des formes dans la terre qui ressemblent à des êtres vivants. Un fossile, pour Goudin, est simplement une créature générée « spontanément », dont la forme est imposée dans le mauvais matériau — la pierre, plutôt que la boue souple — et donc qui est incapable de vivre et de se mouvoir comme le ferait un animal, et ce, même s'il partage la forme d'un animal.

Cette force cosmique est aussi un lieu commun dans la tradition platonicienne, s'étendant de Marsilio Ficino au XV^e siècle jusqu'aux figures importantes parmi les platoniciens de Cambridge tel qu'Henry More au XVII^e siècle. Ficino se demande rhétoriquement : « Depuis le commencement de toutes les choses générées, les influences célestes n'ont-elles pas accordé de merveilleux cadeaux dans la concoction de la matière et de sa réunion finale ? » Pour lui, la génération spontanée n'est rien de plus que ce qui survient quand les rayons des cieux concoctent à partir de la matière appropriée des organismes complexes. Puis, Ficino continue en demandant : « N'y a-t-il pas d'innombrables grenouilles et animaux similaires, quand les cieux le favorise, qui bondissent hors du sable en un moment ? Tel est le pouvoir des cieux par rapport au matériau approprié. » Il continue en ajoutant de nombreux autres phénomènes impliquant l'influence du pouvoir de la vision et de l'imagination afin de rendre compte du fait, *a fortiori*, que les rayons célestes ont le pouvoir d'influencer la forme terrestre des choses vivantes.

Je mets de côté les fantaisies accomplies par un coup d'oeil soudain et un amour très passionné instantanément enflammé par des rayons émanant des yeux... Je ne mentionnerai pas la rapidité avec laquelle un œil enflammé afflige quiconque le regarde, ni comment une femme ayant ses menstruations affecte un miroir en se regardant dedans... Qu'est-ce qu'un chien enragé peut accomplir sans une morsure apparente ? [...] À la lumière de tout cela, allez-vous dénier que les rayons des yeux des cieux qui nous regardent et nous touchent réalisent des merveilles en un instant ? Si une femme enceinte peut marquer instantanément par le toucher une partie du corps de l'enfant à naître avec l'empreinte de quelque chose qu'elle désire, n'allez-vous alors jamais douter que les rayons nous touchent d'une telle manière ou qu'ils peuvent accomplir diverses choses²⁴ ?

Ficino maintient que « la taille immense, le pouvoir et le mouvement des choses célestes supposent que tous les rayons de toutes les étoiles pénètrent

23. Antoine Goudin, *Philosophie suivant les principes de saint Thomas*, trad. T. Bourard, Paris, 1864 (éd. originale Paris, 1668), 301 : *Des corps mixtes inanimés, dit fossiles*. Je remercie Roger Ariew d'avoir porté ce travail à mon attention.

24. Marsilio Ficino, *Three books on life*. Éd. et trad. avec introduction et notes, de Carol V. Kaske and John R. Clark (Binghamton, Medieval and Renaissance Texts and Studies, 1989), 323-325.

à un moment dans la masse de la terre ». Ces rayons pénètrent le centre de la terre. Par l'intensité « des rayons, le matériau de la terre — sec et sans humidité — s'éveille immédiatement, et une fois éveillé est vaporisé et dispersé à travers des canaux dans toutes les directions et soufflé à l'extérieur par des flammes et du souffre ». Il décrit ce feu comme étant « très foncé » et « ressemblant à une flamme sans lumière », utilisant ainsi la même phrase dont Descartes donnera l'écho en décrivant le « feu » qui brûle dans le cœur au moment où le fœtus s'éveille.

Cette vision de l'influence céleste trouve un écho au XVII^e siècle à travers le platonicien de Cambridge, Henry More. Dans *The Immortality of the Soul*, More maintient que le soleil et les étoiles sont les « Êtres les plus intellectuels du monde », et qu'ils ont

rempli toute la Terre avec un mouvement vital, donnant naissance à d'innombrables sortes de fleurs, d'herbes et d'arbres hors du sol. Ils ont aussi engendré la plupart des genres de créatures vivantes. Ils ont empli les mers de poissons, les sols de bêtes et les airs de volatiles ; la matière terrestre a été formée aussi facilement dans les formes vivantes de ces animaux par la puissante marque de l'imagination du soleil et des étoiles, que l'embryon du vers est marqué par l'imagination de la mère qui le porte²⁵.

Dans le même travail, More défend une sympathie entre les corps terrestres et astraux comme une conséquence de leur sujétion mutuelle à « l'Esprit de la nature ». More maintient que l'esprit universel « est prêt à changer sa propre activité et sa production, indifféremment du monde ou de la forme quand l'occasion l'y engage, et donc de préparer un édifice, ou au moins les plus grands traits et délimitations conséquentes, peu importe l'âme spécifique, et ce, partout où la matière permettra ses opérations²⁶ ».

Ce mode de génération n'est pas, à proprement parler, « spontané », au point que les cieux fonctionnent littéralement comme un père, en imposant une forme à une matière réceptive. La théorie de la génération spontanée en tant qu'imposition de forme semble disparaître précisément au moment même où la théorie de la génération sexuelle se transforme, d'une imposition de forme par un agent paternel, en un processus purement thermomécanique (Descartes) ou en un déclenchement de quelque primordium (Harvey, et plus tard les préformationnistes, spermistes, ou ovistes). En d'autres mots, la génération des grenouilles à partir des rayons cosmiques a perdu la faveur, non pas parce que cela est ridicule, mais parce que les rayons ont été pensés comme jouant un rôle trop analogue au principe formel et actif auparavant attribué à la semence. La génération non sexuelle était antérieure au mécanisme et jamais conçue de manière strictement spontanée ; et, inversement, la génération sexuelle mécaniste était conçue comme spontanée au moins dans une certaine mesure. Avec la montée du mécanisme, la génération spontanée

25. Henry More, *The Immortality of the Soul*. A. Jacob (ed.) Dordrecht : Nijhoff, 1987, 70.

26. *Ibid.*, 257.

en est venue à être vue principalement comme une hétérogénèse, une pure conséquence de la putréfaction : ici, les nouveaux organismes ne sont pas développés à partir d'une matière prévue mais sans forme, en fait ils sont plutôt le sous-produit de la pourriture d'organismes morts.

4. Hétérogénèse et la descendance des espèces

Athanasius Kircher écrit vivement dans son *Scrutinium pestis* de 1658 :

Toute chose vivante produit de sa propre pourriture quelques animaux congruents et différents des autres. Cela nous est prouvé par l'expérimentation actuelle de différentes herbes, et cela est vrai pour les grains qui sont transformés en vers volants. Cela est aussi juste pour certains animaux plus ou moins organisés. Un bœuf mort et pourri est transformé en abeilles [...]. Les chevaux vivants et morts produisent des guêpes et des scarabées qui prennent comme nourriture le sang des animaux qui leur ont donné la vie, à leur plus grand agacement. Les êtres humains (aussi bien que quelques brutes) génèrent des bêtes de lit, des poux et des puces, lesquels sont des compagnons intimes produits par la nature pour retirer le sang corrompu. Un corps mort, empli de pourriture, devient une pouponnière pour les vers. Les restes d'insectes, quand ils sont pourris, produisent des animaux de même nature²⁷.

Nous pourrions suggérer que ce nouvel accent sur l'hétérogénèse, par contraste avec la transmission de formes par les corps célestes, était largement le résultat de deux développements distincts du début de la période moderne. Le premier a été l'invention du microscope et l'éveil résultant de l'ubiquité de la vie organique. Cette découverte a eu une influence importante sur la façon dont les gens concevaient la question du point de départ ultime de la génération et semble aussi avoir motivé certaines personnes à repenser la possibilité du prédicat « vivant » dans la nature : après la découverte des micro-organismes dans la boue, dans les étangs, il est apparu de plus en plus que toute nouvelle créature générée spontanément ne pouvait être générée que par une autre créature ou à partir de ses restes. Le deuxième facteur, mentionné trop brièvement ci-dessus, était que l'accent sur la génération comme imposition de forme — issue indifféremment du père ou des cieux — à partir d'une matière sans forme, en est venu à être minimisé, puisque l'observation microscopique révélait que ce qui paraissait sans forme pour un simple œil humain (sans l'aide d'un microscope) possédait souvent une structure compliquée où, plus tôt, on ne percevait qu'une homogénéité.

Un facteur important dans la nouvelle conception de la génération spontanée que nous avons décrite a été la montée de la technologie microscopique. Un autre facteur a été le libertinisme sceptique qui travaillait précisément avec la possibilité que les types biologiques aient des origines terrestres plutôt que divines et que la multiplicité des formes de vie que nous observons

27. Athanasius Kircher, *Scrutinium pestis* ; cité dans Harry Beal Torrey, « Athanasius Kircher and the Progress of Medicine » *Osiris*, V, 1938, 246-275, 258.

autour de nous soient le résultat de la chance. Au XVI^e siècle, Girolamo Cardano²⁸ et Julius Scaliger²⁹ ont tous les deux entretenu la possibilité que les cadavres des gros animaux puissent produire des animaux nouveaux et parfaits, et peut-être même des êtres humains. Dans le siècle suivant, plusieurs seront inspirés par la découverte des micro-organismes pour identifier une large variété de changements microscopiques comme des cas d'hétérogénèse. Ainsi, Leibniz écrit en 1669 à son mentor Jakob Thomasius : « Quand le fer rouille, il surgit un genre de forêt en une minute ; rouiller est donc une altération du fer, mais aussi une génération de petits buissons (touffes de matière)³⁰. » Peu importe que le changement naturel soit une altération ou une génération, en d'autres mots, il s'agit d'une question de perspective. Quelques décennies plus tard, Margaret Cavendish explique que la « génération » de certains organismes peut être considérée entièrement comme une altération de leur source matérielle :

J'ai mentionné dans mes *Philosophical Letters* qu'aucune créature animale ne peut être produite par métamorphose, qui est un changement dans les mouvements de certaines parties de la matière ; mais (j'exprime aussi au même endroit) que de tels animaux se produisent l'un par l'autre, et la production de l'un n'est pas causée par la destruction de l'autre ; de telles créatures ne peuvent être produites par une stricte métamorphose, sans une transmigration ou une réplication des parties du géniteur ; mais les insectes comme les asticots, et plusieurs autres sortes de vers et de mouches, et tous ceux qui n'ont pas de géniteur de leur espèce mais qui sont élevés à partir du fromage, de la terre et de la boue, etc., ne peuvent être produits que par métamorphose et non par réplication des parties³¹.

Au moment où les théories de la génération spontanée passaient de la vision de l'imposition de la forme sur la matière terrestre par des forces cosmiques, à des théories de l'hétérogénèse comme sous-produit de la pourriture, elles se sont trouvées éliminées dans certains cercles libertins par les théories du naturel, comme opposées au supranaturel, et les théories des origines des espèces, y compris humaines. Selon l'historien Giuliano Gliozzi, dans son étude monumentale des théories de la différence raciale du début, la majorité des défenseurs de la génération spontanée des humains au XVI^e siècle étaient des padouans lourdement immergés dans la biologie aristotélicienne. Parmi les plus importants et controversés de ceux-là, mentionnons Lucilio Vanini, que Gliozzi voit comme « liant le polygénétisme exprimé dans le naturalisme du XVI^e siècle avec la version qui prendra une nouvelle forme dans les milieux

28. Girolamo Cardano, *De propria vita* (Amsterdam, 1654).

29. Julius Scaliger, *Exotericarum exercitationum liber quintus decimus, de Subtilitate ad Hieronymum Cardanum* (Paris, 1557).

30. G. W. Leibniz, *Die philosophischen Schriften von G. W. Leibniz*, vol. I. C. I. Gerhardt (dir.), Berlin, 1875-90, 19.

31. Margaret Cavendish, *Observations upon Experimental Philosophy*. Eileen O'Neill (dir.), New York, Cambridge University Press, 66f.

du libertinisme français dans la première moitié du XVII^e siècle³² ». Le polygénétisme, la vision selon laquelle les différents groupes raciaux humains possèdent des origines distinctes, implique dans la plupart de ses variations une relation aux origines naturelles d'au moins certains groupes d'humains, puisque la seule création divine est celle décrite dans la Genèse. En d'autres mots, les juifs et les chrétiens ont été créés, mais les païens ont été formés à partir de la boue ou sont vus comme un sous-produit de la putréfaction. Il s'agit d'une vision modérée au moins dans une certaine mesure : à proprement parler, cela sauve les êtres humains d'une origine terrestre opposée à une origine surnaturelle mais restreint la classe des êtres humains à un bien plus petit groupe que les missionnaires et leurs supporters colonialistes le désiraient.

Dans un dialogue, Vanini, dans le personnage de Jules César, rapporte l'opinion de Diodorus de Sicile, selon qui « le premier homme est né par chance de la boue de la terre³³ ». Vanini y utilise des techniques narratives de distanciation permettant de ne pas identifier ce qu'il affirme comme sa propre vision, car ce n'est pas une chose qu'il a l'intention de faire entrer dans le débat légitime. Jules César répond à l'objection que nous avons toujours vu les petits animaux comme les grenouilles et les souris apparaître spontanément, et qu'il s'agit davantage d'une question de quantité de matière putride disponible pour la génération que d'une limitation inhérente à ce qui peut être produit de cette manière. Pour corroborer cela, il cite une autre variation de l'adage, *Ex Africa semper aliquid novi*, suggérant que dans le Nil il existe des bancs de vase suffisamment larges pour produire des « animaux énormes » s'ils sont chauffés par le soleil. Vanini propose, encore à travers Jules César, que « le premier homme a été généré par la corruption des primates, des cochons et des grenouilles », « puisqu'il existe une grande ressemblance entre ces espèces en ce qui concerne leur peau et leurs habitudes³⁴ ». Par la suite, il modifie cette vue pour suggérer que peut-être seuls les Éthiopiens sont des descendants des primates, puisqu'ils ont, selon lui, une pigmentation similaire. Le racisme manifeste de cette suggestion ne doit pas nous empêcher de discerner sa motivation réelle, c'est-à-dire adopter une position de compromis entre la véritable vision blasphématoire que tous les êtres humains, y compris les chrétiens européens, sont issus de la chance, et la vision que tous les humains, peu importe la diversité raciale, sont les descendants d'Adam.

La vision mystérieuse de Vanini semble plutôt être que les animaux sont générés à partir de la matière pourrie d'autres animaux qui sont, d'une certaine manière, comme leurs « progénitures », mais pas membres de la même

32. Giuliano Gliozzi, *Adamo e il Nuovo Mondo. La nascita dell'antropologia come ideologia coloniale: dalle genealogie bibliche alle teorie razziale (1500-1700)* (Florence, 1977); trad. : *Adam et le nouveau monde. La naissance de l'anthropologie comme idéologie coloniale: des généalogies bibliques aux théories raciales* (Paris, Théétète, 2000, 427).

33. Lucilio Vanini, *De Admirandis Naturae Reginae Deaque Mortalium Arcanis libri quatuor* (Paris, 1616) ; cité dans Gliozzi, 2000, 427.

34. Gliozzi, 2000, 429.

espèce. Ainsi, son hétérogénèse est en même temps une explication de la descendance des formes de vie. La matière organique pourrie qui était avant un singe peut donner naissance à un homme en vertu des propriétés partagées de ces deux espèces. Vanini n'a jamais expliqué clairement ce qu'il considérait comme des lignées différentes de descendance, mais il révèle au moins l'interconnexion de la possibilité que les créatures émergent de la terre et de la boue, d'un côté, et qu'elles partagent une lignée de descendance commune, de l'autre : si une créature peut émerger *de novo* à partir de la pourriture sans être le résultat de quelque « conception » plutôt céleste ou séminale, alors la création supranaturelle est réfutée répétitivement partout dans le monde à chaque moment.

Vanini n'est pas le seul libertin à discerner la connexion entre la génération spontanée, d'un côté, et les origines naturelles de l'humanité, de l'autre. Quelques-uns, comme François de la Mothe le Vayer, croient que c'est seulement en faisant appel à la génération spontanée que nous pouvons expliquer la variété du genre humain, des créatures, évidemment sans âme, qui habitent aux limites du monde connu³⁵. Comme il l'écrit : « La nature est capable de produire d'elle-même — sans que les hommes tombent dans une bestialité exécrable — des animaux qui nous ressemblent de telle manière qu'ils nous obligent à dire qu'il existe parfois une plus grande différence entre un homme et un autre qu'entre nous et eux. » La Mothe le Vayer affirme que cela n'est pas difficile à concevoir si nous décidons de nous appuyer sur la valeur de plusieurs autorités de l'Antiquité et du Moyen Âge, telles que Platon, Aristote, Épicure, Lucrèce, et Avicenne, qui admettent tous que « [bien avant], la Terre nous avait produit à partir d'elle comme [elle avait fait] les animaux ». La Mothe le Vayer n'a pas simplement répété la théorie hétérogénétique des naturalistes du XVI^e siècle, mais a aussi offert une théorie développementale selon laquelle « les hommes ne viennent pas originellement à l'existence dans le stade parfait auquel nous les voyons³⁶ ».

5. Embryogénèse mécanique vue comme spontanée

Nous avons vu qu'au début de la période moderne la génération spontanée a cessé d'être conçue comme une imposition de forme par les cieux dans une matière terrestre réceptive et appropriée, et, à la place, en est venue à être pensée comme la conséquence d'une chance, ou ce que nous pourrions appeler un sous-produit de la putréfaction, et cela a été mis en lumière par ceux qui souhaitaient promouvoir la possibilité d'origines naturelles pour chaque genre biologique, en opposition avec la création divine *ex nihilo*. De manière intéressante, une reconsidération très similaire de la génération *sexuelle* s'est imposée simultanément : qu'est-ce que la vision de l'embryogénèse de Descartes,

35. François de la Mothe le Vayer, *Cinq dialogues faits à l'imitation des Anciens* (Francfort, 1716).

36. Gliozzi, 2000, 432.

si ce n'est un effort pour expliquer comment, à travers une série de causes mécaniques, des fluides homogènes peuvent se coaguler et grandir, en se séparant en organes spécialisés, et à la fin devenir un ? En effet, Descartes pense qu'il s'agit de la façon la plus aisée de justifier la responsabilité de Dieu concernant les naissances anormales, tout en l'exaltant proprement en lui attribuant la sagesse de régler tous les phénomènes seulement par un nombre restreint de lois éternelles de la nature. Il résume son approche de l'embryologie dans les *Primae cogitationes* comme suit :

J'espère que certains diront dédaigneusement qu'il est ridicule d'attribuer un phénomène aussi important que la procréation humaine à de telles causes mineures. Mais quelles causes plus grandes que les lois éternelles de la nature peuvent être requises ? Avons-nous besoin de l'intervention directe de l'esprit ? Quel esprit ? Dieu lui-même ? Pourquoi alors des monstres naissent-ils³⁷ ?

Ce que Descartes souhaitait réaliser est bien connu. Mais ce qui a été beaucoup moins remarqué dans les études récentes, c'est que cet effort a été perçu par plusieurs comme réduisant les espèces sexuellement générées à un niveau ontologique vague où plusieurs étaient contents de ranger les bernacles et les souris, ainsi l'embryogénèse cartésienne fait effectivement de chaque genre biologique le produit de causes mineures et met ainsi les espèces dans une certaine crise. Cela est reconnu rapidement par Malebranche, qui note que même si Descartes a pu rendre compte de la formation des organes en général,

jamais, il me semble [...] les femmes et les animaux ne peuvent facilement produire des petits de la même espèce. Bien que l'on puisse donner quelques explications sur la formations du fœtus en général, comme Descartes a suffisamment tenté de le faire, il demeure très difficile [...] d'expliquer pourquoi une jument ne donne pas naissance à un veau, ou pourquoi un œuf de poule ne contient pas une perdrix ou un oiseau d'une nouvelle espèce³⁸.

Peu après, Leibniz fait l'éloge de la découverte du spermatozoa par Leeuwenhoek, et de sa théorie selon laquelle la créature préexistante est reçue par ce petit « vers », sur ce sol que les microscopistes hollandais ont restauré pour la dignité des hommes, contre les premiers penseurs — c.-à-d. les embryologistes mécanistes — qui avaient dit que le père « ne fait rien de plus que la pluie ». Pour Leibniz, accepter l'épigénèse mécanique comme opposée à son propre préformatisme spermiste n'est vraiment pas différent d'accepter la vision de Vanini selon laquelle les animaux peuvent émerger de la matière appropriée dans les bonnes circonstances environnementales et

37. Descartes, *Primae cogitationes circa generationem animalium* (Amsterdam, 1701, 15). « *Exspecto cur aliquis caperata fronte dicat esse ridiculum, rem tanti momenti, quanta est hominis procreatio, fieri ex tam levibus causis ; sed vero quas velint graviore quam Naturae leges aeternas ; forte ut ab aliqua mente fiant, a qua autem, an immediate a Deo, cur ergo aliquando fiunt monstra ?* »

38. Nicolas Malebranche, *De la recherche de la vérité*, in *Œuvres complètes de Malebranche*, vol. I, Geneviève Rodis-Lewis (dir.), Paris, Vrin, 1962, 243.

météorologiques. Même si cela peut survenir, il semblerait que ce ne soit pas une raison satisfaisante de justifier que le même engendre régulièrement le même : où il existe une possibilité que des organismes émergent de matière non vivante comme le résultat d'un processus thermomécanique, puisque les *formes* de la vie organique ne peuvent plus être vues comme les produits de la transmission d'un genre préservé, dans une essence immatérielle d'espèce, d'une génération à une autre.

Pour compliquer l'affaire, ce n'est pas seulement qu'il n'est pas facile de savoir pourquoi, pour les mécanistes, il existe des espèces, mais aussi pourquoi, dans une espèce, il doit avoir deux sexes différents. En liaison avec cela, Catherine Wilson a remarqué³⁹ que s'il n'y a plus de dichotomie entre la forme et la matière dans laquelle la distinction mâle-femelle peut être tracée, alors pourquoi n'y aurait-il pas simplement une parthénogénèse universelle au lieu d'une reproduction par deux sexes distincts ? Comment pouvons-nous expliquer le dyphormisme sexuel à l'intérieur du contexte du mécanisme ? Pour un hylomorphiste, un sexe est responsable de la forme de la progéniture, et l'autre produit la matière. Mais à l'intérieur du contexte du mécanisme, la raison pour laquelle Dieu devrait ne pas avoir arrangé l'épigenèse ou la préformation dans un corps de base type n'est pas claire.

En résumé, c'est Descartes et non Aristote qui doit être au moins un peu en cause dans le problème soulevé par Hull, et pas seulement concernant les créatures générées spontanément, mais aussi celles générées sexuellement. Si nous admettons le problème des espèces, Descartes ne pourrait que conclure, avec Darwin, que leurs origines sont aussi « mineures » que les origines de chaque organisme individuel. Aristote possède une manière satisfaisante d'expliquer un large éventail de phénomènes naturels : toutes les choses naturelles sont une combinaison de matière et de forme ; il y a toute sorte de façons dont une forme particulière peut venir à être inhérente à la matière réceptive appropriée, parmi lesquelles il y a la génération sexuelle, la concoction solaire, etc. La forme se sépare toujours d'elle-même de la matière et ses extensions possibles, en imitation de la pure forme divine, et donc des organismes les plus nobles — ceux capables de se réunir à travers leur propre mouvement pour copuler quand cela est nécessaire — qui seront divisés en mâle et femelle. Mais après que la métaphysique de la forme et de la matière est tombée, il n'existe plus de raison claire expliquant pourquoi, pour Descartes, il y a plutôt un dimorphisme dans les espèces générées sexuellement, *ou* la reproduction régulière d'espèce — ce qui avant aurait pu être décrit comme une répllication formelle — parmi les mâles et les femelles d'une espèce. Avant Descartes, un scholastique aristotélicien aurait pu affirmer par une de ces deux voies : d'un côté, que les anguilles et les bernacles sont de simples réorganisations de la matière qui ne requièrent aucune transmission de forme, et que cela diffère fondamentalement des chevaux et des humains, dont la génération est aussi

39. Dans sa correspondance personnelle.

toujours une *reproduction* ; de l'autre côté que les anguilles et les bernacles sont engendrées à travers la transmission de forme par un flux astral. Mais Descartes ne pourrait prendre aucune de ces voies pour rendre compte de la génération des anguilles et des bernacles ou celle des chevaux et des hommes. La génération de toute créature résulte seulement de « causes mineures », sans aucune transmission de forme.

6. Conclusion

Pourquoi la perspective de la génération par des causes mineures est-elle troublante ? Même Darwin a diagnostiqué le problème de manière perspicace quand il a noté dans *The Descent of Man* que « la naissance des espèces et de l'individu est une partie égale de cette grande séquence d'événements que nos esprits refusent d'accepter comme le résultat d'une chance aveugle⁴⁰ ». Aristote n'était pas troublé en cette matière, puisque pour lui les espèces des créatures générées spontanément ou sexuellement n'étaient jamais des genres réifiés supranaturels et éternels, mais plutôt des régularités morphologiques émergeant en réponse à des régularités environnementales. La réification d'espèces éternelles a débuté avec les exigences des Écritures imposées par la science spéculative des origines, et a été consolidée par la nouvelle « métaphysique appliquée » de la taxonomie botanique qui a commencé au XVI^e siècle. Même dans la vision fixiste des espèces, la génération « spontanée » ne constitue pas nécessairement une rupture dans l'ordre de la nature, puisque, comme les théories du flux cosmique le montrent, il existe une vision valable de la génération non sexuelle en termes de communication d'une forme par un agent — seulement celui-ci n'est pas un agent biologique. La véritable rupture est venue avec l'embryologie mécaniste, qui prive toute forme de génération, sexuelle autant que spontanée, de toute explication téléologique, et qui a milité fortement contre la vision que toute créature individuelle, humaine autant qu'animale, possède des origines naturelles opposées à des origines supranaturelles. Et, comme Darwin l'a observé, une fois que les origines naturelles sont admises pour la génération des individus, le créationnisme est sapé, et il devient possible de penser qu'une espèce entière peut avoir des origines naturelles. Une telle possibilité ressemble à ce qui était sous-jacent aux théories hétérogénétiques des origines de la *megafauna* parmi des libertins du début de la modernité tels que Vanini et La Mothe le Vayer : leur but n'était pas en premier lieu de promouvoir une vision raciste des origines des non-européens ni de simplement tomber dans des fantaisies en se portant garants d'un phénomène naturel qu'ils n'avaient jamais clairement vu. En fait, avec les ressources conceptuelles mises à leur disposition, ils ont plutôt cherché à introduire une solution naturaliste aux origines supranaturelles des espèces dictées par la Genèse. Le fait que la génération spontanée ait joué un rôle si important dans ce basculement

40. Charles Darwin, *The Descent of Man*, in *Darwin : A Norton Critical Edition*. Philip Appleman (ed.), New York, W. W. Norton, 1979, 202f.

du supranaturalisme vers le naturalisme devrait nous amener à repenser la relation historique entre le problème de la génération des choses vivantes, d'un côté, et de l'origine des espèces vivantes, de l'autre.

La prudence de Descartes face à la question de l'infini en mathématiques

Jean-Baptiste Jeangène Vilmer

Volume 34, numéro 2, automne 2007

URI : <https://id.erudit.org/iderudit/016991ar>

DOI : <https://doi.org/10.7202/016991ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Société de philosophie du Québec

ISSN

0316-2923 (imprimé)

1492-1391 (numérique)

[Découvrir la revue](#)

Citer cet article

Jeangène Vilmer, J.-B. (2007). La prudence de Descartes face à la question de l'infini en mathématiques. *Philosophiques*, 34 (2), 295–316.
<https://doi.org/10.7202/016991ar>

Résumé de l'article

La question de l'infini cartésien est vaste et polymorphe, de la métaphysique à la philosophie des sciences en passant par la philosophie pratique. Mais c'est en mathématiques que l'attitude de l'auteur est la plus ambivalente et paradoxale, car il n'y a pas, chez Descartes, d'infini en mathématiques. Le but de cet article est d'analyser les manifestations et les raisons de cette prudence cartésienne. Pour ce faire, nous procédons en deux temps. D'abord, nous constatons l'absence d'infini en mathématiques à travers l'examen de l'infinitésimal. Ensuite, nous montrons qu'il n'y a, dans les mathématiques cartésiennes, qu'un indéfini, en discutant l'existence et le statut du « plus grand nombre ». Ce parcours dévoilera les deux motifs de la prudence cartésienne : la contrainte métaphysique et la rigueur méthodologique.

La prudence de Descartes face à la question de l'infini en mathématiques

JEAN-BAPTISTE JEANGÈNE VILMER

Université de Montréal / McGill University / EHESS

jb.jeangene.vilmer@umontreal.ca

RÉSUMÉ. — La question de l'infini cartésien est vaste et polymorphe, de la métaphysique à la philosophie des sciences en passant par la philosophie pratique. Mais c'est en mathématiques que l'attitude de l'auteur est la plus ambivalente et paradoxale, car il n'y a pas, chez Descartes, d'infini en mathématiques. Le but de cet article est d'analyser les manifestations et les raisons de cette prudence cartésienne. Pour ce faire, nous procédons en deux temps. D'abord, nous constatons l'absence d'infini en mathématiques à travers l'examen de l'infinitésimal. Ensuite, nous montrons qu'il n'y a, dans les mathématiques cartésiennes, qu'un indéfini, en discutant l'existence et le statut du « plus grand nombre ». Ce parcours dévoilera les deux motifs de la prudence cartésienne : la contrainte métaphysique et la rigueur méthodologique.

Mots-clés : Descartes, mathématiques, infini, indéfini, infinitésimal, nombre.

ABSTRACT. — The question of the Cartesian infinite is vast and polymorphic, from metaphysics to the philosophy of science and practical philosophy. But it is in mathematics that the attitude of the author is the most ambivalent and paradoxal because, for Descartes, there is no infinite in mathematics. This article aims to analyze the signs and the reasons of this cartesian prudence. We will proceed in two steps. Firstly, we notice the absence of the infinite in Cartesian mathematics through the examination of the infinitesimal. Secondly, we will demonstrate that there is, in Cartesian mathematics, only an indefinite, by discussing the existence and the status of the «biggest number». This examination will reveal the two motives of cartesian prudence : metaphysical constraint and methodological rigour.

Keywords : Descartes, mathematics, infinite, indefinite, infinitesimal, number.

L'infini est une grille de lecture privilégiée de l'œuvre cartésienne. De la métaphysique à la philosophie pratique en passant par les sciences, la distinction que Descartes opère entre infini (*infinitus*) et indéfini (*indefinitus*) traverse et éclaire les textes, de Dieu au monde, de l'univers à la volonté, de l'atome au plus grand nombre. Une difficulté, cependant, persiste : l'infini en mathématiques. Quand on connaît à la fois l'importance et l'ambivalence du rôle des mathématiques dans l'environnement cartésien, l'attitude de l'auteur, souvent sévère, toujours critique, envers les problèmes posés par ses contemporains, on ne s'étonnera pas que l'infini y occupe une place à l'arrière-scène. La position de Descartes face à la question de l'infini en mathématiques est complexe,

paradoxe, prudente. Pour le dire en un mot : *il n'y a pas, chez Descartes, d'infini en mathématiques.*

À plusieurs endroits, l'auteur fait de sévères critiques à l'encontre des mathématiques, évoquant « les problèmes creux avec lesquels les arithméticiens ou les géomètres ont coutume d'amuser leurs loisirs¹ » : « Rien n'est plus vain que de s'occuper de nombres abstraits et de figures imaginaires, au point de sembler vouloir se contenter de connaître de pareilles bagatelles². » En 1630, il écrivait déjà à Mersenne : « Pour des problèmes, je vous en enverrai un million pour proposer aux autres, si vous le désirez ; mais je suis si las des mathématiques, et en fais maintenant si peu d'état, que je ne saurais plus prendre la peine de les soudre moi-même³. »

L'attitude de Descartes, loin d'être un pur rejet, dissimule en vérité un simple reproche : si l'auteur ne s'occupe guère des mathématiques et en fait un seul loisir, c'est en remarquant combien sont vains les problèmes qu'elle pose, c'est-à-dire combien son abstraction manque l'essentielle visée *pratique* qui doit être celle de l'esprit humain. Les mathématiques cartésiennes, car elles sont malgré tout, ont effectivement, elles, une visée résolument pratique et *ordinaire*⁴. Et s'il en est ainsi, c'est que les mathématiques, pour Descartes, ne sont rien de plus qu'une *fabrication* de l'esprit — thèse qui domina l'école française de la fin du XIX^e siècle et du début du XX^e (Borel, Poincaré, Lebesgue), dont on retrouve encore aujourd'hui les traces dans le constructivisme, et qui oppose ses mathématiques du *faire* à celles de l'*être*, défendues notamment par Cantor et Dedekind.

Dans un tel contexte, on comprendra aisément que la question de l'infini, que déjà Descartes se refuse à expliciter d'une manière générale, en raison de son essentielle incompréhensibilité, soit, en mathématiques, plus encore qu'ailleurs, évitée, dans la lignée de la tradition grecque⁵. L'auteur fait la promesse de ne rien déterminer de cette question :

Ainsi nous ne nous embarrasserons jamais dans les disputes de l'infini ; d'autant qu'il serait ridicule que nous, qui sommes finis, entreprissions d'en déterminer quelque chose, et par ce moyen le supposer fini en tâchant de le comprendre ; c'est pourquoi nous ne nous soucierons pas de répondre à ceux qui demandent si la moitié d'une ligne infinie est infinie, et si le nombre infini est pair ou non pair, et d'autres choses semblables, à cause qu'il n'y a que ceux qui s'imaginent que leur esprit est infini qui semblent devoir examiner telles difficultés⁶.

C'est précisément parce que Descartes n'entend rien *déterminer* de la question de l'infini en mathématiques, que nous pourrions assurer, d'une part, qu'il

1. *Regulae* IV ; AT X 373.

2. *Regulae* IV ; AT X 375.

3. À Mersenne du 15 avril 1630 ; AT I 139.

4. Voir *Regulae* IV ; AT X 373-374.

5. Comme le souligne Belaval, 1960, p. 302.

6. *Principes* I, 26 ; AT IX-2 36.

n'y a pas, chez Descartes, d'infini mathématique et, d'autre part, que les rares réponses que l'auteur ne parvient pas à taire touchant cette question ne laissent voir que son refus d'en *déterminer* quoi que ce soit, c'est-à-dire l'*in-détermination* de l'*in-défini*⁷.

En premier lieu, s'il n'y a pas, chez Descartes, d'infini mathématique, c'est que l'infini cartésien — l'*infinita substantia* — exclut que l'on puisse parler d'un *autre* que lui, qui serait un infini non véritable, et qui pourrait œuvrer en mathématiques. Leibniz ne sera pas si exigeant et proposera que l'on puisse disputer de l'infini en mathématiques, à condition de savoir que cet infini n'est pas le véritable : à Des Bosses, qui lui oppose, d'une manière éminemment cartésienne, « 4- Notre esprit, de ce qu'il est fini, ne peut rien savoir de certain sur l'infini, et partant il convient de n'en jamais disputer », il répond :

Les mathématiciens, si je ne m'abuse, ont déjà réfuté la quatrième, et j'ai moi-même donné quelques spécimens de la science de l'infini. En attendant, je pense qu'à proprement parler l'infini formé de parties n'est ni une unité ni un tout, et qu'on ne le conçoit comme quantité que par pure fiction de l'esprit. Seul l'infini sans parties est un, mais il n'est pas un tout ; cet infini est Dieu⁸.

Mais Descartes, quant à lui, interdit que l'*infinita substantia* puisse côtoyer, et être comme limité par lui, un infini en son genre qui ne serait pas l'infini en tous les genres. La notion d'« infini mathématique » est absurde, dans la mesure, comme le rappellera le cartésianisme de Fénélon, où « l'infini ne peut admettre ni nombre ni augmentation⁹ ». Et c'est pourquoi il n'y a pas, à proprement parler, d'*infini* en mathématiques.

Et en second lieu, si Descartes reste donc indécis sur la question de l'infini en mathématiques, c'est aussi, tout simplement, parce que la mathématique « accoutume à se repaître de vérités (à reconnaître la vérité)¹⁰ » et que, sur des problèmes tels que ceux qui demandent « si la moitié d'une ligne infinie est infinie, et si le nombre infini est pair ou non pair, et d'autres choses semblables¹¹ », on ne saurait prétendre dire une vérité, puisque la vérité, dans ce cas, est au-delà des capacités de notre esprit fini. Or, en ne faisant de la question de l'infini en mathématique que celle de l'indéfini, Descartes s'autorise, non à dire une quelconque vérité sur ces problèmes, mais seulement à décrire l'impuissance dans laquelle il est de ne pouvoir le faire.

Nous bâtissons donc la thèse selon laquelle, chez Descartes, étudier la question de l'infini en mathématiques ne revient qu'à dévoiler et expliquer les signes de sa fuite, en deux moments : d'abord, nous verrons comment il n'y a pas, chez Descartes, d'infini mathématique par le biais de la position

7. Sur ce point, voir notre article « La véritable nature de l'indéfini cartésien », à paraître.

8. Leibniz, 1960, t. II, p. 313.

9. Fénélon, 1712, 2^e partie, chap. V, art. 1, § II.

10. *Entretien avec Burman* ; AT V 177 ; Descartes, 1981, p. 142.

11. *Principes* I, 26 ; AT IX-2 36.

cartésienne touchant la question de l'infinitésimal. Ensuite, nous concluons qu'il n'y a, chez Descartes, et sur l'infini en mathématiques, qu'un *indéfini*, en discutant l'existence et le statut du « plus grand nombre ».

I. L'infinitésimal

Si l'on pense presque systématiquement à Leibniz, en ce qui concerne la question de l'invention du calcul infinitésimal, on oublie trop souvent trois sources. Premièrement, le travail initiateur de Fermat qui, dès 1632, dans son *De Maximis et Minimis*, utilise ce qui pourrait être tenu pour le premier calcul infinitésimal, à savoir la première méthode d'algébrisation permettant de trouver la tangente à une courbe : dans ses calculs, Fermat use effectivement d'une quantité infinitésimale, notée e , et qui, bien que non nulle, peut être supprimée ; ainsi passe-t-il de l'*adégalité* « $b (2a + e)$ » à l'*égalité* « $b = 2a$ ». Reste que l'infinitésimal n'est pas chez Fermat la plus belle manifestation de la question de l'infini en mathématiques : il s'agirait plutôt de la descente infinie (ou indéfinie : Fermat sur ce point utilise la même précaution de vocabulaire), laquelle peut être considérée comme le point de départ du courant constructiviste sur la question, comme on peut le voir chez Yvon Gauthier. Deuxièmement, Pascal, dont Leibniz lui-même reconnaîtra l'influence sur son propre calcul¹², et qui écrit à Fermat, comme à son allié, au sujet de son adversaire le Chevalier de Méré :

Je n'ai pas le temps de vous envoyer la démonstration d'une difficulté qui étonnait fort M. de Méré, car il a très bon esprit, mais il n'est pas géomètre ; c'est, comme vous savez, un très grand défaut, et même il ne comprend pas qu'une ligne mathématique soit divisible à l'infini et croit fort bien entendre qu'elle est composée de points en nombre fini et jamais je n'ai pu l'en tirer ; si vous pouviez le faire, on le rendrait parfait¹³.

Et troisièmement, Newton, dont la méthode des fluxions n'est pas non plus à négliger¹⁴.

Et s'il est si répandu de faire naître le calcul infinitésimal avec Leibniz, c'est qu'il est tout autant remarquable que Descartes l'a à ce titre effleuré, ne pouvant revendiquer pour sa part que la paternité de la géométrie moderne,

12. Faisant référence à une lettre de M. Dettonville de la fin de 1658 (Amos Dettonville est le pseudonyme que Pascal prend fin 1658) à M. de Carcavi et, parlant de lui-même, Leibniz écrit, dans son *Histoire et origine du Calcul différentiel* : « De retour d'Angleterre en France vers 1673, l'auteur se mit, sur le conseil de Huygens, à étudier l'Analyse cartésienne, et pour recevoir une introduction à la géométrie des quadratures, il consulta : la *Vue d'ensemble de la géométrie* de l'estimable Fabri, Grégoire de Saint-Vincent et un petit livre de Dettonville (c'est-à-dire de Pascal) ». Le texte de Pascal en question est celui de la *Lettre de A. Dettonville à Monsieur de Carcavy*, imprimée entre décembre 1658 et janvier 1659, et reproduite dans Pascal, 1908-1914, t. VIII, p. 325-384.

13. Pascal à Fermat du 29 juillet 1654, dans Pascal, 1908-1914, t. III, p. 375-393.

14. Cléro, 1989-90, et Gandt, 1992, soulignent combien, dans l'invention du calcul infinitésimal, Newton a joué un rôle aussi important que celui de Leibniz.

c'est-à-dire algébrique (ou analytique¹⁵), comme le rappelle Louis Couturat, répartissant ainsi les mérites : « Descartes, le père de la Géométrie moderne, et Leibniz, l'inventeur du Calcul infinitésimal¹⁶ ». Il s'agit donc pour nous de comprendre dans quelle mesure on peut reconnaître à Descartes pour une petite part d'en avoir été le précurseur, et pour une grande part de l'avoir formidablement — et délibérément — manqué.

A. La réticence de Descartes face à la question de l'infinitésimal

En quoi Descartes peut-il être vu comme un précurseur du calcul infinitésimal, voilà qui peut être réglé rapidement, car si l'on doit reconnaître qu'il s'approche parfois très près des prémisses de l'infinitésimal, le fait qu'il feigne de ne point le remarquer pour l'éviter soigneusement ampute systématiquement les ambitions qu'on lui pourrait attribuer. Certains interprètes, comme Sergescu, persistent malgré tout à lui accorder, sur ce point, un rôle non négligeable : « Il faut dire que Descartes (1596-1650), sans s'occuper de la notion d'infini mathématique, a joué un très grand rôle dans le développement de tous ces calculs infinitésimaux ; on verra plus loin que sans lui on n'aurait pas pu avoir Leibniz¹⁷. » Pour quelles raisons, donc, peut-on reconnaître que Descartes, d'une certaine manière, a conditionné le calcul infinitésimal à venir ?

On pourrait voir au moins deux arguments. D'une part, dans la mesure où le calcul différentiel, qui exige des méthodes infinitésimales, est lui-même construit sur la base de l'algèbre cartésienne, on peut indirectement accorder à Descartes, certes à son corps défendant, un rôle dans la genèse du calcul infinitésimal, même sans qu'il ait jamais lui-même joint l'usage des séries infinies à son algèbre pour engendrer ce qui aurait pu devenir le calcul différentiel¹⁸. Et, d'autre part, on peut aussi attribuer à Descartes la découverte d'un principe essentiel du calcul infinitésimal, pour avoir été le premier à voir que l'intégration est une opération inverse de la détermination de la tangente à une courbe (c'est-à-dire la différenciation), dans une lettre à Florimond de Beaune¹⁹.

Néanmoins, et quant à se demander si Descartes, pour ces deux raisons, possédait en conséquence une sorte d'équivalent du calcul infinitésimal avant l'heure, Vuillemin prévient fort justement que

l'on a coutume, depuis, Lazare Carnot²⁰, de prétendre que Descartes possédait un équivalent du Calcul infinitésimal dans la méthode des indéterminées. Cette

15. Bien qu'étant la plupart du temps dite « analytique », sa véritable *nature* est sans conteste algébrique : l'analytique ne touchant que sa *lecture*. Voir Brunschvicg, 1904, p. 770 : « La géométrie de Descartes [...] est proprement une algèbre » ; et 1927, p. 286 : « Il crée une géométrie qui est avant tout une algèbre. »

16. Couturat, 1896, p. 265.

17. Sergescu, 1955, p. 70.

18. C'est la conclusion de Houzel, 1997, p. 35.

19. À Florimond de Beaune du 20 février 1639 ; AT II, 513-514.

20. Voir Carnot, 1797 et, sur lui, Gillispie et Iushkevich, 1979.

affirmation n'est véritable qu'à une double condition : d'abord faire abstraction du fondement que Carnot prête à cette méthode, y voyant faussement à l'œuvre une compensation d'erreurs ; ensuite en éliminer toute notion d'« équation imparfaite » et d'approximation, étrangères à l'Algèbre²¹.

Et il se trouve que Descartes, précisément, se rend aveugle de la naissance du calcul infinitésimal pour y refuser l'*approximation*, selon les exigences de son algèbre propre. C'est ainsi qu'il y a bien plus à dire, concernant la question de l'infinitésimal chez Descartes, en examinant la manière dont il a su, non y répondre, ou même la poser, mais tout simplement la *contourner*. On retiendra trois expressions.

Premièrement, par la seule *attitude* de Descartes, qui reproche à ses correspondants de travailler l'infinitésimal, ou encore leur dénonce les travaux de ceux qui s'y occupent. L'attitude générale de Descartes sur la question des infiniment petits en mathématiques est effectivement, et sans conteste, la plus simple *réticence*, comme en témoigne sa correspondance avec Fermat et sur Galilée.

La querelle qui oppose Descartes à Fermat²² révèle précisément en quoi la critique cartésienne de la conception de la tangente comme d'un maximum ou minimum d'un segment sécant à la parabole est celle de l'infinitésimal et, plus largement, de la question de l'infini en mathématiques. Ce que Descartes refuse à Fermat est l'*évanouissement* d'une quantité quelconque : en passant de l'adégalisation qui approxime *e* comme *tendant* vers 0 à l'égalité qui l'y rend identique pour pouvoir la rendre évanescence, Fermat semble considérer une *même* quantité à la fois comme *différente* de 0 et comme *nulle*, ce qui implique une contradiction, et est problématique. Un siècle plus tard, Berkeley formulera la même critique des méthodes infinitésimales en doutant notamment de la validité d'une mathématique qui utiliserait une grandeur *non nulle mais tenue pour nulle*. Dans un court texte intitulé, *Of Infinites*, et présenté le 19 novembre 1707, Berkeley montre qu'un infinitésimal « n'est absolument rien²³ ». En 1734, dans *The Analyst*, il développera méthodiquement de nombreuses objections à l'encontre des méthodes infinitésimales de toutes sortes.

Descartes, en un mot, ne se permet pas ce *saut* de l'infiniment petit au néant, puisqu'entre les deux il ne saurait y avoir une différence de degré — c'est précisément la raison d'être de l'infinitésimal, qui reste hors du 0 pour ne jamais épuiser tous les degrés qui, théoriquement, mènent à lui — mais de nature. C'est pourquoi la correspondance de Descartes avec Fermat montre que le premier pense du second qu'il est bon mathématicien (on ajoutera : sans doute, même, supérieur à Descartes), mais piètre philosophe.

21. Vuillemin, 1960, p. 62.

22. Sur la correspondance Descartes — Fermat, voir Mahoney, 1973, p. 170-193.

23. Berkeley, 1985, p. 146.

Quant à Galilée²⁴, Descartes écrit à Mersenne que son procédé annonçant le calcul infinitésimal lui semble n'être qu'un sophisme :

[Page 19²⁵] Il examine la colle qu'il ajoute avec le vide pour la liaison des parties des corps, et il l'attribue à d'autres petits vides qui ne sont nullement imaginables. Et ce qu'il dit (p. 22), pour prouver ces petits vides, est un sophisme ; car l'hexagone qu'il propose ne laisse rien de vide en l'espace par où il passe, mais chacune de ses parties se meut d'un mouvement continu, lequel décrivant des lignes courbes qui remplissent tout un espace, on ne doit pas les considérer, comme il fait, en une seule ligne droite. Et il n'importe qu'en sa figure les parties de la ligne droite, IO, PY, etc., ne soient point touchées par la circonférence H I K L, car elles le sont en récompense par d'autres parties de la superficie A B C, et ainsi ne sont non plus vides que les parties OP, YZ, etc. [Page 28] C'est aussi un sophisme que son argument, pour prouver qu'un point est égal à une ligne ou à une superficie. Car *in forma* on ne peut conclure, sinon que la ligne ou superficie n'est pas un plus grand corps solide que le point, et non qu'elle n'est pas plus grande absolument. [Page 31] Il manque en tout ce qu'il dit de l'infini, en ce que, nonobstant qu'il confesse que l'esprit humain, étant fini, n'est pas capable de le comprendre, il ne laisse pas d'en discourir tout de même que s'il le comprenait²⁶.

On notera que Galilée lui-même reconnaît que notre entendement fini ne peut comprendre adéquatement ces problèmes touchant l'infini, mais, contrairement à Descartes, qui peut-être est plus exigeant, il invoque notre aspiration vers cet infini pour défendre le fait d'y travailler malgré tout : « Rappelons-nous que nous traitons d'infinis et d'indivisibles, inaccessibles à notre entendement fini, les premiers à cause de leur immensité, les seconds à cause de leur petitesse. Pourtant nous constatons que la raison humaine ne peut s'empêcher de sans cesse y revenir²⁷. »

Deuxièmement, l'esquive cartésienne se manifeste également dans le problème de la quadrature de la cycloïde, pour lequel Descartes ne s'embarrasse pas de l'infini. On lit dans les lettres à Mersenne des 27 mai et 27 juillet 1638 que, pour conclure à l'égalité de deux surfaces (l'arc de la cycloïde et un cercle appelé générateur), Descartes ne conçoit pas chacune des surfaces comme une somme (infinie) de lignes droites²⁸, ce qui exigerait de créer un algorithme d'intégration,

24. Sur Descartes et Galilée, voir Enriques, 1937 ; Banfi, 1957 ; et Marion, 1981, p. 203-227. Plus précisément, sur la critique que fait Descartes de Galilée, voir Shea, 1978.

25. Descartes commente le livre de Galilée *Discorsi e dimostrazioni matematiche, intorno a due nuove scienze* (Discours concernant deux sciences nouvelles) récemment paru le 6 mars 1638.

26. À Mersenne du 11 octobre 1638 ; AT II, 382-384. Alquié, 1992, p. 94, n. 2, objecte que Descartes agit exactement de même : « cette remarque est curieuse, Descartes agissant exactement de la même façon ». Il faut lui répondre que, si c'était vraiment le cas, nous aurions peut-être une doctrine systématisée de l'infini chez Descartes. Mais, au contraire, il semble que Descartes n'en discute pas beaucoup. Et pour ce qu'il en discute, il le fait *tout en signalant constamment* l'incompréhensibilité de l'infini, ce qui n'est pas négligeable.

27. Galilée, 1970, p. 26.

28. Notamment en raison de l'infinie divisibilité de la matière : on ne cessera jamais de trouver une droite plus étroite que la précédente.

mais il se contente de faire passer par n'importe quel point de leur hauteur une parallèle qui sera égale dans l'une et dans l'autre des figures (de telle sorte qu'il pourra conclure à leur égalité). Et, sur ce point, il est si proche de Cavalieri²⁹ — l'« inventeur » d'une telle méthode — qui résout le problème en balayant de la même façon et par une règle les deux surfaces, qu'on pourrait lui reconnaître, comme le fait Belaval, de l'avoir même mieux compris que ne le fera Leibniz³⁰. Pascal, à l'inverse, s'insurgera que l'on ne veuille point donner de sens à l'expression « somme de droites », et sera en conséquence bien moins prudent de s'embarrasser des disputes de l'infini, n'hésitant pas à parler de la « somme de *tous* les arcs » d'un cercle ou encore d'une « somme d'une *multitude indéfinie* de lignes », alors même que les définitions de telles expressions restent floues³¹.

Troisièmement, il faut enfin examiner l'exclusion des passages à la limite³². C'est certainement là l'exemple le plus manifeste de la manière dont Descartes a soigneusement évité l'infinitésimal et tout ce qui était présupposé par lui ou pouvait y conduire. On le remarque à plusieurs endroits.

En métaphysique, tout d'abord, parce que le passage à la limite, qui ici est celui d'une quantité infinitésimale au néant du 0, serait celui du fini à l'infini, alors même que la première règle, touchant l'infini, est celle du tout ou rien, qui interdit qu'une échelle de degrés, même de longueur infinie, puisse le joindre au fini et le décliner en une infinité d'infinis dégradés³³.

En physique, également, à travers la discontinuité entre repos et mouvement et la négation que la vitesse doit passer par tous les degrés pour le mouvement accéléré³⁴, sans compter la solution à l'« Achille » de Zénon qui, en montrant comment la somme des termes d'une série convergente ne *tend* pas

29. Ce même Cavalieri qui, huit ans plus tard, et toujours dans une lettre à Mersenne, se verra, cette fois, et selon l'auteur, redevable de Descartes : « Ce Cavalieri ne fait autre chose que de démontrer, par un nouveau moyen, des choses qui ont déjà été démontrées par d'autres. » Descartes reconnaît alors : « Ce nouveau moyen n'est autre que l'un de ceux dont je me suis servi pour démontrer la roulette, en supposant que deux triangles curvilignes différents étaient égaux, pour ce que toutes les lignes droites tirées en même sens en l'un qu'en l'autre étaient égales » (à Mersenne le 20 avril 1646 ; AT IV, 395).

30. Belaval, 1960, p. 302.

31. En cela, ne déroge-t-il pas à sa propre règle qui demande de « n'employer dans la définition des termes que des mots parfaitement connus ou déjà expliqués » ? (Pascal, 1908-1914, t. IX, p. 279.)

32. Le passage à la limite est, par exemple, l'équation suivante : $1/2 + 1/4 + 1/8 + \dots + 1/2n + \dots = 1$. Lorsque l'on demande alors : quel est le sens de l'égalité ? on devrait répondre : *a priori*, elle a bien peu de sens, puisque la sommation ne peut être achevée. C'est pourquoi, après Cauchy, on écrira : *limite* de cette somme = 1.

33. On prendra « dégradés » dans les deux acceptions : en *dégradant* l'infini, c'est-à-dire en le diluant, par degrés, dans le non infini, on le *dégrade*, du point de vue de sa valeur : il perd sa perfection, et du même coup son infinité.

34. Descartes, dans les lettres à Mersenne des 17 novembre et 7 décembre 1642, réfute l'idée qu'un corps qui se met en mouvement doive passer par tous les degrés de vitesse, en donnant l'exemple d'un grand corps dur qui percute dans son mouvement un petit corps au repos : ce dernier prend alors *immédiatement* sa vitesse finale.

vers une limite finie, mais est *égale* à une quantité finie, s'extrait de toute compromission avec l'idée de limite.

Il ne faut pas croire ensuite que la courbe proposée à de Beaune, évoquée tout à l'heure, et qui semble utiliser la méthode des limites, le fasse réellement, puisque, comme le rappelle Yvon Belaval, « telle qu'elle se présente chez lui, elle procède par exhaustion de « parties » finitésimales — et non infinitésimales ; elle reste étrangère à la notion de convergence qui ne sera dégagée que par Gregory et reprise par l'anticartésien Wallis³⁵ ».

Enfin, le refus d'utiliser la notion de « passage à la limite » doit finalement se comprendre à partir de l'idée de limite elle-même, que Descartes ne se résigne pas à accepter claire et distincte, comme le note Vuillemin :

L'idée de limite, dont le philosophe sent au moins confusément qu'elle fonde le Calcul infinitésimal, et qui implique celle d'approximation à l'infini, ne saurait être reçue par lui comme exacte ou distincte. Les courbes transcendantes assignent donc à l'entendement tel qu'il le définit en Mathématiques une limite infranchissable, qui leur confère dans le système de Descartes un rôle analogue à celui que jouent chez Kant les idées pures de la raison que nous ne saurions nous représenter dans une intuition sensible³⁶.

C'est d'ailleurs pourquoi Belaval reproche justement à Gueroult d'avoir fait de la limite une notion cartésienne dans son *Descartes selon l'ordre des raisons*, le tirant vers Leibniz d'une manière illégitime³⁷. Partant de là, le raisonnement est simple : l'idée de limite n'est pas elle-même claire et distincte. Or la Règle VIII nous apprend que : « Si dans la série des objets de recherche il s'en présente un dont notre entendement ne puisse acquérir une intuition satisfaisante, il faut s'arrêter là : on ne doit point examiner ceux qui suivent, mais se dispenser d'une peine superflue³⁸. » Donc, il se comprend aisément que Descartes ne puisse accepter tout ce qui suit du problème de la limite, c'est-à-dire, d'une manière générale, la question de l'infinitésimal elle-même.

Aussi est-il absurde d'écrire, comme le fait J. D. North, que le concept cartésien de limite non seulement est à l'origine de celui de l'infini mais encore influencera Newton, dont il est dit : « *In playing with words, making limit a negative concept and therefore infinite a positive one, he is following in Descartes' footsteps*³⁹. » Il y a ici au moins trois erreurs. Premièrement, donner au concept de limite une place fondamentale chez Descartes. Deuxièmement, déduire (« *therefore* ») la positivité de l'infini de la négativité de la limite alors même que Descartes ne cesse de répéter la priorité d'un infini

35. Belaval, 1960, p. 301.

36. Vuillemin, 1960, p. 95.

37. Belaval, 1955, p. 442-443 : « On pourrait contester l'authenticité cartésienne de certaines notions. Nous ne sommes pas sûr que, parlant de limite (I, 61, 220-221, etc.), M. Gueroult n'introduise en son commentaire une notion leibnizienne. »

38. *Regulae* VIII ; AT X, 392. Cette règle constitue aussi une justification à ce qu'il ne faille point tenter de comprendre l'infini.

39. North, 1983, p. 122.

dont la positivité ne signifie d'ailleurs rien d'autre (de telle sorte que « déduire la positivité de... » est une contradiction dans les termes, puisque la positivité est, par définition, première). Troisièmement, déclarer Newton cartésien pour avoir repris cette structure qui n'est pas chez Descartes !

En conclusion, si Descartes semblait avoir à sa portée une bonne partie des outils nécessaires à la génération du calcul infinitésimal, il n'en fit rien. Paralysé par le manque de rigueur de cette fiction seulement utile, il refuse de « négliger comme nulles »⁴⁰ — selon l'expression de Pascal qui, lui, l'acceptera sans peine — les quantités infiniment petites. C'est la méthode cartésienne qui dédaigne ici le peu de rigueur du « négliger » de l'évanouissement et demande à la démonstration mathématique d'être un continu — ou continûment légitimée — sans jamais laisser se faire, ne serait-ce qu'un instant, un *saut* illégitime, fût-il pragmatique. Il applique, en quelque sorte, sa *rhétorique de l'attention*, qui est une *rhétorique contre la relâche*⁴¹, au domaine des mathématiques⁴².

B. Réfutation d'un soi-disant cas d'utilisation du passage à la limite en métaphysique

Vuillemin remarque qu'un certain texte des *Réponses aux Quatrièmes Objections* semble utiliser le passage à la limite⁴³, alors que celui-ci est refusé aux mathématiques. Si, comme nous le défendons, la fuite de Descartes est liée à des impératifs métaphysiques, pourquoi donc, au sein même de sa métaphysique, y aurait-il précisément ce avec quoi elle est prétendument incompatible ? Il nous faut examiner l'argument de Vuillemin de plus près.

Dans les *Quatrièmes Objections*, Arnauld refuse la preuve cartésienne par les effets : le principe de causalité ne s'applique qu'aux choses finies (il faut

40. « On n'augmente pas une grandeur continue lorsqu'on lui ajoute, en tel nombre que l'on voudra, des grandeurs d'un ordre d'infinitude inférieur. Ainsi les points n'ajoutent rien aux lignes, les lignes aux surfaces, les surfaces aux solides ; ou, pour parler en nombres — comme il convient dans un Traité d'Arithmétique —, les racines ne comptent pas par rapport aux carrés, les carrés par rapport aux cubes et les cubes par rapport aux carro-carrés. En sorte qu'on doit *négliger comme nulles*, les quantités d'ordre inférieur » (Pascal, 1908-1914, t. III, p. 367). Voir aussi : « J'estime qu'il n'est pas légitime et conforme à la géométrie d'admettre la méthode de démontrer par les indivisibles ou (comme j'ai coutume de les nommer) par les hétérogènes, que le célèbre géomètre Bonaventure Cavalieri a mis en lumière. Cette méthode célèbre passe des lignes aux surfaces, des superficies aux solides, et une égalité ou proportion trouvée pour les lignes donne une conclusion appliquée aux superficies. Par cette manière de raisonner on n'aboutit absolument à rien de certain à moins qu'on ne la ramène à l'homogénéité » (cité par Russo, 1962, p. 308).

41. Sur la relâche, voir par exemple les lettres à Elisabeth, de mai ou juin 1645, AT IV, 220, et du 6 octobre 1645, AT IV, 307.

42. Parce que « je ne puis pas ouvrir les yeux des lecteurs, ni les forcer d'avoir de l'attention aux choses qu'il faut considérer pour connaître clairement la vérité [...] je ne puis pas donner de l'esprit aux hommes, ni faire voir ce qui est au fond d'un cabinet, à des gens qui ne veulent pas entrer dedans pour le regarder » (à Mersenne, du 21 janvier 1641 ; AT III, 283-285), l'auteur s'adresse à l'*attentus lector*, au « *mentis purae et attentae* » (*Regulae*, III, AT, X, 368). La notion de « *Rhetoric of attention* » domine le chapitre III de Carr, 1990. Voir aussi Rodis-Lewis, 1950, p. 182.

43. Vuillemin, 1960, note III, p. 144-146.

que la cause et l'effet soient distincts : la notion de *causa sui* est donc contradictoire dans les termes).

Descartes répond ceci :

Certes, il me semble que M. Arnauld a fait en ceci la même chose que si (après Archimède, parlant des choses qu'il a démontrées de la sphère par analogie aux figures rectilignes inscrites dans la sphère même, aurait dit : si je pensais que la sphère ne pût être prise pour une figure rectiligne, ou quasi rectiligne, dont les côtés sont infinis, je n'attribuerais aucune force à cette démonstration, parce qu'elle n'est pas véritable, si vous considérez la sphère comme une figure curviligne, ainsi qu'elle est en effet, mais bien si vous la considérez comme une figure rectiligne dont le nombre de côtés est infini). Si, dis-je, M. Arnauld, ne trouvant pas bon qu'on appelât ainsi la sphère, et néanmoins désirant retenir la démonstration d'Archimède, disait : si je pensais que ce qui se conclut ici, se dût entendre d'une figure rectiligne dont les côtés sont infinis, je ne croirais point du tout cela de la sphère, parce que j'ai une connaissance certaine que la sphère n'est point une figure rectiligne. Par lesquelles paroles il est sans doute qu'il ne ferait pas la même chose qu'Archimède, mais qu'au contraire il se ferait un obstacle à soi-même et empêcherait les autres de bien comprendre sa démonstration⁴⁴.

Vuillemin commente :

Descartes répond par un raisonnement qui utilise le passage à la limite. Comme Archimède démontre des propriétés pour le solide rectiligne inscrit dans la sphère et l'applique à la sphère même, conçue comme un solide ayant une infinité de faces, de même le philosophe est légitimé, si le principe de causalité peut être à bon droit appliqué aux choses et aux idées des choses finies, à l'appliquer à l'infini même et à l'idée que nous en recevons⁴⁵.

Or nous contestons la pertinence d'une équivalence entre « le passage à la limite en mathématiques », lequel est quantitatif, et « le passage à la limite en métaphysique », qui, lui, est qualitatif : Descartes ne fait qu'utiliser un raisonnement qui n'a de « passage à la limite » que la forme. Il n'est donc pas ici question du problème de la limite tel qu'il se manifeste en mathématiques.

Par ailleurs, Vuillemin ajoute que Descartes semble utiliser un raisonnement qui apparaîtra explicitement chez Leibniz sous le nom de *principe de continuité* et se manifeste ainsi dans le texte leibnizien :

Lorsque la différence de deux cas peut être diminuée au-dessous de toute grandeur donnée, *in datis*, ou dans ce qui est posé, il faut qu'elle puisse se trouver aussi diminuée au-dessous de toute grandeur donnée *in quaesitis*, ou dans ce qui en résulte ; ou, pour parler plus familièrement, lorsque les cas (ou ce qui est donné) s'approchent continuellement et se perdent enfin l'un dans l'autre, il faut que les suites, ou événements (ou ce qui est demandé) le fassent aussi⁴⁶.

44. *Réponses aux Quatrièmes Objections* ; AT IX-1 189.

45. Vuillemin, 1960, 145. Alquié, 1950, p. 225-226, parle lui aussi de « passage à la limite, ainsi opéré en ce qui concerne le principe de causalité ».

46. Leibniz, *Nouvelles de la République des Lettres*, mai 1687, dans Leibniz, 1960, p. 744.

Le principe de continuité leibnizien enveloppe ainsi un passage aux limites : Leibniz applique à sa métaphysique les procédés inventés pour le calcul infinitésimal. Mais il ne nous semble pas non plus adéquat d'attribuer ce principe de continuité à Descartes, sur le problème du fini et de l'infini, puisque la continuité entre ces deux notions est chez lui inacceptable, en ce qu'il n'y a pas entre elles une différence de *degré*, mais de *nature*. Cela est confirmé par Belaval, qui infirme l'attribution de la notion de limite que fait Gueroult à Descartes, par le fait qu'une telle notion présuppose un principe de continuité que Descartes est loin de considérer : « Cette notion implique un principe de continuité que Descartes, non seulement n'a pas généralisé à tout le domaine mathématique (puisque'il ne peut passer des courbes algébriques aux courbes transcendentes), mais encore qu'il n'a pas songé à étendre à sa Mécanique⁴⁷. » Il faut conclure que Vuillemin, comme Gueroult, tire ici Descartes vers Leibniz.

En conclusion, il ne semble pas que ce prétendu cas d'utilisation de passage à la limite en Métaphysique soit inquiétant pour notre affaire.

II. L'indéfini : existence et statut du « plus grand nombre »

On sait donc comment Descartes a contourné la question de l'infinitésimal, comme celle de l'infini en mathématiques. Reste que, malgré tout, il ne peut éviter de la rencontrer, ici et là, dans un traité de géométrie comme au détour d'une correspondance métaphysique. Comment, dans ce cas, réagit-il ? Il ne répond à la question que par l'*indétermination* de l'*in-défini*. Et l'exemple parfait d'un tel comportement se trouve particulièrement manifeste dans le problème du nombre infini, ou du « plus grand nombre ». Nous verrons, tout d'abord, comment Descartes, par la finitude de l'entendement condition du nombre, écarte d'emblée la notion même de « nombre infini », dans le but de mieux substituer à l'infinité de la formule classique, et ici refusée, le simple maximum d'un « plus grand nombre ». Il s'agira ensuite de préciser l'in-défini de ce plus grand nombre, pour ne pas pouvoir prouver son existence : car c'est effectivement de cette position concernant son existence que découlera son statut, à savoir celui de nombre indéfini.

A. Comment la finitude de l'entendement infirme la notion de « nombre infini »

À la base des mathématiques se trouve le nombre. À la base de la question de l'infini en mathématiques, donc, se trouve celle du nombre infini. Or, le nombre relève de l'entendement. Si donc il peut seulement ensuite être dit lui-même au fondement des mathématiques, c'est que, par continuité, les mathématiques, à travers le nombre, ont pour condition l'entendement. Or, et c'est de là que vient le problème lorsque se pose la question de l'infini en mathématiques, l'entendement est fini : « ma nature [...] est finie et bornée⁴⁸ ».

47. Belaval, 1955, p. 443.

48. *Méditation* III ; AT IX-1 37.

La conclusion ne saurait tarder : n'ont de sens ni un « nombre infini » ni toute investigation « infinitisante » en mathématiques.

C'est donc bien la finitude de l'entendement qui est le nerf du rejet cartésien de l'infini mathématique, comme le confirmera Cantor dans une critique acerbe :

Les mots « entendement fini » que l'on entend si souvent n'offrent, je le pense, aucun caractère pertinent : si bornée que soit en vérité la nature humaine, elle porte pourtant en elle, c'est inhérent, une très grande part d'infini, et j'irais jusqu'à dire que si elle n'était pas elle-même infinie sous bien des rapports, on ne saurait expliquer la ferme assurance et la pleine certitude où nous nous savons tous unis, s'agissant de l'être de l'absolu⁴⁹.

Le fait que Cantor, pour asseoir sa position infinitiste, soit amené à s'en prendre à la finitude de l'entendement cartésien non seulement confirme, d'une part, que, d'une manière générale, l'entendement est condition du nombre et, d'autre part, que, chez Descartes, la finitude de l'entendement est condition du rejet de l'infinité du nombre, mais aussi explique pourquoi Cantor préférera Spinoza, chez qui l'entendement est infini — puisqu'il *comprend*⁵⁰ l'infini — pour jeter les bases philosophiques de son infinitisme mathématique tel qu'il est à l'œuvre dans sa théorie des ensembles⁵¹.

B. Qu'on ne peut prouver l'existence du « plus grand nombre » : indéfinité
Descartes rejette donc d'emblée la notion de « nombre infini ». Il n'en faudra pas moins à Leibniz, qui sur ce point partage les vues cartésiennes, pour en déduire sa simple finitude : si le nombre n'est pas infini, c'est qu'il est fini. C'est effectivement en niant l'existence de l'infini catégorématique, infini dénombrable (« Il n'est pas donné d'Infini catégorématique, c'est-à-dire ayant en acte des parties infinies formellement⁵² »), que Leibniz en vient à nier la notion de nombre infini : « Il n'y a point de nombre infini, ni de ligne ou autre quantité infinie, si on les prend pour de véritables tous [...] Les écoles ont voulu dire cela en admettant un infini syncatégorématique, comme elles parlent, et non pas

49. Cantor, 1932, p. 176-177.

50. À prendre dans les deux sens : si précisément l'entendement comprend (*comprehendere*) l'infini, c'est pour le comprendre (inclusion, qui serait identité puisqu'on ne peut enfermer l'infini : l'entendement *est* lui-même infini).

51. On notera cependant que l'infinitude de l'entendement n'empêche pas Spinoza d'établir l'impossibilité d'un nombre infini, puisque « l'infini et le nombre s'excluent » (Gueroult, 1966, p. 401) : « Ni le nombre, ni la mesure, ni le temps, puisqu'ils ne sont que des auxiliaires de l'imagination, ne peuvent être infinis, car autrement le nombre ne serait plus nombre, ni la mesure, mesure, ni le temps, temps ; de là on voit clairement pourquoi, confondant ces trois êtres de raison avec les choses réelles dont on ignore la vraie nature, on nie fréquemment l'infini en acte » (Lettre XII à Louis Meyer, dans Spinoza, 1954, p. 1097). Si l'infini et le nombre s'excluent, c'est que, pour le coup, comme chez Descartes, seul l'entendement peut concevoir l'infini, alors que le nombre est généré par l'imagination qui, elle, ne peut s'imaginer l'infini.

52. À Des Bosses, le 1^{er} septembre 1706, dans Leibniz, 1960, t. II, p. 314.

l'infini catégorématique⁵³. » Et, de cette négation, il déduit l'affirmation de sa finitude : « Il appartient à l'essence du nombre, de la ligne, et d'un tout quelconque, d'être borné⁵⁴. » La position de Descartes est plus complexe pour au moins deux raisons.

Premièrement, son rejet de la *notion* de « nombre infini » n'implique pas le rejet de son *existence*, qu'il ne trouve pas *a priori* absurde. À Burman qui demande : « Mais il y aurait dans l'hypothèse envisagée un nombre infini ? », il aurait répondu : « Et où serait l'absurdité ? »⁵⁵. Hannequin s'empresse d'en déduire que le plus grand nombre cartésien, pour cette raison et le fait que, chez Descartes, un infini puisse être plus grand qu'un autre, n'annonce rien de moins que le transfini cantorien :

Ainsi Descartes ne trouve absurde *a priori* ni l'existence d'un nombre infini [...] ou même celle de plusieurs nombres infinis différents les uns des autres ni la possibilité de concevoir entre eux des rapports qui ne répondent point nécessairement aux rapports des nombres finis ; et nous nous contentons de dire qu'en cela il ne faisait qu'énoncer des propositions rendues plus que plausibles par de récentes spéculations sur l'infini proprement dit ou sur le *transfini*⁵⁶ [voir notamment les travaux de M. George Cantor] .

C'est que, de la même manière que l'existence d'un nombre infini ne serait pas absurde, il n'est pas absurde qu'un infini soit plus grand qu'un autre : on pourrait par exemple penser que la série « infinie » des nombres n'a pas de sens, que nous la posons comme telle, sans même savoir ce qu'elle signifie, puisque la série des nombres entiers et celle des nombres pairs constituent deux infinis dont l'un est le double de l'autre — Cantor, avant Russell, a justement montré qu'il y avait bijection entre ces deux ensembles — ce qui semble absurde, pour beaucoup, dont Spinoza lui-même : « Si chaque partie est infinie, il y a donc un infini deux fois plus grand qu'un autre infini, ce qui est également absurde⁵⁷. » Mais, pour Descartes, au contraire, il n'y a rien d'absurde à ce qu'un infini soit plus grand qu'un autre :

Mais, à propos de l'infini, vous m'en proposiez une question en votre lettre du 14 mars, qui est tout ce que j'y trouve de plus qu'en la dernière. Vous disiez que, s'il y avait une ligne infinie, elle aurait un nombre infini de pieds et de toises, et par conséquent que le nombre infini des pieds serait six fois plus grand que le nombre des toises⁵⁸. — *Concedo totum* [« j'accorde tout cela »]. — Donc ce dernier n'est pas infini. — *Nego consequentiam* [« Je nie la conséquence »]. — Mais un infini ne peut être plus grand que l'autre. — Pourquoi non ? *Quid*

53. Leibniz, 1990, p. 124.

54. À Des Bosses le 11 mars 1706, dans Leibniz, 1960, t. II, p. 304.

55. *Entretien avec Burman* ; AT V, 155 ; Descartes, 1981, p. 52.

56. Hannequin, 1896, p. 447, note.

57. *Ethique*, prop. 15, scolie, in Spinoza, 1954, 380. Le scolie de la prop. 15 discute de manière substantielle de divers problèmes touchant l'infini, dans le but d'établir l'infini de l'éternelle, en tant qu'elle *exprime* un Dieu infini.

58. 1 toise = 6 pieds.

absurdi ? [« Qu'y a-t-il d'absurde en cela ? »]. Principalement s'il est seulement plus grand *in ratione finita, ut hic ubi multiplicatio per 6 est ratio finita, quæ nihil attinet ad infinitum* [« Selon un rapport fini, comme ici où la multiplication par 6 est un rapport fini qui ne touche en rien à l'infini »]. Et de plus, quelle raison avons-nous de juger si un infini peut être plus grand que l'autre, ou non ? Vu qu'il cesserait d'être infini, si nous le pouvions comprendre⁵⁹.

On ne peut identifier un nombre infini comme absolu avec un nombre « infini » qui serait relatif, puisque six fois plus grand qu'un autre. De là, on tirera deux remarques.

D'une part, cet argument logique trouve un fondement dans la méthode cartésienne des moyennes proportionnelles décrite par la Règle VI des *Regulae* : « Je remarque en effet, pour commencer, qu'il n'a pas été plus difficile de trouver le double de 6 que le double de 3 ; et que pareillement, en toutes choses, une fois trouvée la proportion qui existe entre deux grandeurs quelconques, on peut trouver d'autres grandeurs en nombre indéfini, qui ont entre elles la même proportion⁶⁰. » En découvrant les moyennes proportionnelles, Descartes fait du nombre une grandeur numérique toujours relative, en tant qu'elle est toujours à la fois le produit d'un nombre et la moyenne proportionnelle d'un autre. De cette manière, le nombre ne peut être « infini » et incompréhensible, puisqu'il est relatif et exprimable⁶¹.

Et, d'autre part, Yvon Belaval exploite cet argument pour mettre en parallèle, chez Descartes, le couple infini/indéfini et celui des ensembles infinis dénombrables/ensembles transfinis :

Descartes semble même pressentir la différence que les mathématiciens ont, depuis, institué entre ensembles infinis dénombrables et ensembles transfinis : des premiers, il donne un exemple lorsqu'il affirme ne rien voir d'absurde à multiplier par 6 un nombre infini de toises, ni même dans la multiplication infinie de l'indéfini qui est égale à l'infini ; et avec ce que l'entendement divin conçoit au-delà du plus grand nombre (*majus numero*), ne pénétrons-nous pas dans le transfini ?

Et, comme pour répondre à Hannequin, et contester qu'il s'agit là effectivement de transfini, Belaval poursuit : « Mais il aurait fallu que Descartes passât du nombre cardinal à la considération de l'ordinal, pour concevoir le transfini⁶² » — ce qui d'ailleurs est inexact.

Deuxièmement, la non-absurdité de l'existence d'un « nombre infini » n'implique pas l'absurdité de son inexistence. Car toute l'originalité de l'*indéfini* cartésien — qui se révélera effectivement occuper le problème du « plus grand nombre » — réside précisément en cette possibilité de contourner la logique

59. À Mersenne le 15 avril 1630 ; AT I, 146-147.

60. *Regulae* VI ; AT X, 385.

61. Sur le problème des moyennes proportionnelles, on peut voir particulièrement l'analyse de Costabel, 1982.

62. Belaval, 1960, p. 226.

binaire qui fait du fini et de l'infini des contradictoires dont la négation de l'un, en l'occurrence du second, impliquerait l'affirmation de l'autre, en l'occurrence du premier. Descartes, de la négation de l'infini du nombre, ne déduit pas sa finitude, puisqu'il ajoute qu'il ne refuse pas moins l'*inexistence* que l'*existence* à ce prétendu « nombre infini » : face à l'idée d'un « nombre infini », Descartes, d'une part, comme on le sait, refuse donc d'en accepter l'existence, puisqu'un tel nombre relève de l'entendement fini et, d'autre part, refuse effectivement avec tout autant de force de la nier en le déclarant contradictoire, ce qui était pourtant relativement courant à l'époque, chez les scolastiques, Leibniz lui-même, ou Mersenne⁶³, à qui Descartes répond ceci : « De cela seul que j'aperçois que je ne puis jamais, en nombrant, arriver au plus grand de tous les nombres, et que de là je connais qu'il y a quelque chose en matière de nombrer qui surpasse mes forces, je puis conclure nécessairement, *non pas à la vérité qu'un nombre infini existe, ni aussi que son existence implique contradiction*, comme vous dites⁶⁴. » Si donc j'ai le pouvoir de nombrer à l'infini — et cette infinie puissance de nombrer n'est autre qu'une manifestation de l'infini de ma volonté — je n'ai jamais pour autant nommé le plus grand nombre. Par conséquent, je ne peux rien déduire de son existence. L'*indécidabilité* de l'existence du plus grand nombre — puisque ni la proposition de l'existence d'un « nombre infini » ni la proposition contraire ne sont démontrables — implique, entre autres, les deux remarques suivantes.

D'abord, le fait que je ne puisse déduire ni l'existence ni la contradiction du « nombre infini » implique que le principe du tiers exclu devienne ici inapplicable. Pour cette raison, Yvon Belaval parle de l'« intuitionnisme » mathématique de Descartes, qu'il dit comparable aujourd'hui à celui de Brouwer, Weyl, Lebesgue ou Borel, dans la mesure où, comme chez eux, le principe du tiers exclu cesse ici d'être valable dans un ensemble infini⁶⁵. Remarquons par ailleurs que Descartes n'accorde de toute façon que peu d'importance à ce principe :

Au premier sens, on peut dire que *impossible est idem simul esse et non esse* est un principe, et qu'il peut connaître l'existence d'une chose, mais seulement à faire que, lorsqu'on la connaît, on en confirme la vérité par un tel raisonnement. Il est impossible que ce qui est ne soit pas ; or je connais que telle chose est ; donc je connais qu'il est impossible qu'elle ne soit pas. Ce qui est de bien peu d'importance, et ne nous rend de rien plus savants⁶⁶.

Et si le principe du tiers exclu est « de bien peu d'importance », c'est que Dieu est *au-dessus* du principe de non-contradiction qui le conditionne : « Dieu ne peut avoir été déterminé à faire qu'il fût vrai, que les contradictoires ne peuvent être ensemble, et que, par conséquent, il a pu faire le contraire⁶⁷. »

63. Les *Secondes Objections*, AT VII, 124, disaient : « *est numerus ille impossibilis* ».

64. *Réponses aux secondes objections* ; AT IX-1, 109-110, spn.

65. Belaval, 1960, p. 221. Ce qui est inexact dans le cas de Borel, qui acceptait l'infini dénombrable (.

66. À Clerselier de juin ou juillet 1646 ; AT IV, 444.

67. À Mesland du 2 mai 1644 ; AT IV, 118.

Ensuite, on notera que Leibniz, dans ses *Animadversiones*⁶⁸, profitera de cette indétermination pour dénoncer une soi-disant contradiction : Descartes, d'une part, refuse de nier ou d'affirmer l'existence du dernier terme dans une série divergente (à juste titre, par définition, car une série divergente n'a pas de limite)⁶⁹ et, d'autre part, affirme la divisibilité de la matière à l'infini ; de telle sorte qu'il y a ensemble un nombre infini *indécidable* et un nombre infini de parties *décidable*. On pourrait répondre à Leibniz en utilisant les textes de Descartes lui-même, et contester qu'il y a effectivement contradiction en distinguant *divisibilité* et *division* : « Pensez que chaque corps peut être divisé en des parties extrêmement petites. Je ne veux point déterminer si leur nombre est infini ou non ; mais du moins il est certain qu'à l'égard de notre connaissance il est indéfini, et que nous pouvons supposer qu'il y en a plusieurs millions dans le moindre petit grain de sable qui puisse être aperçu de nos yeux⁷⁰ » ; ou encore : « De ce que je ne puis comprendre [*comprehendere*] une infinité de divisions en une quantité finie, il ne s'ensuit pas que l'on puisse venir à une dernière, après laquelle cette quantité ne puisse plus être divisée ; mais bien il suit seulement que mon entendement, qui est fini, ne peut comprendre [*capere*] l'infini⁷¹. » Dans ces textes est manifeste la distinction entre la possibilité d'être divisé (divisibilité) et la division effective. La contradiction soulevée par Leibniz n'a plus lieu d'être si l'on répartit l'indécidabilité du nombre infini dans la *division* et la décidabilité d'un nombre infini de parties dans la seule *divisibilité*, qui est aussi peu concurrente de la première que la puissance de l'acte.

C'est ainsi que l'indécidabilité de l'existence du « nombre infini » lui donne un statut d'indéfini — ce dont on ne peut prouver les bornes — puisqu'on ne peut le prouver infini : « Et il suppose aussi qu'il ne peut y avoir de nombre infini, etc., ce qu'il ne saurait *prouver* non plus⁷². » Pour cette raison, et de la même manière que « parce que nous ne saurions imaginer tant d'étoiles que Dieu n'en puisse créer davantage, nous supposerons que leur nombre est indéfini, et ainsi du reste⁷³ », le « plus grand nombre » reste indéfini. *L'Entretien* confirmera cette conclusion — tout en manifestant, une fois de plus, la dépendance des mathématiques à l'égard de la physique — en ajoutant, comme une preuve supplémentaire : « Ce nombre infini ne se rencontre-t-il pas dans la division de la quantité, etc. ?⁷⁴ », alors que la divisibilité de la matière en question est *indéfinie* et non infinie.

68. *Animadversiones* de Paris, art. 26-27.

69. Le passage du *nombre* infini à la *suite* — ou *série* — infinie se fait naturellement, puisque la question de l'infini, en mathématiques, peut elle-même être réduite au fait que, de chaque nombre, et aussi grand qu'il soit, on pourra toujours trouver un nombre plus grand encore, ne serait-ce qu'en lui ajoutant 1, et ainsi indéfiniment.

70. *Le Monde* XI ; AT XI, 12.

71. *Réponses aux Premières Objections* ; AT IX-1, 85.

72. À Mersenne le 28 janvier 1641 ; AT III, 294, spn.

73. *Principes* I, 26 ; AT IX-2, 36.

74. *Entretien avec Burman* ; AT V, 155 ; Descartes, 1981, p. 52.

Conclusion : les raisons de la prudence cartésienne

La question de l'infinitésimal exprime de manière paradigmatique la prudence de Descartes face au problème de l'infini en mathématiques. Conclure sur ce point implique maintenant d'explorer les raisons supposées de l'esquive cartésienne, afin de comprendre pourquoi Descartes s'est tant évertué à contourner soigneusement toute manifestation de l'infini en mathématiques : quel est donc le moteur de la prudence cartésienne ?

Certains voudraient y voir des raisons internes aux mathématiques ou à la physique, dans le seul but de trouver chez Descartes une autorité de plus à la légitimation de l'indépendance du scientifique par rapport au philosophique, voire même du logique par rapport au métaphysique. Mais dans le cadre de l'histoire de la philosophie, laquelle commande de ne pas interpréter l'auteur au-delà de ce qu'on y peut lire, et particulièrement dans le cas de Descartes, *il est absolument incontestable que l'indéfini mathématique présuppose l'indéfini physique, lequel présuppose à son tour l'infini divin*, et que, de cette manière, on ne peut raisonnablement y comprendre une quelconque indépendance qu'en la posant soi-même de sa lecture. Nous limitant, pour notre part, aux textes cartésiens, il est aisé de montrer, de trois manières, comment la fuite de Descartes face à la question de l'infini en mathématiques est liée à des impératifs métaphysiques, comme tous les grands commentateurs de Descartes le reconnaissent sans peine.

D'abord, on peut réaliser la dépendance de l'indéfini mathématique envers l'infini divin, par le seul choix du vocabulaire qui nous contraint à ne parler de l'infini mathématique qu'en terme d'indéfini, parce que Dieu seul, qui est passé avant, pouvait ravir l'infini. L'indéfini étant ainsi nommé *par défaut*, pour désigner tout ce qui nous *semble* infini sans être Dieu, donc sans l'être positivement, l'indéfini mathématique souffre d'une dette envers l'infini divin au moins *linguistique*.

Ensuite, lorsqu'Yvon Belaval écrit : « On remarque quelles idées se retrouvent toujours dans l'esprit de Descartes lorsqu'il aborde le problème du plus grand nombre : l'idée de Dieu, celle d'une puissance de nombrer, celle du nombre lui-même et celle, enfin, de l'existence de ce nombre⁷⁵ », il faut comprendre l'ordre de l'énumération comme n'étant pas hasardeux. L'idée de Dieu est en effet d'abord là, toujours antécédente, comme l'exige l'ordre des raisons dans les *Méditations* : c'est uniquement parce que j'ai en moi l'idée de Dieu et que de cette manière je peux démontrer son existence, que je pourrai faire de Dieu un principe épistémologique dont la véracité permettra, dans une dialectique descendante (application), de refonder la possibilité même de la science, dont le nombre fait partie. C'est ainsi que « cette puissance de nombrer confirme la preuve de Dieu. C'est par elle que je puis produire les nombres, mais cette cause immédiate du nombre est elle-même l'effet d'une cause supérieure, car je ne suis pas cause de moi, ni cause de ma conservation⁷⁶ ». De

75. Belaval, 1960, p. 222.

76. Belaval, 1960, p. 223-224. Voir les *Réponses aux Seconde Objections*, AT VII, 139.

ce point de vue, parce que l'existence de Dieu est condition de mon idée du nombre, l'infini divin est condition de l'indéfini mathématique.

Enfin, cela est aussi aisément démontrable par l'indéfini physique, en utilisant la transitivité du syllogisme suivant : On sait que l'indéfini physique dépend de l'infini divin, or l'indéfini mathématique dépend de l'indéfini physique, donc l'indéfini mathématique dépend de l'infini divin.

Que l'indéfini physique dépende de l'infini divin, nous l'avons montré ailleurs, et notamment à travers l'utilisation de Dieu dans les preuves de l'indéfini de l'univers⁷⁷. C'est pourquoi Françoise Monnoyeur n'hésite pas à déclarer : « La considération sur l'infinité du monde est envisagée par rapport à l'infinité divine et se trouve ainsi en étroite dépendance de celle-ci⁷⁸. » En accord avec Michel Fichant⁷⁹, toute interprétation positiviste de la physique cartésienne, qui à vrai dire ne se fait plus guère depuis Louis Liard, qui la caractériserait par « l'absence de toute idée métaphysique »⁸⁰, doit être abandonnée, si l'on veut être fidèle aux textes cartésiens. Le métaphysique, dans l'œuvre de Descartes, ne peut en être détachée sans blessures⁸¹.

Que l'indéfini mathématique dépende ensuite de l'indéfini physique, c'est la définition même du célèbre géométrisme cartésien, qui fait poursuivre Belaval : « Chez Descartes, tout l'infini quantitatif, dont traitent les mathématiques, est réductible à l'étendue géométrique⁸². » On peut à cet égard remarquer que si Descartes a adopté la *recta linea* comme type de la quantité, c'est précisément pour ne pas connaître le nombre irrationnel en tant que nombre. Celui-ci, selon la définition de Cauchy, étant la limite des diverses fractions qui en fournissent des valeurs de plus en plus approchées, on comprend que Descartes n'ait pu, contrairement à Leibniz, participer à sa génération, d'une part pour le problème que lui pose la notion de limite, laquelle inclut le nombre irrationnel à la question de l'infini mathématique qu'il ne veut point résoudre, et d'autre part parce qu'il aurait fallu, pour cela, substituer au quantitatif cardinal le qualitatif ordinal. Et cette même étendue géométrique présuppose la matière et le mouvement de la physique : « Chez Descartes, cet infini — ou cet indéfini — mathématique, Dieu l'explique seulement en tant que cause : cet infini nous est donné dans l'essence de la matière et du mouvement, encore « plus aisé à concevoir que la ligne des géomètres »⁸³, parce qu'il engendre les lignes⁸⁴. » On comprend ainsi que « Descartes ne posait le problème du plus grand nombre que par rapport au monde⁸⁵ ».

77. Voir notre article « Descartes : indéfini de et dans l'univers », à paraître.

78. Monnoyeur, 1992, p. 85.

79. Fichant, 1998, p. 65.

80. Liard, 1882, p. 66.

81. Voir notamment Gilson, 1930, p. 176.

82. Belaval, 1960, p. 223-224.

83. *Le Monde*, VII ; AT XI, 40.

84. Belaval, 1960, p. 223-224.

85. Belaval, 1960, p. 272.

On peut donc légitimement conclure que, par extension, l'infini mathématique dépend de l'infini divin : « l'infini mathématique présuppose un infini de nature spirituelle⁸⁶ ». Vuillemin confirme : « Descartes peut donc bien être regardé comme l'un des fondateurs du Calcul intégral, techniquement parlant. Mais *l'interdit métaphysique* qu'il faisait peser sur la classe des transcendantes lui interdisait de donner droit de cité à sa découverte⁸⁷. »

Ainsi, Descartes a développé sa pensée scientifique dans un espace soumis à ne pas pouvoir dépasser la voûte métaphysique ; et c'est précisément par cette limitation qu'en partie « s'explique la fuite de Descartes devant toute spéculation mathématique sur l'infini, bien que son génie eût commencé à en explorer les voies⁸⁸ ». Cela ne suffit pas, pourtant. La prudence épistémologique de Descartes face à la question de l'infini, partagée d'ailleurs par nombre de penseurs de Copernic à Kant, est aussi et surtout la manifestation d'un esprit attaché à ne pas transgresser l'ordre des raisons⁸⁹. Le moteur de sa prudence reste une contrainte, mais elle est double : voûte métaphysique et rigueur méthodologique.

Bibliographie

Descartes

Nous citons Descartes dans l'édition de référence d'Adam et Tannery en 12 volumes, Paris, Vrin, 1897-1913, dans sa nouvelle présentation par P. Costabel et B. Rochot, Paris, Vrin-CNRS, 1966. L'Entretien avec Burman est publié dans une édition distincte : Descartes, *Entretien avec Burman*, édition de Jean-Marie Beyssade, Paris, PUF, 1981.

Littérature secondaire

- Alquié, Ferdinand. *La découverte métaphysique de l'homme chez Descartes*, Paris, PUF, 1950.
- . Édition et notes de Descartes, *Œuvres Philosophiques*, t. II, Paris, Bordas, 1992.
- Banfi, Antonio. « Galilée, Descartes et Vico », in *Descartes*, Cahiers de Royaumont, Philosophie n. II, Paris, Editions de Minuit, 1957, p. 376-392.
- Belaval, Yvon. « Descartes selon l'ordre des raisons, d'après M. Gueroult », *Revue de métaphysique et de morale*, 1955, p. 442-443.
- . *Leibniz critique de Descartes*, Paris, Gallimard, 1960.
- Berkeley, George. *Œuvres*, vol. I, édition de G. Brykman, Paris, PUF, 1985.

86. Belaval, 1960, p. 224.

87. Vuillemin, 1960, p. 73, souligné par nous (spn). Plus généralement, « Leibniz d'une part, Newton et Kant de l'autre fonderont leurs systèmes sur les ruines du cartésianisme en accordant la première place, celui-là à l'idée de virtualité, ceux-ci aux constructions mécaniques, que Descartes, mathématicien, avait naturellement rencontrées pour résoudre des problèmes techniques, mais que la cohérence de sa méthode métaphysique l'avait contraint de ne traiter qu'aux marges des découvertes qu'il jugeait essentielles à sa pensée. »

88. Gueroult, 1968, p. 521.

89. Sur Copernic notamment, voir Seidengart, 2006. Je remercie l'évaluateur anonyme de *Philosophiques* d'avoir attiré mon attention sur ce point.

- Brunschvicg, Léon. « La révolution cartésienne et la notion spinoziste de la Substance », *Revue de métaphysique et de morale*, 1904, p. 755-798.
- . « Mathématique et métaphysique chez Descartes », *Revue de métaphysique et de morale*, 1927, p. 277-324.
- Cantor, Georg. *Gesammelte Abhandlungen Mathematischen und Philosophischen Inhalts*, Berlin, J. Springer, 1932.
- Carnot, Lazare. *Réflexions sur la métaphysique du calcul infinitésimal*, 2 vol., Paris, Duprat, 1797.
- Carr, Thomas A. *Descartes and the Resilience of Rhetoric*, Carbondale, Southern Illinois University Press, 1990.
- Cléro, Jean-Pierre. « La méthode des fluxions de Newton », dans *Divers aspects de l'infini en mathématiques et en philosophie*, Cahiers pédagogiques de philosophie et d'histoire des mathématiques, ministère de l'Éducation nationale, académie de Rouen, 1989-90, fasc. 2, p. 169-190.
- Costabel, Pierre. « La solution par Descartes du problème des moyennes proportionnelles (Règle VI) », dans *Démarches originales de Descartes savant*, Paris, Vrin, 1982, p. 49-62.
- Couturat, Louis. *De l'infini mathématique*, Paris, Alcan, 1896.
- Enriques, Federico. « Descartes et Galilée », *Revue de métaphysique et de morale*, 1937, p. 221-235.
- Fénélon, François de Salignac de la Motte. *Démonstration de l'existence de Dieu, tirée de la connaissance de la Nature, et proportionnée à la faible intelligence des plus simples*, Paris, Jacques Estienne, 1712.
- Fichant, Michel. *Science et métaphysique dans Descartes et Leibniz*, Paris, PUF, 1998.
- Galilée, Galilei Galileo. *Discours et démonstrations mathématiques concernant deux sciences nouvelles*, traduit par Maurice Clavelin, Paris, A. Colin, 1970.
- Gillispie, Charles Coulston, et Adolf Pavlovich Iushkevich. *Lazare Carnot savant et sa contribution à la théorie de l'infini en mathématiques*, Paris, Vrin, 1979.
- Gilson, Étienne. *Études sur l'histoire de la formation du système cartésien*, Paris, Vrin, 1930.
- Gueroult, Martial. « La lettre de Spinoza sur l'infini », *Revue de métaphysique et de morale*, 1966.
- ¹/_M³. *Spinoza : Dieu, Ethique I*, Paris, Aubier-Montaigne, 1968.
- Houzel, Christian. « Descartes et les courbes transcendantes », dans J. Biard & R. Rashed, dir., *Descartes et le Moyen Âge*, Paris, Vrin, 1997.
- Gandt, François de. « Newton : la justification des infiniment petits et l'intuition du mouvement », dans F. Monnoyeur, dir., *Infini des mathématiciens, infini des philosophes*, Paris, Berlin, 1992, p. 147-157.
- Hannequin, Arthur. « La preuve ontologique cartésienne défendue contre la critique de Leibniz », *Revue de métaphysique et de morale*, 1896, p. 433-458.
- Leibniz, Gerhardt Wilhelm. *Die Philosophischen Schriften von Gottfried Wilhelm Leibniz*, édition de C. D. Gerhardt en 7 vol., Berlin, 1875, réed. Georg Olms Hildesheim, 1960.
- . *Nouveaux Essais sur l'entendement humain*, édition de Jacques Brunschwig, Paris, GF, Flammarion, 1990.
- Liard, Louis. *Descartes*, Paris, G. Baillière et C^{ie}, 1882.
- Mahoney, Michael Sean. *The Mathematical Career of Pierre de Fermat*, Princeton, Princeton University Press, 1973.

- Marion, Jean-Luc. *Sur la théologie blanche de Descartes*, Paris, PUF, 1981.
- Monnoyeur, Françoise. « L'infini et l'indéfini dans la théorie cartésienne de la connaissance », dans F. Monnoyeur, dir., *Infini des mathématiciens, infini des philosophes*, Paris, Berlin, 1992.
- North, John David. « Finite and Otherwise : Aristotle and some Seventeenth Century Views », dans William R. Shea, dir., *Nature Mathematized*, Dordrecht, Reidel, 1983.
- Pascal, *Œuvres de Pascal*, édition de Léon Brunschvicg, Pierre Boutroux et Félix Gaze, 11 vol., Paris, Hachette, 1908-1914.
- Rodis-Lewis, Geneviève. *L'individualité selon Descartes*, Paris, Vrin, 1950.
- Russo, François. « Pascal et l'analyse infinitésimale », *Revue d'histoire des sciences*, vol. 15, 1962, p. 303-320.
- Seidengart, Jean. *Dieu, l'univers et la sphère infinie : penser l'infinité cosmique à l'aube de la science classique*, Paris, Albin Michel.
- Sergescu, Petre. « L'infiniment petit mathématique du Moyen Âge au XIX^e siècle », dans *L'infini et le réel*, Centre international de synthèse, Dix-huitième semaine de synthèse, Exposés et discussions, Paris, Albin Michel, 1955.
- Shea, William R. « Descartes as critic of Galileo », dans R. E. Butts & J. C. Pitts, dir., *New Perspectives on Galileo*, Dordrecht, Reidel, 1978, p. 139-159.
- Spinoza. *Œuvres Complètes*, édition de Roland Caillois, Madeleine Francès et Robert Misrahi, Paris, Gallimard, 1954.
- Vuillemin, Jules. *Mathématiques et métaphysique chez Descartes*, Paris, PUF, 1960.

Le rêve matérialiste, ou « Faire par la pensée ce que la matière fait parfois »

Charles T. Wolfe

Volume 34, numéro 2, automne 2007

URI : <https://id.erudit.org/iderudit/016993ar>

DOI : <https://doi.org/10.7202/016993ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Société de philosophie du Québec

ISSN

0316-2923 (imprimé)

1492-1391 (numérique)

[Découvrir la revue](#)

Citer cet article

Wolfe, C. T. (2007). Le rêve matérialiste, ou « Faire par la pensée ce que la matière fait parfois ». *Philosophiques*, 34 (2), 317–328.
<https://doi.org/10.7202/016993ar>

Résumé de l'article

Cet article vise à expliciter la notion de « rêve matérialiste » à partir d'une réflexion sur l'ouvrage de Diderot, *Le Rêve de D'Alembert*. Quel lien y a-t-il entre le matérialisme philosophique proclamé dans ce livre et la forme du rêve qui donne un caractère inédit à la présentation de cette philosophie? Une approche purement textuelle montrerait, déjà, une indissociabilité particulière entre forme et contenu; mais l'approche proposée ici s'attache à la manière dont une certaine idée du rêve sert de condition nécessaire *conceptuelle* à l'élaboration du matérialisme de Diderot : à l'affirmation de l'unicité du monde matériel sans perdre de vue l'état mental d'un sujet rêvant.

Le rêve matérialiste, ou « Faire par la pensée ce que la matière fait parfois »*

CHARLES T. WOLFE

Unit for History and Philosophy of Science, University of Sydney
c.wolfe@usyd.edu.au

RÉSUMÉ. — Cet article vise à expliciter la notion de « rêve matérialiste » à partir d'une réflexion sur l'ouvrage de Diderot, *Le Rêve de D'Alembert*. Quel lien y a-t-il entre le matérialisme philosophique proclamé dans ce livre et la forme du rêve qui donne un caractère inédit à la présentation de cette philosophie ? Une approche purement textuelle montrerait, déjà, une indissociabilité particulière entre forme et contenu ; mais l'approche proposée ici s'attache à la manière dont une certaine idée du rêve sert de condition nécessaire *conceptuelle* à l'élaboration du matérialisme de Diderot : à l'affirmation de l'unicité du monde matériel sans perdre de vue l'état mental d'un sujet rêvant.

ABSTRACT. — This article seeks to explicate the notion of what I term the 'materialist dream', on the basis of an examination of Diderot's work *D'Alembert's Dream*. What might the connection be, between the materialist philosophy put forth as the content of the work, and the form of the dream which gives a highly idiosyncratic character to the presentation of this philosophy ? A purely textual, literary analysis would itself show a particular kind of inseparability between form and content here ; but the analysis suggested here focuses on how a certain idea of the dream serves as a necessary *conceptual* condition for the elaboration of Diderot's materialism : for the affirmation of the unity of the material world without losing sight of the mental states of a dreaming subject.

C'est un état bien singulier que celui du rêve. Aucun philosophe que je connaisse n'a encore assigné la vraie différence de la veille et du rêve¹.

Quand la philosophie se trouve dans une impasse dont la science ne peut pas la sortir, face à une limite ou un « dehors », il lui reste une botte secrète, une sortie cachée : le rêve. Pourtant, elle doit également se garder de ne pas fonder sa réflexion sur une rêverie, car elle tomberait alors dans la sphère du privé même et n'aurait aucune force de conviction sur laquelle s'appuyer quand elle chercherait à convaincre un interlocuteur. Certes, depuis Locke, nous connaissons

* Cette réflexion sur le lien entre matérialisme et rêve au sein du *Rêve de D'Alembert* de Diderot doit beaucoup aux analyses de Alain Gigandet sur Diderot et Lucrèce, et de Jean-Claude Bourdin sur la forme du *Rêve*, ainsi qu'aux travaux du groupe de recherches auquel j'ai appartenu sur le *Rêve* au sein du CERPHI (ENS LSH, 2000-2005), dirigé par Jean-Claude Bourdin, Colas Duflo, et Annie Ibrahim, travaux qui ont abouti à la publication de *l'Encyclopédie du Rêve de D'Alembert*, Paris, Éditions du CNRS, 2006.

1. Diderot, « Promenade Vernet », *Salon de 1767, Œuvres complètes*, d. Dieckmann-Proust-Varloot, Paris, Hermann, 1975- (cité DPV), vol. XVI, p. 230.

le rôle des expériences de pensée dans la construction d'un argument philosophique, mais celles-ci servent principalement une fin déflationniste de réduire l'argument d'autrui à des bases absurdes. On imagine mal un Socrate rêvant : « Sais-tu, mon cher Ménon, qu'hier j'ai rêvé d'un esclave capable de résoudre les plus vieux problèmes de la géométrie ? ».

Un projet systématique se porterait mal s'il était fondé sur un rêve ; mais une philosophie spéculative et éclectique — au sens pluraliste et non fondationnaliste, donc asystématique donné à ce mot dans l'article « ÉCLECTISME » de l'*Encyclopédie* de Diderot et D'Alembert² — qui souhaitera pousser plus loin la réflexion sans rattacher chaque idée neuve à une chaîne de vérités préexistante, pourra les exprimer par un rêve. Mais justement, ce que j'appellerai ici le rêve proprement matérialiste n'est pas un pur artifice littéraire, un emballage pour un contenu théorique totalement distinct de sa forme. Dans ce sens, un « rêve » possède une dimension corporelle vécue que n'a pas un « songe », ce que l'histoire de ces expressions semble confirmer : pendant longtemps, « songer » est associé à « penser », alors que « rêver » est associé à « délirer »³. Ainsi, la notion que cet article vise à expliciter, celle de « rêve matérialiste », est tout le contraire des « songes » ou « rêveries » philosophiques ; ceux-ci sont soit des artifices visant à présenter un ensemble d'idées déterminées — les mythes dans la *République* de Platon, ou justement le *Songe de Platon* de Voltaire — soit des prolongations de l'espace conceptuel du moi, chez Descartes ou Rousseau.

Je pense que c'est dans ce sens qu'il faut interpréter la phrase assez mystérieuse de l'article « RÊVE » qu'on considère comme étant pour l'essentiel de Diderot (on le trouve dans l'édition de référence de ses *Œuvres complètes*) : « L'histoire des rêves est encore assez peu connue, elle est cependant importante, non seulement en médecine, mais en métaphysique, à cause des objections des idéalistes » ; que signifie cet « à cause des objections des idéalistes » ? Lisons la suite :

Nous avons en rêvant un sentiment interne de nous-mêmes, & en même-temps un assez grand délire pour voir plusieurs choses hors de nous ; nous agissons nous-mêmes voulant ou ne voulant pas ; & enfin tous les objets des rêves sont visiblement des jeux de l'imagination⁴.

Autrement dit, la métaphysique matérialiste a besoin du concept de rêve pour pouvoir y assimiler le « sens interne » ou « l'intentionnalité » des idéalistes, en les naturalisant, car le rêve et surtout ses objets sembleraient être des objets

2. *Encyclopédie de Diderot et d'Alembert*, 35 vol., Paris, Briasson, 1751-1780 ; réédition, Stuttgart/Bad Cannstatt, Frommann, 1966, V, 273a (dorénavant cité *Enc.*, suivi du volume et de la page).

3. L'expression « songe » est plus ancienne, datant au moins du XII^e siècle. Voir V. C. Petrovich, *Connaissance et Rêve(rie) dans le discours des Lumières*, New York, Peter Lang, 1996, p. 6.

4. « REVE, s. m. (*Métaphysique*) : songe qu'on fait en dormant. Voyez SONGE » (*Enc.*, XIV, 228).

intentionnels par excellence — mais il est également un phénomène ou processus dont on peut rendre compte en termes pleinement naturalistes. L'article, très bref, se conclut avec une définition tout à fait sensualiste, qu'on retrouvera dans l'article bien plus long du médecin suisse Samuel Formey intitulé « SONGE » (auquel je reviendrai plus loin⁵) : « Les choses qui nous ont le plus frappé durant le jour apparaissent à notre âme lorsqu'elle est en repos ; [...] la cause des rêves est donc toute impression quelconque, forte, fréquente & dominante » (« RÊVE », *op. cit.*).

Tout métaphysique qu'il est, un rêve est plus dangereux qu'un songe (seul le premier est considéré par l'*Encyclopédie* comme appartenant à la métaphysique) : plus matériel, plus charnel, plus terrien en ce qu'il est plus proche du délire et donc d'une causalité psychique réelle. Ce rêve inventé est alors un rêve philosophique, qui partage avec le rêve vécu la fonction de rendre pensable l'impensable, ou, pour le dire autrement, de donner une expression, une matérialité, à des idées qui seraient restées entièrement théoriques et abstraites si elles avaient été exprimées dans un « traité », un « précis », une « ontologie »...

* * *

Dans la pensée matérialiste, il existe quelques petits rêves, tel celui de Cyrano de Bergerac, *L'Autre Monde ou États et empires du Soleil et de la Lune*⁶, qui s'apparente plutôt à ce que je nommais « songe », et un grand rêve : le dialogue en trois parties de Diderot nommé *Rêve de D'Alembert*⁷, rédigé en 1769 mais jamais publié de son vivant — ce qui montre, en passant, qu'il n'emploie pas la forme polyphonique du rêve pour « dissimuler » des idées, comme on l'a parfois dit⁸. L'idée générale de ce texte qui ne peut pas être exprimée tant qu'elle n'a pas pris la « consistance » du rêve est celle de la *sensibilité universelle de la matière*, c'est-à-dire d'une *inhérence* dans la matière de la sensibilité. Le caractère « vital » ou même « vitaliste » du matérialisme de Diderot, pour reprendre une expression employée d'abord par Jean Wahl, tient dans son attribution de propriétés organiques telle la sensibilité (dont la définition est reprise principalement aux avancées physiologiques de Thomas Willis et surtout Albrecht von Haller) à *la matière elle-même*⁹.

5. Formey est également l'auteur d'un *Essai sur le sommeil* que l'on trouve dans le tome II de ses *Mélanges philosophiques*, Leyde, chez E. Luzac, 1754.

6. Cyrano de Bergerac, *Estats et Empires de la Lune* [approx. 1648] et *Histoire des Estats et Empires du Soleil* [1655], in *Œuvres complètes*, vol. 1, M. Alcover, dir., Paris, Champion, 2000.

7. Je cite le *Rêve de D'Alembert* dans l'édition Versini (*Œuvres*, tome 1 : *Philosophie*, Paris, Laffont, coll. « Bouquins », 1994), indiquée 'V' suivi du numéro de la page.

8. Pour l'analyse textuelle suivie du *Rêve de D'Alembert*, qui appuie, je l'espère, l'analyse conceptuelle proposée ici, je renvoie aux études de H. Dieckmann, A. Vartanian et G. May dans *Diderot Studies* n° 17 (1973), un numéro spécial sur le *Rêve*, ainsi qu'à J. Chouillet, « La poétique du rêve dans les *Salons* », *Stanford French Review* vol. 8, n° 2-3 (1983).

9. J. Wahl, *Tableau de la philosophie française*, Paris, Gallimard, 1962, p. 54. Sur les transformations de la notion hallérienne de sensibilité chez Diderot, voir F. Duchesneau, « Diderot et la physiologie de la sensibilité », *Dix-huitième siècle*, vol. 31 (1999).

Cette attribution est présentée parfois comme une définition (l'une des propriétés de la matière serait de sentir), parfois — notamment ici dans le *Rêve* — comme une expérience de pensée, qui joue dans ce texte un rôle fondateur. Répondant à l'incrédulité du personnage D'Alembert, le personnage Diderot propose le célèbre *Gedankenexperiment* de la statue de marbre qui « s'animalise » petit à petit, devenant terre, puis chair : « par une suite d'opérations purement mécaniques, je réduirais le premier génie de la terre à une masse de chair inorganisée » (V 614). Cependant, D'Alembert éveillé — car il n'est précisément *pas encore dans un état de rêve* — n'en accepte pas pour autant que « la pierre sent ». La tactique employée par Diderot, à travers le personnage du dialogue qui porte son nom, est de revenir en arrière vers un état de la conscience plus primitif, non pas parce qu'il est plus simple, brut ou malléable, mais à cause de son « animisme résiduel » : il n'y pas de distinction tranchée entre le vivant et le non-vivant dans l'univers du rêve.

Cette insistance chez Diderot sur la plénitude du rêve eu égard aux limitations spatio-temporelles du monde éveillé, sa dimension onirique, vient remplacer sa fonction prémonitoire traditionnelle¹⁰ ; celle-ci disparaît au XVIII^e siècle, comme on le voit avec les cas de la comète ou du monstre, qui de « merveille » ou « prodige » deviendra une simple « anomalie », là encore suivant un processus de naturalisation¹¹. Cette plénitude, cette matérialité différencie le rêve du songe ou de la supposition théorique pure, qui n'est qu'une abstraction, celle-ci étant jugée par Diderot, ici à travers le personnage du médecin Bordeu, comme « un signe vide d'idée » (V 667). De plus, on peut considérer que l'état de rêve, dans lequel les fonctions polarisantes de l'état de veille (telles que la concentration, la focalisation, etc.) sont éteintes, est le microcosme qui correspond, au niveau macrocosmique, à cette sensibilité universelle de la matière. Le sujet rêvant est (comme) la matière sensible, dans un état de réceptivité pure ; la rêverie est présentée ici comme un abandon des frontières du moi (« je suis un peloton de points sensibles [...] tout presse sur moi et je presse sur tout », V 638).

Le rêve, c'est donc du concret. Chaque fois que Diderot en invente un (d'abord ceux de Mangogul et Mirzoza dans *Les Bijoux indiscrets*, vingt ans avant le *Rêve de D'Alembert*), il le fait afin de construire une situation dans laquelle le « temple des systématiciens », au « royaume des hypothèses », sera détruit, ici par le jeune géant Expérience. Cependant, cette destruction a bien lieu *dans* le rêve, mais elle produit sa fin. Ce premier rêve de Diderot n'est pas un opérateur du « plus » de réalité, à la différence du fonctionnement « fictif » du rêve proprement matérialiste, qui permet de libérer un personnage des contraintes (physiques, perceptives) du monde réel, du monde dans lequel

10. Diderot écarte la fonction prémonitoire du rêve dans *Les bijoux indiscrets*, chap. « Les songes », DPV III, p. 182.

11. Voir A. Ibrahim, « Le statut des anomalies dans la philosophie de Diderot », *Dix-huitième siècle*, vol. 15 (1983), et Ibrahim, dir., *Qu'est-ce qu'un monstre ?*, Paris, PUF, 2005 ; C.T. Wolfe, dir., *Monsters and Philosophy*, Londres, Kings College Publications, 2005.

nous vivons, afin d'exprimer les « suppositions » de la philosophie spéculative. Car, ne l'oublions pas, le rêve matérialiste doit faire progresser la philosophie, en tant qu'elle est à la recherche d'ordres multiples, de liaisons, de symétries au-delà de l'ordre du visible ; ainsi, face à l'*invisibilité* de l'ordre interne du vivant, elle doit « rêver » les divers concepts-hypothèses proto-scientifiques : les « molécules » ou unités minimales du vivant, les polypes ou êtres « mitoyens » entre les règnes dont la propriété la plus frappante est de se régénérer, les générations spontanées et l'épigenèse de l'embryon qui signifient tous deux aux yeux de Diderot l'indécidabilité et le dynamisme interne de la matière vivante...

Ce rôle expérimental du rêve le rapproche de l'activité de l'« interprète de la nature », cette figure baconienne qui est présentée dans l'ouvrage de Diderot *Pensées sur l'interprétation de la nature* (1753). L'interprète de la nature peut affirmer par « conjectures » la cohérence de l'univers, y compris sa clôture causale, avant ou sans que la science expérimentale soit en mesure de la confirmer. D'ailleurs, Diderot associe lui-même l'activité cognitive nommée « conjecture », c'est-à-dire mettre en rapport spéculatif différents éléments perçus, aux « rêves d'un malade », à la divination. Le flux d'idées dans une conversation à plusieurs voix, leur(s) mécanisme(s) d'association, sont comme ceux des rêves : « Les rêves d'un malade en délire ne sont pas plus hétéroclites [mais cependant] il n'y a rien de décousu ni dans la tête d'un homme qui rêve, ni dans celle d'un fou¹². » Dans le rêve matérialiste, on ne vérifie et on ne falsifie pas des propositions, mais on accumule des faits, toujours plus de faits ! Toutefois ces faits ne sont ni des sense-data ni des éléments constitutants d'une théorie scientifique.

Le rêve n'est pas la science même ; si nous avions accès aux fondements réels de la nature, à « ce qui se passe » en elle, nous n'aurions pas besoin d'analogies, de conjectures, de suppositions : de rêves. Mais inversement, et d'une manière qui rappelle tout autant l'inférentialisme humien (selon lequel « la connexion nécessaire dépend de l'inférence, au lieu que l'inférence dépende de la connexion nécessaire¹³ »), toute théorie de la nature est, selon Diderot, fondée sur une analogie, comme par exemple l'analogie de la machine pour expliquer le fonctionnement du corps (après tout, même Descartes est bien obligé d'admettre, à propos du corps, qu'au mieux « on *peut fort bien comparer* les nerfs de la machine que je vous décris aux tuyaux des machines¹⁴ »).

Cependant, même si l'interprète de la nature n'a pas accès à ses fondements réels et doit donc produire des conjectures dont le statut analogique les rapproche du reste des mécanismes d'association, sa réflexion est naturaliste au sens où (i) elle appartient pleinement à l'univers naturel et (ii) elle se situe

12. Lettre de Diderot à Sophie Volland, 20 octobre 1760, in *Correspondance*, vol. 3, G. Roth, dir., Paris, Éditions de Minuit, 1957, p. 173.

13. D. Hume, *Traité de la nature humaine* (1739), liv. I, ch. iii, § 6.

14. *L'Homme*, AT XI, 130 ; c'est moi qui souligne.

sur un plan de continuité avec l'émergence du discours scientifique, et non pas au-dessus ou en opposition à lui. Dans ce sens, l'interprétation de la nature doit contribuer à la « théorie *une* » sur la « machine » de l'univers « *qui est une* », comme dit Diderot ; « la découverte de cette théorie est d'autant plus éloignée que la machine est plus compliquée¹⁵ ». Or nos intellects ne sont pas capables de dépasser leurs limites psychologiques et spatio-temporelles. « Si nous possédions le recueil complet des phénomènes, il n'y aurait plus qu'une cause ou supposition » (*ibid.*). Mais nous ne possédons pas ce « recueil » ; nous n'avons pas accès aux causes premières, seulement à la multiplicité des effets. Ce que Diderot appelle dans le même texte « suppléer au silence » de la nature, c'est donc nécessairement supposer et produire des conjectures, c'est rêver.

Le « grand tout » de la nature (c'est ainsi que Diderot désigne la totalité de l'univers naturel) comporte des blancs que l'histoire naturelle n'a pas remplis ; elle inventera en conséquence des « cabinets d'histoire naturelle ». Ceux-ci, ainsi que l'explique l'article de l'*Encyclopédie* qui porte ce titre (et qui est de Diderot), sont voués à combler, en tant que créations de l'esprit, les lacunes dans notre connaissance de la nature elle-même : « L'ordre d'un cabinet ne peut être celui de la nature ; la nature affecte partout un désordre sublime. [...] [U]n cabinet d'histoire naturelle est fait pour instruire ; c'est là que nous devons trouver en détail et par ordre, ce que l'univers nous présente en bloc¹⁶. » Produire un système de la nature, c'est inventer un arrangement particulier des « masses » ou « blocs » que la nature nous présente, une fiction, au moins à ses fondements, une supposition.

Les « suppositions » plus ou moins scientifiques de l'interprète de la nature lui permettent d'anticiper sur des découvertes scientifiques non encore réalisées, sans que son activité mentale soit une sorte de mythopoïèse spontanée, plus authentique ou première que la systématisation scientifique. Une telle distinction catégorielle n'a pas lieu d'être pour Diderot, d'une part parce qu'historiquement la constitution de la science comme ensemble théorique ordonné n'existera pas avant le XIX^e siècle, mais d'autre part parce que le naturalisme de Diderot est aussi une sorte de holisme épistémologique, dans lequel un énoncé subjectif tiré de l'expérience sensorielle du sujet appartient, au même titre qu'un énoncé expérimental, au réseau total d'énoncés : la « théorie une » de la « machine [sc. de l'univers, CW] qui est une ».

C'est dans ce contexte théorique, et au sein d'une ontologie matérialiste, qu'il faut comprendre la rêverie du personnage de M^{lle} de Lespinasse, dans le *Rêve*. Je cite l'échange entre son personnage et celui de l'autorité, le médecin Bordeu :

15. *Essai sur les règnes de Claude et de Néron*, II, § 97, DPV XXV, pp. 387-389.

16. « Cabinet d'histoire naturelle », in Diderot, *Œuvres*, DPV VI, p. 240. On trouve des développements complémentaires dans les articles « Animal » et « Encyclopédie ». J'analyse ce naturalisme curieusement dépendant d'un artificialisme dans « "Cabinet d'histoire naturelle", or: the interplay of nature and artifice in Diderot's naturalism » (ms.).

BORDEU. — Qu'est-ce qui circonscrit votre étendue réelle, la vraie sphère de votre sensibilité ?

MADemoiselle DE LESPINASSE. — Ma vue et mon toucher.

BORDEU. — De jour ; mais la nuit, dans les ténèbres, lorsque vous rêvez surtout à quelque chose d'abstrait ; le jour même, lorsque votre esprit est occupé ?

MADemoiselle DE LESPINASSE. — Rien. J'existe comme en un point ; je cesse presque d'être matière, je ne sens que ma pensée ; il n'y a plus ni lieu, ni mouvement, ni corps, ni distance, ni espace pour moi ; l'univers est anéanti pour moi, et je suis nulle pour lui.

BORDEU. — Voilà le dernier terme de la concentration de votre existence ; mais sa dilatation idéale peut être sans bornes. Lorsque la vraie limite de votre sensibilité est franchie, soit en vous rapprochant, en vous condensant en vous-même, soit en vous étendant au-dehors, on ne sait plus ce que cela peut devenir.

MADemoiselle DE LESPINASSE. — Docteur, vous avez raison. Il m'a semblé plusieurs fois en rêve [...] que je devenais immense. [...] Que mes bras et mes jambes s'allongeaient à l'infini, que le reste de mon corps prenait un volume proportionné ; [...] que j'escaladais le ciel, ou que j'enlaçais les deux hémisphères (V 648-649).

La rêverie de cette femme qui croit quitter son corps et s'étendre aux dimensions de l'univers s'inscrit dans une philosophie matérialiste qui a montré que l'individu, l'unité du moi, était une fiction : c'est dans ce sens que le rêve matérialiste n'est pas un songe cartésien ou une rêverie rousseauiste, car il *ne part pas du moi*. Si je suis une infinité de parcelles de matière vivante, et que tout l'univers vit, je suis aussi tout l'univers : de même que la statue de marbre qu'imaginait Diderot au début du livre était déjà potentiellement de la terre, donc une plante, donc un animal, donc moi-même qui me nourrit de la chair de cet animal, eh bien, « depuis l'éléphant jusqu'au puceron, depuis le puceron jusqu'à la molécule sensible et vivante, l'origine de tout, pas un point dans la nature entière qui ne souffre ou qui ne jouisse » (V 637). Cela conduit à un paralogisme : je ne peux pas sentir tout l'univers ! On voit là une faiblesse potentielle de l'argument qui fonderait le matérialisme sur le rêve : au contraire, comme l'a remarqué un penseur méconnu de cette époque, le fait même qu'on puisse rêver de choses inconnues, de sensations inédites, et donc de « matières » inexistantes, devrait suffire à prouver que tout n'est pas matériel¹⁷ ! Chez Diderot, la solution tiendra à la distinction entre la « contiguïté » purement spatiale et la « continuité » proprement organique : entre moi et Jupiter, il n'y a que de la contiguïté, alors qu'entre moi et le bout de mes doigts il y a de la « chair » continue. L'interaction entre différentes parties d'un organisme est possible parce que ses parties sont continues plutôt

17. C'est le raisonnement du *Précis des arguments contre les matérialistes* d'Isaac de Pinto, 2^e édition, La Haye, P.-F. Gosse, 1775, p. 101.

que simplement contigües ; elles sont continues parce qu'elles possèdent la propriété de la sensibilité. L'unité du moi est produite par l'interaction d'une multiplicité de fibres organiques, composant le « faisceau » (V 625).

Quand un philosophe matérialiste écrit un *Rêve*, il ne s'agit donc pas seulement d'un savant artifice (straussien) permettant à l'auteur de dissimuler son point de vue ou de le faire progresser de manière dialogique (bakhtinienne) par les influences respectives de personnage sur personnage ; sa pensée n'étant qu'un mode de la matière sensible, du grand tout, le point de vue de l'auteur-sujet ne peut être autre que « polypeux », pour reprendre la formule de Jacques Proust — je rappelle le passage des lettres de Diderot au sculpteur Falconet auquel Proust fait allusion :

L'argument du philosophe n'est qu'un squelette, celui de l'orateur est un animal vivant, c'est une espèce de polype. Divisez-le et il en naîtra une quantité d'autres animaux. C'est un hydre à cent têtes. Coupez une de ces têtes, les autres continueront à s'agiter, de vivre, de menacer. L'animal terrible sera blessé, mais il ne sera pas mort¹⁸.

Le rêve matérialiste cherche à « faire par la pensée ce que nature fait quelquefois » (V 643)¹⁹ ; et son contenu n'est donc pas un matérialisme pré-constitué « habillé » en rêve, pas plus qu'il n'est assimilable à un méta-discours fondé sur l'imagination — une pensée « poétique » qui serait soi-disant plus « profonde » qu'une pensée scientifique, comme je l'ai indiqué plus haut. Simplement, les mécanismes du rêve permettent d'aller *plus loin*. C'est le cas quand Bordeu, perplexe face à tant d'intertextualité, présente dès le démarrage du second dialogue — M^{lle} de Lespinasse parle, prononce une série de phrases diverses, mais on ne sait pas si c'est réellement elle qui s'exprime ou si elle est en train de « citer » le personnage D'Alembert qui a déliré dans son sommeil, délire qui la pousse à demander conseil au médecin Bordeu —, demande : « Est-ce vous qui parlez ? », et qu'elle répond : « Non. *C'est le rêveur* » (V 625).

Dans le même sens, il est remarquable que lorsqu'un personnage de Diderot rencontre une difficulté conceptuelle majeure, il aime bien clore en disant : Bonsoir²⁰, non seulement ici dans le *Rêve* mais également dans l'*Entretien d'un père avec ses enfants* ou dans la partie des *Salons de 1767* connue sous le nom de la « Promenade Vernet » (« Bonsoir mon ami » puis, quelques pages plus loin : « Adieu mon ami. Bonsoir et bonne nuit²¹ »). Simplement, comme le sait tout lecteur du *Rêve*, ici le « bonsoir » mène à un nouveau niveau de la réflexion ; pas seulement de la narration, mais de l'enchaînement théorique. Quand, à la fin du premier dialogue, c'est-à-dire l'entretien entre les deux philosophes, D'Alembert dit « Adieu mon ami, bonsoir et bonne

18. Diderot, lettre XV à Falconet, septembre 1766, DPV XV, p. 145. Voir J. Proust, « Diderot et la philosophie du polype », *Revue des sciences humaines*, vol. 14, n° 182 (avril-juin 1981).

19. Dans le texte, Bordeu l'affirme sur le mode impératif : « Faites ».

20. R. Mortier, « Rhétorique et discours scientifique dans le *Rêve de D'Alembert* » [1976], in *Le cœur et la raison*, Oxford, Voltaire Foundation, 1990, p. 241.

nuit », le personnage Diderot répond : « Vous plaisantez, mais vous rêverez sur votre oreiller de cet entretien » (V 621). En d'autres termes, ce dont je n'ai pas pu vous convaincre durant notre dialogue philosophique traditionnel, vous allez finir par y croire grâce à la... matérialité du rêve et de ses associations. La parole est inséparable de la matière organique ; une performance verbale et un état purement corporel sont liés par une seule et même causalité, au sein du « grand tout ».

* * *

Ayant décrit les grandes lignes de ce qui caractérise le rêve matérialiste, il nous reste à évacuer quelques malentendus à son sujet, en le confrontant à une objection classique, avant de conclure.

Notamment, l'idée même de rêve matérialiste ne serait-elle pas une contradiction interne, comme on le pense parfois ? Cette objection est la suivante : dans un matérialisme où tout ce qui est réel, il ne peut pas y avoir de notion d'*artifice*, donc du faux, donc du rêve, sauf au sens faible où l'on peut « mesurer » le pouls, la température ou les mouvements visuels d'un sujet et dire : à l'heure H1 il s'est mis à rêver. Remarquons que ce sens « faible », que l'on retrouve dans l'approche neurophysiologique contemporaine du rêve, « repéré trop schématiquement par la sécrétion bio-électrique du sommeil paradoxal²² », est déjà présent à l'époque de Diderot, comme l'a montré Michel Porret²³, chez le médecin suisse Samuel Formey, dans l'article « SONGE » de l'*Encyclopédie*. Pour Formey, le rêve est toujours lié à une *sensation*, et on devrait pouvoir recomposer le rêve à partir d'un complexe de sensations, en faisant l'inventaire des sensations qui affectent le dormeur. Cette approche matérialiste du rêve — qui n'est donc pas le rêve matérialiste au sens où j'emploie l'expression, pour désigner la position de Diderot — marie le mécanisme cartésien (solides, fluides) à une prise en compte de l'individualité de la personne ; elle met ensemble un premier niveau qui serait celui d'un appareillage mécanique, et un deuxième niveau qui serait celui du contenu biographique. Si nous nous tournons vers l'attitude matérialiste contemporaine face au rêve, celle-ci n'y verra pas non plus grand-chose de réel,²⁴ le réduisant au processus par lequel des événements émotionnants pour un animal étant produits la veille son sommeil paradoxal augmente en conséquence de 20 à 30 % dans la nuit suivante²⁵.

21. *Salon de 1767*, DPV XVII, pp. 230, 237.

22. B. Cyrulnik, in D. Lecourt, dir., *La pensée est-elle un produit de la sélection naturelle ?*, Paris, PUF, 1996, p. 32.

23. Conférence inédite au X^e Congrès international des Lumières, Dublin, juillet 1999.

24. Par exemple, D. M. Armstrong, *A Materialist Theory of the Mind*, Londres, Routledge Kegan Paul, 1968, pp. 303-305.

25. B. Cyrulnik, *op. cit.* Un autre point de vue « réductionniste » consiste à se demander, non plus comment donner une explication neurophysiologique des états oniriques, mais si ceux-ci constituent des « expériences » au même titre que l'expérience consciente. C'est ce qu'affirmera Daniel Dennett (« Are dreams experiences ? », *Philosophical Review*, vol. 85, n° 2 [1976]) en réaction à la théorie « externaliste » des rêves, d'inspiration wittgensteinienne, avancée par Norman Malcolm (*Dreaming*, Londres, Routledge Kegan Paul, 1959) : les rêves ne sont rien d'autre que des récits inventés une fois qu'on s'est réveillé.

Mais afin d'approcher cette grande question : comment le matérialisme peut-il penser l'artifice ?, afin de répondre à l'objection, il nous faut rapidement examiner le rapport du rêve particulier formulé par Diderot à sa *forme*, étant donné qu'il s'agit d'un artéfact *matérialiste*. L'expression de « rêve matérialiste » telle que je l'ai employée ici devrait souligner que cette forme, en l'occurrence le rêve, est *irréductiblement liée au cadre matérialiste dont elle est issue*.

Dans une analyse de la « métaphore comme expérience » dans le *Rêve de D'Alembert*, Michel Delon fait une remarque très précise et utile à notre propos :

La métaphore [...] s'articule doublement sur le projet matérialiste. L'effort pour constituer une philosophie matérialiste passe par le recensement de divers savoirs scientifiques [...] et la réflexion épistémologique, mais cherche à aller au-delà de ce front des connaissances par l'intermédiaire de la fiction. Cette fiction utilise les voies de l'expérience imaginaire ou de la métaphore. Comme, par ailleurs, l'objectif d'une telle recherche est de mettre en évidence la continuité qui fait passer de la matière à la sensibilité, puis de la sensibilité à la pensée [...] et fonde ainsi le monisme, la métaphore qui établit des analogies, rapproche des champs distincts, devient une figure privilégiée. Elle appartient à la démarche matérialiste qui tendrait, en un second temps, à la réduire scientifiquement à une simple métonymie²⁶.

Pour le dire en une phrase, par l'invention métaphorique on fonde le monisme matérialiste. Cela peut surprendre, mais souvenons-nous de la remarque du personnage Diderot à D'Alembert : « Vous rêverez sur votre oreiller de cet entretien », c'est-à-dire que l'auteur Diderot inscrit toute expression, philosophique autant que littéraire, dans une substance unique qui serait la matière vivante — et cette substance, les sujets « limités » que nous sommes la connaissent mieux en rêve que par la raison consciente. Le *Rêve de D'Alembert* met en scène la possibilité que la matière puisse penser, à la fois par le refus d'une séparation ferme entre sentir et penser, mais aussi par la capacité qu'a l'esprit, par la rêverie (c'est-à-dire un processus indissociable de divagation-spéculation-supposition), de s'étendre à l'univers entier, donc à la matière elle-même. C'est le sens de l'impératif de Bordeu : « Faites par la pensée ce que nature fait quelquefois » (V 643).

Mais il faut bien voir le « second temps » dont parle Delon : une fois existant, le matérialisme tendrait à vouloir nier le rêve en tant que fiction. Le problème s'est déjà posé une fois dans l'histoire de la philosophie, et même dans l'histoire du matérialisme, avec le *De rerum natura* de Lucrèce, comme l'a montré Alain Gigandet²⁷. Pas seulement parce que c'est un poème (d'autres

26. M. Delon, « La métaphore comme expérience dans le *Rêve de D'Alembert* », in A. Becq et al., dir., *Aspects du discours matérialiste en France autour de 1770*, Caen, Université de Caen / Paris, J. Touzot, 1981, p. 290.

27. Voir son excellente analyse de Lucrèce intitulée « Matérialisme et poésie », in J.-Cl. Bourdin, dir., *Les matérialismes philosophiques*, Paris, Éd. Kimé, 1997.

auteurs, comme Nietzsche, ont eux aussi innové sur le plan formel en philosophie), mais parce qu'il cherche à exprimer par ce poème la vérité d'une philosophie, l'épicurisme, qui s'inscrit dans la lignée réductionniste de l'atomisme ancien, pour lequel les formes esthétiques n'avaient guère de sens. Nous voilà face au problème mentionné plus haut : comment concevoir l'artifice dans une pensée matérialiste ?

En effet, le matérialisme, considéré comme système où tout est réel et actuel, ne pourrait pas concevoir le possible ni même le faux, au sens fort. « Avec le temps, tout ce qui est possible dans la nature, est », dit Diderot ; les espèces existantes sont celles qui n'ont pas été exterminées²⁸. Dans un univers entièrement actuel, il ne peut pas y avoir d'« erreur », puisqu'une erreur n'est qu'un mode déficient du réel, elle n'a pas d'existence propre ni de modalité, au sens où la logique modale étudie des propriétés *possibles* et non actuelles : selon la formule de Hilary Putnam, « *modal properties are not part of the materialist's furniture of the world* »²⁹. Et l'artifice ? La notion d'artifice sous-entend une distinction entre le vrai et le faux ou le naturel et l'artificiel, qui ne semble pas permise au matérialisme, pour deux raisons, qui viennent en partie d'être indiquées : d'abord sa dimension « productiviste » ou « actualiste » : tout ce qui existe est réel et dans ce sens « vrai » (difficile dans ce cadre de donner un statut particulier à la fiction : l'*Illiade*, dira d'Holbach, n'est que le résultat d'un « assemblage » fortuit de molécules, de « dés pipés par la nature »³⁰), ensuite son refus de critères normatifs extra-territoriaux, c'est-à-dire appartenant à une sphère qui ne serait pas celle de la nature elle-même : si un être existe, nous n'avons pas de critères avec lesquels l'évaluer autre que l'existence.

Mais si la nature de Diderot est aussi celle des cabinets d'histoire naturelle, nature monstrueuse et rêvée — malgré la continuité causale globale — alors nous ne sommes plus dans un univers naturaliste au sens normatif, justement (l'attitude lucrétienne classique consistant à évaluer tous les phénomènes du point de vue de la nature, pour ensuite les *réduire* et *éliminer* ce qui n'y est pas réductible, comme par exemple la superstition), puisqu'on reconnaît que le monde tel que nous le percevons n'est qu'un « masque »³¹. Au contraire, ce matérialisme « moderne » (au sens où Diderot emploie l'expression polémique mais surtout programmatique de « spinosistes modernes » dans l'article

28. Diderot, *Sur deux mémoires de D'Alembert*, § 1, *Le calcul des probabilités* (1761), DPV II, p. 351 ; *Éléments de physiologie* (inachevés, rédigés essentiellement à la fin des années 1770), DPV XVII, 295.

29. H. Putnam, « Why There isn't a Ready-Made World », in *Realism and Reason : Philosophical Papers 3*, Cambridge, Cambridge University Press, 1987 ; repris in P. Moser & J.D. Trout, dir., *Contemporary Materialism*, Londres, Routledge, 1995, p. 237. Évidemment, l'univers de Diderot étant également un univers de « vicissitude » qui nie radicalement l'essence, on a pu dire à l'inverse que tout chez Diderot est modalité et contingence (J. Proust, « Diderot et la philosophie du polype », *op. cit.*, notamment pp. 22-24).

30. *Système de la Nature* (1770, 2^e édition 1781), réédition, J. Boulad-Ayoub, dir., Paris, Fayard, coll. « Corpus », livre II, ch. v, note 41.

31. Diderot, *Éléments de physiologie*, DPV XVII, p. 296.

« SPINOSISTE »³², pour les différencier des « spinosistes anciens », qui sont des fondationnalistes en métaphysique) accepte la dimension poïétique de l'être humain, qui transforme le monde autour de lui. Une fois cette dimension reconnue, la différence entre la nature et l'artifice disparaît. Un livre, un désir ou un mensonge sont tout aussi réels qu'une molécule ou un volcan — surtout quand ils sont pensés comme modes de la matière universellement vivante et sentante (donc pensante).

* * *

Nulle réduction d'un état mental à une normalité biochimique ou plus précisément neurochimique dans le grand rêve matérialiste ; au contraire, tout ce qui est « fou » et « excessif » est la condition même de la progression d'un expérimentalisme, qui ne se passe pas des conjectures, suppositions et autres constructions mentales. Faisant par la pensée ce que la matière fait parfois, le rêve matérialiste est une démultiplication des points de vue par laquelle la pensée s'étend aux dimensions de l'univers, imaginant des univers empédocléens multiples (« Dans Jupiter ou dans Saturne des polypes humains », V 630) tout en réaffirmant la sensibilité universelle de la matière (« Qui sait ce qu'est l'être pensant et sentant en Saturne ? », V 635), au lieu d'en revenir au monologue d'un philosophe, éveillé ou somnolent.

³² *Enc.* XV, 477a. Pour un excellent commentaire de cet article voir Alexandre Métraux, « Über Denis Diderots Physiologisch Interpretierten Spinoza », *Studia Spinozana* vol. 10 (1994).

La modélisation des comportements non conséquentialistes en théorie du choix rationnel

Learry Gagné

Volume 34, numéro 2, automne 2007

URI : <https://id.erudit.org/iderudit/016995ar>

DOI : <https://doi.org/10.7202/016995ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Société de philosophie du Québec

ISSN

0316-2923 (imprimé)

1492-1391 (numérique)

[Découvrir la revue](#)

Citer cet article

Gagné, L. (2007). La modélisation des comportements non conséquentialistes en théorie du choix rationnel. *Philosophiques*, 34 (2), 329–352.
<https://doi.org/10.7202/016995ar>

Résumé de l'article

Nous tentons, dans cet article, de déterminer la place des comportements non conséquentialistes, notamment le respect des valeurs et des normes, dans la théorie du choix rationnel. Au départ, il n'y a pas de limites à ce qui peut constituer une préférence ou une valeur d'utilité; tout comportement non conséquentialiste peut être réduit à un comportement conséquentialiste. Un bref examen de certains modèles rationnels des normes sociales nous montre, d'une part, que la réduction conséquentialiste du conformisme laisse inexpliqués certains phénomènes importants, et d'autre part, que l'emploi d'utilités autres qu'instrumentales pose de sérieux problèmes à la théorie. Pour terminer, nous discutons de deux autres approches de la théorie du choix rationnel faisant une large part aux comportements non conséquentialistes. Nous concluons que si, pour une même action, le type de motivation (conséquentialiste ou non) influence en soi les résultats de l'action collective, alors il nous faut dans ce cas sortir du cadre de la rationalité, sans toutefois complètement l'abandonner.

La modélisation des comportements non conséquentialistes en théorie du choix rationnel¹

LEARRY GAGNÉ

Université du Québec à Montréal

gagnel@awale.qc.ca

RÉSUMÉ. — Nous tentons, dans cet article, de déterminer la place des comportements non conséquentialistes, notamment le respect des valeurs et des normes, dans la théorie du choix rationnel. Au départ, il n'y a pas de limites à ce qui peut constituer une préférence ou une valeur d'utilité ; tout comportement non conséquentialiste peut être réduit à un comportement conséquentialiste. Un bref examen de certains modèles rationnels des normes sociales nous montre, d'une part, que la réduction conséquentialiste du conformisme laisse inexpliqués certains phénomènes importants, et d'autre part, que l'emploi d'utilités autres qu'instrumentales pose de sérieux problèmes à la théorie. Pour terminer, nous discutons de deux autres approches de la théorie du choix rationnel faisant une large part aux comportements non conséquentialistes. Nous concluons que si, pour une même action, le type de motivation (conséquentialiste ou non) influence en soi les résultats de l'action collective, alors il nous faut dans ce cas sortir du cadre de la rationalité, sans toutefois complètement l'abandonner.

ABSTRACT. — In this paper, we attempt to find the place of non-consequentialist behavior, like respect of norms and values, in rational-choice theory. At the onset, there is no limit as to what can constitute a preference or a utility value ; any non-consequentialist behavior can be reduced to a consequentialist one. A brief survey of several rational models of social norms shows us, on one hand, that consequentialist reduction of conformity leaves some important phenomena unexplained, and, on the other hand, that recourse to non-instrumental utilities generates serious problems within the theory. We end with a discussion of two alternative approaches to rational-choice theory that leave a large place for non-consequentialist behavior. We conclude by saying that if, for the same action, the type of motivation behind it (consequentialist or not) has by itself an influence on the outcome of collective action, then in such cases we have to step out of rational-choice theory, without however abandoning it.

Les succès de la théorie du choix rationnel en sciences sociales sont indéniables. Nous constatons toutefois que les modèles rationnels s'avèrent plus convaincants dans des champs où la motivation des agents est manifestement conséquentialiste et maximisatrice (les transactions de marché par exemple) que dans des champs où ce genre de motivation est moins évident (les pratiques religieuses, le choix d'un époux, etc.). Certains adeptes de la théorie répliqueront que les motivations réelles ne les intéressent pas, car leur méthode consiste à faire

1. J'aimerais remercier Robert Nadeau et Dominique Leydet pour leur soutien et leurs précieux commentaires lors de la rédaction de cet article.

« comme si » les agents étaient rationnels. D'autres affirmeront que les motivations non conséquentialistes importent, il s'agit de les traiter comme utilités spéciales au sein de la théorie. Dans cet article, nous nous demandons si, effectivement, la théorie du choix rationnel peut traiter des comportements non conséquentialistes de manière satisfaisante.

La théorie du choix rationnel s'intéresse avant tout aux comportements conséquentialistes : je vise un certain but, et j'applique les meilleurs moyens permettant de l'atteindre, considérant les coûts et les bénéfices des différentes possibilités. Un comportement non conséquentialiste est de nature « je fais X parce qu'il le faut ». Prenons l'exemple de deux témoins d'un crime questionnés par un juge, dans une cause où dire la vérité pourrait s'avérer incriminant. Le témoin A n'exhibe pas de préférences *a priori* pour l'honnêteté mais choisit de dire la vérité, car il croit que les sanctions contre une fausse déclaration seraient plus coûteuses que les inconvénients de l'honnêteté². Le témoin B est mû par une valeur d'honnêteté ; il choisit de dire la vérité « parce qu'il le faut », comme on dit. Ces deux choix peuvent être modélisés comme une maximisation d'utilité : A recherche la plus petite peine possible, et B cherche à maximiser l'honnêteté. Du point de vue de la rationalité, a-t-on raison d'aborder ces deux agents de la même manière ? On remarque au départ que le fait que l'agent maximise son utilité ne constitue pas en soi une preuve suffisante de sa rationalité ; il faut également qu'il choisisse les moyens adéquats aux fins recherchées³. Or nous constatons que A choisit la vérité comme moyen en vue d'une fin indépendante, l'évitement de sanctions, alors que B confond moyens et fins, car il choisit la vérité dans le but d'être honnête, c'est-à-dire quelqu'un qui dit la vérité. C'est ce dernier type de choix que nous qualifions de « non conséquentialiste », car même s'il est d'une certaine manière maximisant, il est effectué pour lui-même, et non en vertu d'un but qui lui est extérieur. Ne serait-il pas plus approprié de considérer l'honnêteté comme une motivation distincte des motivations instrumentales ? Au lieu de voir l'action honnête comme un moyen calculé en vue de satisfaire une fin recherchée, l'honnêteté, nous pourrions concevoir une telle action simplement comme une manière d'agir qui ne se veut pas nécessairement rationnelle.

Les comportements de nature non conséquentialiste qui nous intéressent avant tout sont ceux qui peuvent être aisément redécrits en termes de rationalité. Nous laisserons de côté les comportements à la fois non conséquentialistes

2. On pourrait également dire qu'il préfère l'honnêteté au mensonge *ceteris paribus*, tout en étant particulièrement sensible à la structure de coûts, de sorte qu'il puisse rapidement choisir de mentir dès que la vérité s'avère désavantageuse.

3. La théorie de l'utilité contient en elle-même le critère de maximisation : lorsque l'agent classe ses options sur une échelle d'utilité, ce n'est logiquement pas dans le but de choisir le second sur la liste. Dans le contexte de l'utilité formelle, la rationalité par la maximisation devient tautologique (Hausman, 1992, 18), d'où la nécessité de recourir à l'adéquation des moyens aux fins de la définition de la rationalité (Demeulenaere, 1996, 240).

et considérés comme irrationnels dans la conception classique de la rationalité⁴. Nous porterons notre attention sur les actions qui ont un but, lorsque ce but se décrit dans le langage de l'action, de sorte que l'action ne semble pas être entreprise pour une autre raison que sa propre réalisation. Le champ des valeurs et des normes nous semble exhiber cette propriété : lorsque l'agent est motivé par une valeur fortement intériorisée, il a tendance à agir au nom de cette valeur, sans égard aux bénéfices potentiels autres que le respect de cette valeur. Le second témoin de la petite histoire mentionnée plus haut dit la vérité, car l'honnêteté est pour lui une valeur ; les sanctions contre une fausse déclaration n'entrent pas dans son raisonnement. Nous nous pencherons sur les valeurs et leur traitement en théorie du choix rationnel, mais uniquement sur cet aspect non conséquentialiste, et non sur la distinction entre intérêts égoïstes et non égoïstes parfois soulevée qui, sans être inintéressante, nous éloignerait de notre propos. Dans ce qui suit, nous effectuerons un survol des types de modélisation rationnelle des comportements non conséquentialistes. Nous débiterons par un débat épistémologique sur la pertinence de distinguer ces comportements. Nous ferons par la suite l'examen d'une série de tentatives en choix rationnel de modélisation des valeurs qui leur confèrent un statut particulier par rapport aux préférences « ordinaires ». Pour terminer, nous approfondirons la discussion à l'aide de deux approches critiques de la théorie du choix rationnel, qui prennent au sérieux la distinction entre conséquentialisme et non conséquentialisme, soit les alternatives proposées par Raymond Boudon et Jon Elster.

La portée de la théorie du choix rationnel

Il faut toujours garder à l'esprit que, dans la théorie classique du choix rationnel, il n'y a aucune limite à ce qui constitue une préférence ou une valeur d'utilité ; donc, en dernière instance, la question des comportements non conséquentialistes ne s'y pose pas, car on peut toujours faire « comme s'ils » étaient conséquentialistes, pour autant qu'ils ne soient pas manifestement irrationnels. Dès lors se pose toutefois la question de la portée empirique de la théorie. On peut l'aborder de deux manières. Prenant l'agent comme point d'ancrage, la critique empirique *en amont* cherche à déterminer si celui-ci est bel et bien rationnel, en d'autres termes s'il est réellement motivé par une norme de rationalité. On peut faire fi de cette critique en amont en adoptant une conception idéale de l'agent et en le supposant rationnel, même si l'on sait très bien que ce n'est pas toujours le cas. C'est l'*hypothèse de rationalité* : l'agent

4. En bref, nous postulons trois catégories de comportements irrationnels (ou non rationnels) : le comportement *forcé*, soit par la présence d'un seul élément de choix, d'une réaction impulsive (y compris les actions causées par les émotions), ou d'une autorité coercitive ; le comportement *habituel*, où l'agent ne change pas ses choix malgré un changement de situation ; et à l'inverse le comportement *aléatoire*, où l'agent change constamment d'idée malgré un environnement stable.

est supposé agir en tout temps en fonction de sa préférence la plus élevée⁵. La critique empirique se déplace alors *en aval* de l'agent ; on cherche à savoir si le modèle rationnel constitué de tels agents idéalisés nous permet effectivement de mieux comprendre certains phénomènes sociaux. La théorie du choix rationnel a tendance à prendre plus au sérieux ce genre de critique. Sur cette question, on peut classer ses variantes en deux catégories : les théories générales et spéciales (Goldthorpe, 2000 : 123-125). Avec les théories générales, on peut se permettre de modéliser à peu près n'importe quelle situation sociale comme une interaction d'agents rationnels ; on pense à « l'impérialisme économique » de l'École de Chicago, représentée par l'œuvre de Becker⁶ et, plus récemment, par Levitt et Dubner (2005). Les théories spéciales, quant à elles, cherchent à modéliser le comportement des agents d'une manière autre que pleinement rationnelle lorsque cette hypothèse ne semble pas tenir la route ; nous croiserons certaines instances de théories spéciales dans la seconde partie. Pour ce qui est de la modélisation des comportements non conséquentialistes, on peut s'attendre à ce que la théorie du choix rationnel, du moins dans ses variantes spéciales, prenne en considération le réalisme des résultats du modèle, et non le réalisme du comportement attribué à l'agent.

La théorie du choix rationnel peut ramener tout comportement non conséquentialiste en comportement rationnel maximiseur en adoptant des motivations et des préférences *ad hoc*. À défaut de se poser la question du réalisme de l'entreprise, nous pouvons tenter de voir si l'hypothèse de rationalité tient bien le coup. L'hypothèse de rationalité est, bien entendu, un principe de charité. Dans l'explication du comportement, on suppose d'abord que l'agent maximise une utilité quelconque, et ce n'est qu'après avoir épuisé les ressources de la théorie du choix rationnel que nous sommes en mesure d'affirmer que le comportement s'avère forcé, habituel, ou aléatoire. Comme les préférences peuvent porter théoriquement sur n'importe quoi, les limites du principe de charité vont dépendre de la modélisation particulière. Si l'on observe que l'agent perd systématiquement son argent à la suite de ses choix, que l'on soit prêt à le déclarer irrationnel ou bien montrant une préférence pour la pauvreté dépend entièrement de ce que l'on veut modéliser.

Partons d'un exemple commun de comportement non conséquentialiste : suivre une tradition. Pour Goldthorpe (2000 : 128-129), le comportement traditionnel, adopter toujours la même stratégie face au même problème, peut s'avérer rationnel si cette stratégie s'avère efficace. Par contre, si la situation de l'agent change suffisamment pour que la stratégie traditionnelle cesse d'être

5. Cependant, certaines variantes de la théorie vont quand même chercher à justifier cette position épistémologique. Par exemple, on peut soutenir que la rationalité est une motivation réelle car, face à sa propre irrationalité, l'agent ressentira un malaise et tentera de se corriger (Follesdal, 1982, 316). Ce genre de postulat psychologique n'est cependant pas nécessaire, tant que l'hypothèse de la rationalité ne concerne que la cohérence des choix (Ferejohn, 2002 : 224).

6. Notamment pour Becker (1986, 111-112), les comportements non optimaux s'expliquent par des coûts « monétaires ou psychiques » qui ont pu échapper à l'observateur. De plus, l'agent n'est pas nécessairement conscient de sa propre motivation rationnelle.

efficace, et que celui-ci persiste dans ses choix, alors, si l'on veut demeurer fidèle à la théorie du choix rationnel, il nous faut postuler soit une irrationalité, soit une préférence pour la tradition, et cette dernière lui apparaît beaucoup trop *ad hoc* pour justifier une modélisation rationnelle. On retrouve un semblable questionnement dans le fameux « paradoxe du vote » : pourquoi se déplacer pour aller voter alors que les coûts sont non négligeables et que le bénéfice potentiel que notre vote fasse une différence au décompte final est si minime ? Brennan (1991 : 89 n. 14), prenant le camp de la rationalité, fait remarquer que l'on peut facilement expliquer le geste de voter à une élection en invoquant une préférence pour le vote. Une telle réduction des motivations, quoiqu'extrême, demeure légitime. À l'inverse, pour Etzioni (1988 : 43-45), la fusion fins-moyens que l'on retrouve typiquement dans les décisions morales constitue une raison suffisante pour les dissocier des décisions instrumentales. La coupure ne nous apparaît toutefois pas aussi nette. Même si l'agent motivé par une valeur forte ne porte pas *habituellement* attention aux coûts et aux bénéfices de l'action, il est aisé, pour n'importe quelle valeur, d'imaginer une situation où les coûts de l'action morale seront tellement élevés que l'on pourra prédire que l'agent intégrera ces coûts dans son raisonnement. Il adoptera alors une attitude *instrumentale* face au respect de la valeur, et sa motivation correspondra à ce que la théorie du choix rationnel avait toujours supposé⁷. Même la moralité « impérative » chez Etzioni ne peut pas être complètement dissociée de la rationalité instrumentale :

People first sense an absolute command to act morally, but that does not mean that they will always heed it. That they are less likely to heed it if the costs are high does not indicate that there is no imperative ; indeed, all other things being equal, it is what drives up the costs. (Etzioni, 1986 : 167)⁸.

L'hypothèse de rationalité étant un principe de charité, on peut le pousser aussi loin qu'on le désire. Ces théoriciens se distinguent avant tout par le degré de charité qu'ils sont prêts à accepter.

Certains admettent leur impuissance à prendre en défaut l'hypothèse de rationalité de l'agent sur une base conceptuelle. Pour Goldthorpe, lorsque le respect de la valeur domine inconditionnellement le choix que recommanderait une analyse en termes de coûts et de bénéfices, « *it would again seem best to acknowledge that the limits of applicability of RAT [Rational Action Theory] are reached* » (Goldthorpe, 2000 : 129). Il reconnaît cependant que le théoricien du choix rationnel pourrait « sauver » un tel comportement comme rationnel. Le choix des mots est important (« *it would seem best...* »), car par là,

7. Pettit (1995) nous fournit une image de la rationalité comme étant des « balises virtuelles » qui viennent limiter nos actions morales lorsqu'elles s'éloignent trop de notre intérêt personnel.

8. La dernière partie de la citation porte à confusion, car elle semble indiquer que le caractère impératif de la valeur augmente son propre coût d'exécution. La proposition a plus de sens si on lit « *...it is what drives up the cost threshold* ». L'impératif rend l'agent plus tolérant aux coûts qu'il endosse en agissant de façon morale.

Goldthorpe admet qu'il n'y a pas de critère objectif de rejet d'un modèle rationnel ; cela devient alors une question de jugement sur ce qui semble fonctionner le mieux comme modèle. D'autres abordent ce problème sensiblement de la même manière. Isaac (1997 : 560-63) définit les « règles fortes » comme des règles de conduite indépendantes des aspects de la situation de choix. Toutefois, les règles fortes ne constituent pas selon lui un problème « logique » pour la théorie du choix rationnel, car on peut toujours interpréter le choix en termes de maximisation, mais la « maximisation abstraite » qu'une telle approche suppose affaiblit l'aspect empirique de la théorie, car elle s'intéresse peu aux motivations. Dans la même veine, Hausman et McPherson affirment : « *Even though it may be possible formally to model the committed agent as maximizing utility, it seems enlightening not to do so* » (Hausman, McPherson, 1993 : 688), à cause du rôle de la moralité dans les motivations de l'agent. Ce sont là des jugements de valeur sur la pertinence de l'hypothèse de rationalité selon les circonstances, et non des critiques conceptuelles. On peut se poser de sérieuses questions cependant sur le caractère tautologique des explications rationnelles des comportements non conséquentialistes, comme expliquer le conformisme par une préférence pour le conformisme (Demeulenaere, 1996 : 247-269).

Il semble bien que si l'on veut demeurer fidèle à la théorie du choix rationnel, il y a toujours moyen de modéliser les valeurs comme des préférences ou des contraintes, et de les inclure dans le calcul d'utilité. Pour Brennan, tant que nous considérons les valeurs comme *données*, et il n'y a aucune raison de faire autrement en choix rationnel, elles ne posent aucun problème de modélisation. Ce n'est, selon lui, que lorsque les valeurs sont choisies, et choisies pour un but autre que la maximisation de son bien-être, qu'elles peuvent devenir problématiques. Ce serait là de toute façon un questionnement sur la formation des valeurs, et le choix rationnel ne s'en préoccupe pas. Nous ne pouvons qu'être d'accord avec sa conclusion : « *An agent's preferences are simply taken to be what they are revealed to be ; a more complex conception is, by the 'norms' of economic methodology, both empirically unnecessary and theoretically illogical* » (Brennan, 1991 : 93). La loi de l'offre et de la demande suppose que les agents soient motivés à payer, ou à obtenir le meilleur prix possible pour leurs biens, même si dans le monde réel les agents peuvent montrer d'autres motivations. On peut exiger que les théoriciens s'intéressent davantage aux motivations, mais il n'y a rien dans la théorie qui les y oblige. La théorie du choix rationnel se distingue formellement de la sociologie interprétative, mais on peut espérer que ses partisans « interprètent », de temps à autre, leurs sujets afin d'ajouter du réalisme à leurs modèles⁹ (Ferejohn, 2002 : 227-28).

9. Satz et Ferejohn (1994) adoptent la position de l'« externalisme modéré » : la théorie n'a pas à s'intéresser aux motivations internes des agents, mais lorsque les motivations modélisées correspondent aux motivations réelles, la théorie ne peut que s'en porter mieux.

Modèles rationnels des valeurs

Nous allons maintenant nous intéresser plus en détail aux efforts entrepris dans les modèles rationnels visant à traiter des valeurs et des normes. Nous regroupons ces modèles selon qu'ils traitent des valeurs comme des utilités standard, c'est-à-dire conceptuellement conformes en tous points aux utilités matérielles, ou comme des utilités non standard possédant des caractéristiques particulières les distinguant des utilités matérielles.

Modèles rationnels standard

Supposons que nous défendions la théorie du choix rationnel contre les critiques en amont ; nous acceptons donc que les agents soient modélisés d'une façon idéalement rationnelle. La critique se déplace alors en aval : les modèles rationnels d'action collective sont-ils toujours adéquats ? L'approche évidente serait d'évaluer le réalisme des résultats du modèle. Cependant, la théorie du choix rationnel ne se veut pas absolument réaliste, et même si le modèle dévie de l'empirisme, ses conclusions peuvent conserver un intérêt non négligeable. Le type de critique que nous offrons ici est différent, et nous croyons qu'il constitue une charge plus sérieuse contre certains modèles rationnels, particulièrement les modèles intégrant rationalité et normes. Cette critique va comme suit : si on suppose que les agents adoptent *consciemment* et *ouvertement* la motivation rationnelle que le modèle leur attribue, est-ce que celui-ci tient toujours ? La question devient pertinente lorsque, dans une dynamique d'action collective, pour une *même* action, l'agent reçoit un « paiement » différent selon que sa motivation se présentait de façon conséquentialiste ou non conséquentialiste. Dans un tel cas, un modèle rationnel qui convertit les motivations non conséquentialistes verra lui échapper des aspects importants de la dynamique. Nous soutenons que les actions collectives issues des normes sociales présentent cette caractéristique ; il y a une distinction à faire entre le conformisme sincère (non conséquentialiste) et instrumental aux normes, notamment dans la réception de telles motivations par autrui.

Nous adopterons une définition des normes assez commune en rationalité, celle de Pettit (1990, 731) : une régularité de comportement R est une norme lorsque presque tout le monde approuve que chacun adopte R (et désapprouve le contraire), et que cette approbation a un effet sur l'adoption de R. Il existe des normes auxquelles il est individuellement et immédiatement rationnel de se conformer ; respecter les feux de circulation, par exemple, est rationnel, peu importe ce que les autres font. Un modèle rationnel serait ici parfaitement adéquat, par contre on préfère souvent dans de telles situations parler d'une « convention » plutôt que d'une « norme ». Les normes dans les modèles rationnels servent le plus souvent à résoudre des problèmes d'action collective pour lesquels il n'est pas immédiatement rationnel de coopérer, mais lors d'un seuil critique où les agents coopèrent, un surplus bénéfique à tous est généré. Ce sont les cas classiques de production de biens publics et de Dilemmes du prisonnier. La coopération rationnelle peut être assurée par des

institutions établissant un système de sanctions¹⁰, ou par une norme. Dans ce dernier cas, on introduit un désir d'approbation sociale (ou un évitement de la désapprobation) dans la fonction d'utilité de l'agent¹¹.

De tels modèles supposent que l'agent se conforme à la norme de manière conséquentialiste. Pour lancer la discussion, reprenons l'exemple de l'honnêteté en introduction. L'honnêteté est souvent perçue dans le choix rationnel comme une norme d'action collective (Akerlof, 1983 ; Pettit, 1990 : 735). Il n'est pas toujours immédiatement rationnel de dire la vérité, mais une société où tout le monde (ou presque) est honnête est meilleure qu'une société de menteurs. Si, de plus, on introduit un système d'approbation et de désapprobation d'autrui, on voit comment l'honnêteté peut s'avérer rationnelle après tout. Mais de quoi aurait l'air une société où la motivation à l'honnêteté serait purement conséquentialiste ? Il est difficile de croire qu'un agent se déclarant honnête dans le but explicite d'être apprécié d'autrui et d'éviter les sanctions soit digne d'approbation. En fait, l'agent maximisera son utilité en *paraissant* sincèrement honnête, donc en paraissant agir de manière non conséquentialiste. Le grain de sable dans l'engrenage d'un modèle rationnel de l'honnêteté est que l'honnêteté conséquentialiste n'est jamais un aussi bon « choix » que l'honnêteté sincère, et, néanmoins, cette première forme d'honnêteté est la seule qu'il puisse recommander.

Riker (1980) s'oppose à une conception de la confiance comme « attitude » (sans doute non conséquentialiste) pour y préférer une conception conséquentialiste :

Dig just below the surface, even of diffuse trust or support or passive alienation, and numerous individual choices based on rational calculation will appear. The conclusion is that people do think about their circumstances, do seek their advantage, and are not precluded by personality or history or social circumstances from changing their judgments as their situations change (Riker, 1980 : 6).

Il faut tout de même reconnaître qu'un agent dont « l'attitude » est de respecter ses promesses sera plus digne de confiance que celui qui ne les respecte que lorsque cela tourne à son avantage. La rationalité stricte de Riker ne peut rendre compte de ce phénomène.

Certains auteurs rationalistes ont reconnu le problème. Une solution proposée consiste à rendre la rationalité de l'action latente. Hardin (1995 : 86) prend l'exemple du slogan « *Black is Beautiful* » chez les Noirs américains. Même si l'adoption par chacun d'un tel slogan s'avère rationnel par l'effet de renforcement de la solidarité de groupe qui en résulte, l'adopter explicitement *pour cette raison* serait contre-productif. Il propose alors un modèle rationnel fonctionnaliste admettant la rationalité latente : le bénéfice collectif sert

10. La branche néo-institutionnaliste de la théorie du choix rationnel traite justement de ce genre de situation (Ostrom, 1986 ; Shepsle, 1989 ; Goodin, 1996).

11. Comme représentants de tels modèles, voir Höllander (1990), Frank (1998), Nee et Ingram (1998) et Bernheim (1994).

d'incitatif à l'adoption du slogan par une boucle de rétroaction, mais sans que l'agent ne s'en aperçoive. Brennan et Pettit (2000) offrent une approche similaire. L'agent adopte une stratégie rationnelle lui permettant de gagner l'estime d'autrui, mais puisque la recherche ouverte d'estime est vouée à l'échec, il découpe sa stratégie en une succession de petites actions qui feront en sorte qu'il finira par oublier sa véritable motivation. C'est ici aussi un modèle fonctionnaliste, car bien que l'agent n'agisse pas pour gagner l'estime d'autrui, si cette estime cesse alors il cessera son action. La rationalité latente constitue une hypothèse auxiliaire à l'hypothèse de rationalité, qui stipule que si la motivation rationnelle influence négativement l'utilité de l'agent, alors il est possible d'être rationnel sans être rationnellement motivé.

On retrouve, dans les modèles évolutionnaires, une autre manière de se départir des motivations rationnelles tout en conservant la rationalité. La théorie évolutionnaire du choix rationnel cherche avant tout à simplifier au maximum le concept de l'agent. Ici, les individus s'apparentent plus à des « hôtes » qu'à des « agents » ; ils acquièrent leurs préférences et leurs stratégies à travers un processus semblable à la sélection naturelle plutôt que par choix volontaire. Dans un jeu évolutionnaire, les joueurs sont tirés au hasard dans une population et placés dans le jeu dans où ils appliquent leurs stratégies et sur lequel ils n'ont à peu près pas d'information a priori ; cette procédure est ensuite répétée un certain nombre de fois, et la stratégie « gagnante » est celle qui se sera propagée le plus dans la population selon un mécanisme de sélection particulier¹². Le principe général de sélection est ce que Bowles et Gintis nomment « répliation différentielle » : « *Durable aspects of behavior, including norms, habits, and rules of thumb, may be accounted for by the fact that they have been copied, diffused, and hence replicated, while other traits have not* » (Bowles, Gintis, 1998 : 212). Dans les modèles évolutionnaires, les stratégies sont *sélectionnées* par un algorithme exogène à l'agent ; elles ne sont pas *choisies* par un processus endogène ; en d'autres mots, on sort l'hypothèse de rationalité de l'agent pour la placer dans l'algorithme. Il n'y a donc même pas lieu de se demander si le choix s'avère conséquentialiste ou non. Nul besoin de préciser que ce genre d'explication s'avère encore plus fonctionnelle que les précédentes.

Modèles rationnels non standard

Certains théoriciens s'intéressent de plus près à la multiplicité des motivations des agents. Insatisfaits de la modélisation rationnelle *ad hoc* des valeurs et des normes, ils assignent des caractéristiques particulières aux préférences dites « morales » par rapport aux préférences plus matérielles, tout en voulant respecter les postulats de la théorie du choix rationnel. Dans le champ économique, nous rencontrons fréquemment des modèles de préférences multiples. Un agent peut choisir d'acheter telle voiture plutôt que telle autre sur le marché parce qu'elle représente le meilleur compromis entre ses

12. Voir Weibull (1998) pour un excellent résumé de la théorie ; également Danielson (2004).

préférences pour la performance, la consommation d'essence et l'apparence, par exemple. On qualifie ces préférences d'*homogènes*, car, partageant la même nature (ici comme caractéristiques de voitures), on peut les classer sur une même échelle globale d'utilité. L'agent peut sacrifier un peu de l'utilité que lui procure la satisfaction de la préférence A pour une plus grande satisfaction de la préférence B. La théorie du choix rationnel est très bien outillée pour aborder ce genre de décision et consacrer un ordonnancement des possibilités en assignant à chacun des choix une valeur globale d'utilité. Mais avec des préférences *hétérogènes*, où l'on cherche à distinguer conceptuellement préférences matérielles et morales, le calcul de l'utilité finale devient une affaire beaucoup plus délicate.

Nous avons retenu deux types de modélisation des préférences hétérogènes dans la littérature rationaliste. Les modèles d'*utilités multiples* établissent une distinction entre utilité égoïste et non égoïste, ce dernier terme pouvant signifier, selon le modèle, utilité altruiste, sociale, ou encore, émanant des valeurs. Les modèles de *préférences hiérarchiques*, pour leur part, conçoivent les valeurs comme des préférences de second ordre, gouvernant celles de premier ordre. Ces approches engendrent des tensions au sein de la théorie du choix rationnel, car elles remettent en question le sens même « d'utilité » et de « préférence ». Il ne faudrait pas confondre les utilités multiples avec ce que nous pourrions nommer les « utilités additives », de la forme $U = UA + UB$, où la fonction d'utilité nous montre ses diverses composantes et comment elles affectent la courbe d'utilité *unique* de l'agent. De tels modèles demeurent parfaitement en règle avec la théorie du choix rationnel¹³. Ce qui nous intéresse dans l'idée d'utilités multiples, c'est que ce sont des utilités fondamentalement conflictuelles, qui ne peuvent directement s'intégrer dans une fonction unique. Etzioni (1986 : 171-72) propose de distinguer « plaisir » et « devoir », le premier concernant les choix intéressés, et le second, le respect de nos valeurs et normes. Son raisonnement est que nous associons nos préférences égoïstes à une recherche de plaisir, alors que le devoir nous enjoint fréquemment à accomplir des tâches qui nous sont déplaisantes. Toutefois, il ne nous indique pas comment calculer un choix final fondé sur ces deux paramètres. Si on les additionne tout bonnement, nous sommes de retour à la théorie standard, et le « devoir » ne jouit alors d'aucune propriété spéciale. Si cela n'est pas désirable, alors on doit expliquer comment deux paramètres présumés incommensurables peuvent aboutir à un choix unique :

To compete with the neoclassical paradigm of rational choice, proponents of multiple utility representations must show how conflicting motivations are

13. Höllander (1990) nous en fournit un bon exemple avec son modèle de la « coopération volontaire » en groupe. Sa fonction d'utilité comprend deux éléments principaux : la contribution individuelle et l'approbation du groupe. Ici, biens matériels et approbation peuvent s'échanger librement. Son modèle d'approbation dans la production de biens publics prédit, à l'équilibre, un certain niveau de contribution et d'approbation correspondante. Cet équilibre est ensuite perçu par tous comme une norme sociale, cadrant ainsi avec le principe des attentes mutuelles de la théorie de l'échange social.

resolved into outcomes. But as soon as this step is taken, they are involved in all of the 'commensurability' of extended utility representations (Isaac, 1997 : 547-548).

Ce que Isaac nomme « utilité étendue » correspond à nos utilités additives. Examinons à présent deux exemples de modèles d'utilités multiples.

Dans le modèle de Heath (2001), les résultats des actions reçoivent une valeur d'utilité (espérée) standard, alors que les actions elles-mêmes se voient attribuer une valeur de « convenance » (*appropriateness*), une « contrainte déontique » qui vient spécifier le degré (numérique) de permission, d'interdiction, ou d'obligation de l'action. Ces contraintes déontiques proviennent des normes sociales et des valeurs individuelles. Avec $A(a)$ pour la convenance de l'action a et $U(a)$ pour l'utilité standard de a , nous obtenons la « fonction de valeur » globale : $V(a) = kA(a) + U(a)$. Ces utilités sont différenciées, car, maintient-on, l'évaluation « normative » de l'action n'est pas la même chose que l'évaluation « instrumentale » de son résultat. Mais pour que $V(a)$ ait un sens, il faut bien que ces utilités se situent sur une même échelle (Heath, 2001 : 374), donc soient commensurables et comparables d'une manière qui demeure inexplicitée. Heath adopte une défense plutôt méthodologique qu'épistémologique :

I think it is very important to keep distinct the contribution that norms of action and desires for outcomes make to deliberation, because some social behavior becomes intelligible only when these two are treated separately. [...] Thus any attempt to merge norms and desires into a unified utility theory function results in a loss of information that in turn renders certain aspects of social interaction needlessly obscure (Heath, 2001 : 378-79).

Au premier regard, tout cela a l'air d'un modèle d'utilités additives, mais Heath maintient que l'évaluation normative ne peut être traitée comme un élément de l'utilité standard¹⁴. Il prétend pouvoir résoudre le Dilemme du prisonnier en attribuant une valeur de convenance élevée à la coopération, mais on peut aboutir au même résultat simplement en augmentant directement l'utilité (standard) de la coopération telle qu'on la retrouve dans la matrice de jeu. Dans la forme développée du jeu, cela reviendrait à écrire la valeur de convenance directement sur les vecteurs plutôt que sur les points terminaux, ce qui est absurde en théorie des jeux. De toute évidence, les valeurs normatives ne jouissent pas ici de caractéristiques spéciales par rapport aux utilités standard.

14. L'argument est que les préférences sur les manières d'agir violent la théorie Von Neumann-Morgenstern de l'utilité (face à deux loteries distinctes aux résultats escomptés identiques, l'agent est supposé indifférent), sauf si elles sont directement intégrées aux résultats des choix ; mais si on les intègre, on perd de l'information précieuse sur les motivations des agents (Heath, 2001 : 378-79). Cet argument tombe à plat lorsque l'on considère que, pour autant que les préférences sur les manières d'agir et les préférences sur les résultats escomptés demeurent formellement comparables, rien ne nous empêche de ventiler une fonction d'utilité VNM afin de montrer l'influence de ces deux facteurs d'utilité. C'est d'ailleurs de cette manière que la propension au risque est habituellement représentée.

Le modèle de la « surcharge morale » de Kuran (1998) constitue une autre tentative de séparation de valeurs d'utilités. Il recouvre les situations de choix où certaines options, sinon toutes, sont évaluées comme immorales par l'agent. Sans entrer dans les détails, la fonction d'utilité déduit la désutilité morale de l'action de son utilité « intrinsèque » ou standard. Même si Kuran est catégorique au sujet du statut particulier de la désutilité morale, soutenant que son modèle « *departs from the classical theory of individual choice, according to which one has a unitary preference ordering* » (Kuran, 1998 : 232), il apparaît nettement, de par sa fonction d'utilité, que la moralité constitue en fait un facteur d'utilité menant, au bout du compte, à un ordonnancement unique des préférences. Les valeurs correspondent ici à des « *judgments about preference orderings or about the choices that preferences have generated* » (Kuran, 1998 : 232), mais telles qu'on les retrouve dans le modèle, elles servent à assigner une utilité à chacun des choix, pas à l'ordonnancement en tant que tel. Le but du modèle est de démontrer qu'une fois un choix payant mais pas très moral effectué¹⁵, son utilité intrinsèque s'estompe peu après, alors que sa désutilité morale perdure, résultant en une « dissonance morale » et amenant des sentiments de culpabilité et de regret. Ce phénomène n'a toutefois rien à voir avec la fonction d'utilité précédant le choix. On doit faire appel à une seconde fonction, *post hoc*, où l'utilité intrinsèque sera décroissante avec le temps, et où, encore une fois, la moralité ne jouera pas un rôle conceptuellement distinct.

Les modèles des préférences hiérarchiques proposent une avenue différente. Ici, les valeurs ne sont pas intégrées dans la fonction d'utilité ; elles servent plutôt à filtrer les préférences immorales. Dans le paradigme du choix rationnel, on peut les considérer comme des préférences de préférences. Le modèle le plus connu est celui de Sen, dans son fameux article « *Rational Fools* » (1982). Parmi les motivations non égoïstes, il y distingue l'altruisme rationnel (la « sympathie ») du choix fondé sur les valeurs (« l'engagement ») de la manière suivante : « *If the knowledge of torture of others makes you sick, it is a case of sympathy ; if it does not make you feel personally worse off, but you think it is wrong and you are ready to do something to stop it, it is a case of commitment* » (Sen, 1982 : 92, italique ajouté). La satisfaction ou la frustration d'une valeur n'affecte pas l'utilité de l'agent ; donc, si l'on considère que l'action morale comporte des coûts réels (temps, argent, etc.) et des bénéfices moraux n'entrant pas dans la fonction d'utilité, l'engagement constitue un choix non maximiseur (Sen, 1982 : 93), ce qui nous conduit à ce passage fréquemment cité, que l'engagement « *drives a wedge between personal choice and personal welfare* » (Sen, 1982 : 94). L'engagement fonctionne par l'ordonnancement des ordonnancements de préférences. L'agent ordonne

15. Kuran pense à un agent qui quitte un restaurant sans laisser de pourboire.

d'abord ses préférences de premier ordre selon certaines échelles d'utilité¹⁶, et ensuite procède à l'ordonnancement de ces ordonnancements en vue de la maximisation de valeurs, si l'on suppose qu'il aspire effectivement à agir moralement (Sen, 1982 : 100-101). Une telle modélisation des valeurs comme contraintes fortes ne peut fonctionner que si le méta-ordonnancement est *lexicographique*, c'est-à-dire, si l'on n'a recours aux ordonnancements de premier ordre que pour trancher les situations d'indifférence laissées par le méta-ordonnancement. La procédure prend fin à l'instant où le modèle retient un seul choix maximal. Une procédure de décision si rigide ne semble pas pouvoir s'accommoder avec la plupart de nos choix quotidiens, là où la distinction entre moralité et intérêt personnel n'apparaît pas aussi claire. La proposition de Sen pour combler ce besoin de compromis moral est de laisser l'agent, selon la situation, court-circuiter certaines préférences au sommet de l'ordonnancement final afin de lui permettre de choisir plus bas (Sen, 1982 : 100), mais cela jette un doute sur la véritable nature d'une contrainte morale que l'agent peut violer apparemment sans restrictions. Comment peut-il alors choisir de se conformer à une valeur tout en gardant à l'esprit son intérêt particulier ? Si l'on veut demeurer fidèle à la théorie du choix rationnel, l'attribution d'utilités aux valeurs, dans une fonction globale d'utilité, semble inévitable, mais on doit alors abandonner le modèle des préférences hiérarchiques.

Ansperger (1998) va tenter une réconciliation des deux méthodes. Afin de régler ce qu'il appelle, suivant Rawls, les « tensions de l'engagement », il propose que figure au sommet du méta-ordonnancement au moins un ordonnancement lexicographique (c'est pour lui la condition d'existence d'un ordonnancement moral), suivi d'ordonnements moraux faibles pouvant faire l'objet de compromis, et, enfin, d'ordonnements selon l'intérêt personnel. Parmi l'ensemble résiduel de choix laissés par les ordonnancements lexicographiques, l'agent peut recourir à une fonction d'utilité où il assignera des valeurs d'utilité à ses valeurs, avec un poids correspondant à leur position dans l'ordonnancement global (Ansperger, 1998 : 205-207). Si ce modèle s'avère correct, c'est-à-dire que le modèle des préférences hiérarchiques de Sen ne peut par lui-même gérer le compromis moral, et donc qu'une fonction quelconque d'utilités additives est requise (Ansperger, 1998 : 206), nous devons conclure que ce modèle n'a rien de spécial à offrir¹⁷. En effet, on pourrait le remplacer par un modèle d'utilités additives dans lequel les impératifs moraux les plus chers à l'agent seraient conçus comme des biens infiniment inélastiques (Brennan, 1989 : 200).

16. Par exemple, dans le modèle de la « rationalité de rôle », l'agent peut avoir un ordonnancement particulier pour chacun des rôles sociaux qu'il porte (Goodin, 1986 : 88-89). L'appréciation d'une situation en tant que mère n'est pas nécessairement la même qu'en tant qu'employée ou croyante.

17. Voir la critique similaire de Wolfelsperger (2001 : 76-77).

Conclusion

Nous avons survolé quelques échantillons de modèles rationnels cherchant à rendre compte des motivations non conséquentialistes. Le problème central rencontré dans les modèles standard est que les motivations non conséquentialistes en elles-mêmes peuvent influencer l'issue du jeu, et donc, que leur réduction à des motivations conséquentialistes perdra de vue cet aspect. Les solutions à tendances fonctionnalistes ne nous apparaissent pas attrayantes. À notre avis un tel affaiblissement de l'hypothèse de rationalité ne sert pas la cause de la théorie du choix rationnel. Pour ce qui est des modèles faisant une place spécifique aux motivations non conséquentialistes, nous constatons qu'ils se retrouvent dans l'une ou l'autre de deux catégories. Soit que les utilités sont commensurables, ce qui signifie que l'on peut toujours réduire le modèle à un modèle d'utilités additives, pleinement compatible avec la théorie standard de l'utilité mais n'offrant aucune conception particulière des utilités morales, soit qu'elles s'avèrent incommensurables, et il apparaît alors impossible de pouvoir construire une fonction d'utilité appropriée sans violer les axiomes de la théorie. Les utilités additives conservent une fonction heuristique, elles nous permettent d'observer l'influence de divers facteurs dans la détermination de l'utilité globale, mais elles ne peuvent conceptuellement distinguer préférences matérielles et valeurs. La commensurabilité implique que toutes les utilités puissent se mesurer à l'échelle du « grand X », « *an abstract entity, without any content, neither pleasure nor consumption. It is merely the great common denominator, an X, into which all other values can be converted or by which all rank-ordering can be systematized* » (Etzioni, 1988 : 164). L'échec des modèles d'utilités multiples et des préférences hiérarchiques semble démontrer que l'on ne peut répondre à la critique en amont de l'hypothèse de rationalité en introduisant de nouveaux types de préférences sans rompre la théorie du choix rationnel.

Deux théories alternatives de la rationalité

Après avoir examiné s'il était possible d'intégrer les comportements non conséquentialistes au sein de la rationalité instrumentale tout en préservant le sens, nous allons nous pencher sur deux théories alternatives de la rationalité qui évitent de ramener ces comportements à un moyen pour leur propre fin et leur donnent une place bien à eux. Les deux auteurs que nous proposons adoptent chacun une stratégie de prime abord opposée : Raymond Boudon cherche à *élargir* la notion de rationalité afin de pouvoir y inclure les comportements non conséquentialistes sans avoir à les réduire à un instrumentalisme quasi-tautologique, tandis que Jon Elster cherche à *restreindre* la portée de la rationalité aux comportements « réellement » conséquentialistes pour ensuite proposer une théorie hors rationalité des comportements non conséquentialistes.

Le « modèle rationnel général » de Boudon

L'approche de Boudon consiste à ramener les comportements non conséquentialistes dans le giron de la rationalité, en élargissant la notion de « rationalité » au-delà de sa signification traditionnelle en théorie du choix rationnel. En gros, l'explication rationnelle ne consiste plus seulement à montrer que l'agent a choisi les meilleurs moyens en vue d'une fin, mais, plus largement, à montrer qu'il avait des raisons, instrumentales ou non, d'agir comme il l'a fait. Il décompose la théorie classique du choix rationnel en six postulats, en ordre d'implication ascendante : P1 (*individualisme*), P2 (*compréhension*, au sens d'intelligibilité de l'action), P3 (*rationalité*, au sens d'action fondée sur des raisons), P4 (*conséquentialisme/instrumentalisme*), P5 (*égoïsme*) et P6 (*maximisation*) (Boudon, 2003a : 19-21). Pour Boudon, la théorie du choix rationnel propose de réduire « toutes les formes de la rationalité à la rationalité instrumentale » (Boudon, 2003a : 133) ; elle « fait fausse route en prétendant accorder un statut général aux postulats du conséquentialisme et de l'égoïsme, qui ne sont pertinents que dans des cas particuliers » (Boudon, 2003a : 121). Il propose comme solution de remplacement le « Modèle rationnel général », moins restrictif, composé uniquement des postulats P1 à P3 (Boudon, 2003a : 49).

Le Modèle rationnel général comprend trois types de rationalité : instrumentale, cognitive et axiologique. La rationalité instrumentale est bien entendu représentée en réintroduisant le postulat P4. La rationalité cognitive concerne la formation des croyances. L'agent agit au nom d'une croyance X, car il a de fortes raisons¹⁸ de croire que X est vrai (ou probable) dans un certain « contexte cognitif » (Boudon, 2003a : 61). Ici, dans la formation de ses croyances, l'agent ne cherche pas à maximiser son utilité, mais bien à déterminer si ses connaissances sont vraies (Boudon, 1998 : 188). Deux remarques s'imposent. D'abord, il faudrait savoir pourquoi Boudon oppose croyances instrumentales et cognitives. Dans les cas des croyances instrumentales, s'il s'agit d'une formation de croyances sans égards à la vérité, dans le seul but de satisfaire ses désirs, alors c'est une irrationalité (une forme de *wishful thinking* ou de *self-deception*). S'il s'agit plutôt d'une collecte d'information en vue d'une action instrumentale, la recherche de la vérité s'impose de façon implicite chez l'agent rationnel. Nous voyons que, dans ces deux cas, il n'y a pas de conception de la rationalité cognitive qui se distingue du choix rationnel. La seconde remarque concerne la relation entre rationalité et vérité. Boudon affirme que la rationalité cognitive permet de qualifier des croyances fausses de « rationnelles » pour autant que l'agent les croient vraies sur la base de raisons fortes. Contre la rationalité instrumentale, Boudon mentionne une « idée reçue », les associations entre

18. Boudon distingue les « bonnes raisons » propres à l'agent des « raisons fortes », qui, elles, sont entretenues par tous ceux qui partagent les mêmes « paramètres contextuels » (Boudon, 2003a : 160). On ne remarque pas de différences conceptuelles à proprement parler ; les raisons fortes sont simplement plus convaincantes pour l'agent que les bonnes raisons, et fournissent donc une base plus solide pour l'explication.

vérité et rationalité, et entre fausseté et irrationalité, « implicitement présentes dans bien des théories » (Boudon, 2003a : 81). Boudon ne mentionne pas d'exemples de ces théories, mais il est clair qu'il ne peut s'agir du choix rationnel qui, à ma connaissance, n'a jamais prétendu, sous quelque forme que ce soit, que les croyances devaient être objectivement vraies pour être rationnelles. Bien au contraire, les croyances au sein du choix rationnel sont *subjectives*, c'est-à-dire que l'agent doit *croire* qu'elles sont vraies autant que possible.

La rationalité axiologique opère de manière semblable à la rationalité cognitive : l'agent est axiologiquement rationnel s'il accomplit X parce qu'il a de fortes raisons de croire que X est juste ou bon. Alors que nous ne constatons pas de différences réelles entre la motivation rationnelle et la recherche de la vérité, la rationalité axiologique nous semble plus prometteuse comme alternative. Les raisons axiologiques se fondent sur le « programme de la morale », qui a pour but de promouvoir la « dignité de l'individu » ; ce programme se veut « indéfiniment déclinable » (Boudon, 2003a : 128), ce qui évite à Boudon d'avoir à imposer une moralité particulière comme rationnelle. Les raisons cognitives et axiologiques sont de nature non conséquentialiste.

Les comportements non conséquentialistes sont donc rationnels, au sens que Boudon donne à ce terme. Ce type de rationalité s'insère entre la rationalité instrumentale et l'irrationalité, telles que conçues dans la théorie classique du choix rationnel. La frontière entre la rationalité instrumentale et la rationalité cognitive/axiologique est assez nette, elle correspond à la frontière du conséquentialisme, représentée par l'ajout du postulat P4. Qu'en est-il alors de la frontière entre ce type de rationalité et l'irrationalité ? Quelle place occupent chez Boudon les comportements irrationnels ? Il semble bien, a priori, que ce soit le domaine des actions qui ne sont pas soutenues par des raisons, ou, en d'autres termes, qui ne respectent pas le postulat P3. Boudon a très peu à dire sur l'irrationalité ; il semble qu'il veuille pousser son concept de rationalité le plus loin possible, englobant un maximum de types de comportements. À la fin d'un article sur le « modèle cognitif » (le précurseur du Modèle rationnel général), il écrit : « *Irrationality should be given its rightful place. 'Traditional' and 'affective' actions also exists. Moreover, all actions rest on a ground of instincts* » (Boudon, 1998 : 200) ; et nous retrouvons ceci comme seule description du phénomène, au milieu d'une définition formelle de la rationalité : « Ainsi, d'une mère qui par énervement gifle son enfant, l'on dira : « Elle n'avait *pas de raisons* de gifler l'enfant, mais... ». Ce comportement est *compréhensible*, il n'est pas *rationnel* » (Boudon, 2003b : 194, en italique dans l'original). Afin de mieux comprendre, observons comment Boudon traite d'un phénomène particulier, les croyances « à moitié », où l'on croit à des mythes (par exemple) sans vraiment y croire, « on ne mettrait pas sa main au feu pour les défendre ; mais, dans bien des cas, on n'a guère d'intérêt pour leur réfutation. Plus : on répugne à les mettre en doute » (Boudon, 2003a : 139). C'est ce que Elster (2004) nomme « *wishful thinking* », et, pour lui, il s'agit clairement d'une irrationalité, car la formation de la croyance se voit biaisée par

le plaisir immédiat qu'elle apporte à l'agent. Boudon, quant à lui, désire éviter la « vision binaire » du couple rationnel/irrationnel en proposant de s'intéresser à une « typologie des modes de conviction ». Pour ce faire, il applique la notion de « satisfaction », au sens de Simon, aux croyances : l'agent a beau chercher la vérité, il se contente volontiers de croyances suffisamment proches. Cet affaiblissement du critère de « bonne raison » cognitive permet d'inclure dans une typologie sommaire des modes de conviction, des phénomènes tels l'idéologie, la « fausse conscience », ou la « mauvaise foi ». Ce sont de bonnes raisons, mais pas des raisons fortes (Boudon, 2003a : 139-141).

Le but du Modèle rationnel général est d'expliquer les comportements des agents en retraçant leurs raisons d'agir comme ils le font. La théorie du choix rationnel recherche une explication plus resserrée, elle cherche à rendre compte des intérêts des agents et ainsi à expliquer les comportements par le choix de moyens appropriés à la maximisation de ces intérêts. Il s'agit moins d'un postulat sur la nature humaine que d'un prérequis essentiel pour fins de modélisation. Comme nous l'avons constaté précédemment, les modélisations rationnelles impliquant les utilités de l'agent s'avèrent problématiques lorsque l'on y intègre des comportements non conséquentialistes. Boudon note avec justesse que les modèles de choix rationnel ont bien du mal à traiter du paradoxe du vote (Boudon, 2003a : 38-42). Considérer que les raisons de voter (ou pas) peuvent être de nature axiologique permet certes un plus grand réalisme, mais on ne voit pas immédiatement comment on pourrait en tirer un modèle explicatif rigoureux. La théorie de Boudon nous indique que ces autres rationalités existent, et qu'elles ont une influence sur les motivations des agents, mais, pour l'instant, cette théorie ne semble pas outillée pour la construction de modèles.

Les « motivations mixtes » de Elster

Dans la théorie de Boudon, l'irrationalité occupe une place minimale. Elle ne s'applique qu'aux comportements dénués de raisons, ce qui ne semble désigner que les actes instinctifs. Nous retrouvons chez Elster une définition beaucoup plus étoffée de l'irrationalité. Sa stratégie consiste en deux étapes : d'abord circonscrire le champ d'application de la théorie du choix rationnel en situant un certain nombre de comportements (dont les comportements non conséquentialistes) hors de la théorie, ensuite postuler une motivation supplémentaire à la motivation rationnelle, mettant en jeu les émotions. Commençons par sa conception de la rationalité. Elster subdivise le champ de la rationalité en théorie « étroite » et « large » (Elster, 1983 : chap. 1). La théorie étroite correspond *grosso modo* à la conception orthodoxe du choix rationnel, telle que nous l'avons abordée jusqu'ici, qui exige la cohérence des préférences (complètes, continues et transitives) et des croyances (non contradictoires).

La théorie étroite n'est pas suffisante pour Elster, car elle ne peut que soutenir que l'agent s'est effectivement comporté de manière rationnelle (ou non), sans se soucier qu'il soit tombé sur le bon choix après mûre réflexion,

par hasard, ou par conformisme aveugle. La théorie large de la rationalité inclut la théorie étroite et s'intéresse en plus aux mécanismes de formation des désirs et des croyances. Dans la théorie elstérienne du choix rationnel, les désirs et les croyances de l'agent doivent présenter une « histoire causale » à laquelle l'agent peut s'identifier¹⁹ (Elster, 1986 : 15). La théorie large se situe quelque part entre la rationalité étroite et une théorie « du bien et du vrai » (Elster, 1983 : 15) ; elle cherche à expliquer les « bons choix » d'une manière plus enrichissante que la simple cohérence. Les croyances bien formées sont une affaire de *jugement*. Si l'acquisition d'information est coûteuse, il apparaît impossible de pouvoir former des croyances optimales par rapport à la situation, car on ne peut savoir à l'avance quelle sera la valeur d'une information que l'on ne possède pas encore. Tout ce que l'agent peut faire, c'est user de son jugement afin de déterminer s'il a acquis suffisamment d'information pertinente au choix qu'il a à effectuer²⁰. Les désirs, quant à eux, doivent avoir été formés de façon *autonome*, c'est-à-dire libres, autant que possible, d'influences causales échappant au contrôle de l'agent, comme les préférences adaptatives au contexte, le conformisme (ou l'anti-conformisme), l'inertie, etc. En fait, sont considérés comme non autonomes les désirs auxquels l'agent ne « s'identifie » pas pleinement, pour reprendre la citation utilisée plus haut.

Les critères de jugement et d'autonomie n'ont pas de définition propre ; ils sont plutôt compris comme des qualificatifs que l'on applique respectivement aux croyances et aux désirs, à la condition qu'ils ne présentent pas certaines formes typiques d'irrationalité (Elster, 1983 : 24). Elster subdivise les comportements irrationnels en deux catégories : les irrationalités non motivées (*cold*), comme les erreurs inférentielles, et les irrationalités motivées (*hot*) (Elster, 1983 : 25-26 ; 1989 : 23-24). Ces dernières supposent que l'agent désire, la plupart du temps, agir de façon rationnelle ; donc qu'il entretient la rationalité comme motivation. À cela s'ajoute une seconde motivation, non rationnelle, les émotions. Elster soutient que les croyances et les désirs influencés par les émotions ne peuvent pas conduire à une action pleinement rationnelle. Les émotions peuvent influencer les désirs, à travers la colère, la jalousie, la honte, etc. « Agir sous le coup de l'émotion » signifie que l'agent n'est pas principalement motivé par l'atteinte réfléchie de ses objectifs, mais qu'il éprouve plutôt des désirs plus ou moins insensibles aux calculs coûts/bénéfices. Les croyances peuvent également être influencées par les émotions. Nous avons alors affaire au *wishful thinking*, lorsque la croyance est formée selon le « principe du plaisir » au détriment du « principe de réalité », ou à la *self-deception*, lorsqu'une croyance déplaisante se voit remplacée par une autre plus plaisante (Elster, 2004).

19. Cela constitue pour Elster une « idée préanalytique » (Elster, 1989 : 6).

20. Cet argument constitue pour Elster le talon d'Achille de la théorie étroite, qui réfute en quelque sorte sa propre hypothèse d'optimalité des décisions. Le « *satisficing* » de Simon constitue une tentative remarquée de réponse à ce problème.

La présence de motivations non rationnelles impose une limite à la portée explicative de la théorie du choix rationnel. Sa primauté normative n'est pas affectée, car, dans cette optique, elle ne fait que suggérer le meilleur moyen pour réaliser certaines fins. Dans la première mouture de ses travaux sur le choix rationnel, Elster suggérait que l'on complète la théorie par une théorie psychologique du comportement et une théorie sociologique des normes sociales (Elster, 1986 : 22-27 ; 1989 : 30-35). Plus tard (Elster, 1999), il ne retiendra que les motivations émotive et rationnelle, la motivation à se conformer aux normes étant désormais conçue de façon surtout émotive, sans totalement exclure la rationalité. La sociologie sert toujours à expliquer l'écologie des normes sociales, mais celles-ci n'ont plus le statut de motivations propres.

Elster n'offre pas de définition des émotions, se contentant de quelques esquisses de catégorisation²¹. L'aspect le plus important, selon lui, de la motivation émotionnelle par rapport à la motivation rationnelle, c'est son côté « viscéral », au sens où l'influence émotionnelle pousse l'agent à dévaloriser les conséquences de ses gestes²² (Elster, 1999 : 287). On ne peut simplement considérer les émotions comme un coût psychique (ou un bénéfice) venant modifier les valeurs d'utilité des autres solutions, et, ainsi, l'ordonnancement des préférences. Pour Elster, le caractère viscéral des émotions fait en sorte qu'en plus d'agir comme facteur d'évaluation des solutions, elles peuvent entraver, voire empêcher la réflexion rationnelle (Elster, 1999 : 304, 413). Sauf, peut-être, dans le cas d'une émotion très intense, l'agent est généralement motivé à la fois par sa raison et ses émotions ; c'est ce que Elster nomme les « motivations mixtes ». Il ne saurait être question « d'émotions rationnelles » dans le cadre de la théorie large de la rationalité. Si un comportement émotif représente effectivement le meilleur moyen pour en arriver à une fin donnée, la théorie étroite peut s'en accommoder, mais un tel comportement viole, du moins en partie, le critère de l'autonomie des désirs, et parfois même le recours au jugement dans la formation des croyances. L'agent veut pouvoir justifier ses gestes par des *raisons*, non seulement des impulsions²³.

La théorie des motivations mixtes fait plus que simplement affirmer l'existence de comportements authentiquement non conséquentialistes, elle leur donne une caractéristique particulière. Chez Elster, le comportement motivé par des émotions ne peut pas se ramener à une préférence pour la satisfaction de telle ou telle émotion, comme cela se passe dans la théorie classique

21. Il soutient qu'il existe trop de confusion théorique dans ce domaine, et que de toute façon c'est l'aspect phénoménal des émotions qui l'intéresse le plus (Elster, 1999 : 241-43).

22. « *Visceral arousal is an important criterion for deciding that a state is an emotion rather than a simple belief-desire complex* » (Elster, 1999 : 247).

23. À mon avis, cela ne signifie pas que l'agent désire toujours agir de façon rationnelle. Celui-ci peut vouloir se laisser porter par des émotions positives. Mais lorsqu'il désire agir rationnellement, un comportement trop influencé par les émotions sera considéré comme un échec. L'expression « garder la tête froide », qui est en fait un conseil de réussite, illustre bien ce propos.

du choix rationnel, car les émotions entravent les capacités réflexives de l'agent. Pour pouvoir rendre compte de ce phénomène, la théorie du choix rationnel devrait concevoir une fonction d'utilité qui non seulement quantifie la satisfaction émotionnelle de l'agent, mais qui considère également que celui-ci se retrouve *moins* en mesure de se servir de cette fonction. Le problème, relevé chez Boudon, de la modélisation des comportements non conséquentialistes demeure. Elster propose de se servir de « mécanismes sociaux », « *frequently occuring and easily recognizable causal patterns that are triggered under generally unknown conditions or with indeterminate consequences* » (Elster, 1998 : 45). Il apparaît évident, dans les deux cas, que prendre les comportements non conséquentialistes au sérieux implique un abandon de la rigueur de la théorie du choix rationnel.

Conclusion

Boudon et Elster partagent une préoccupation semblable pour les comportements non conséquentialistes. Là où ils ne semblent pas s'accorder, c'est sur la conception de la rationalité. Alors que pour Boudon il suffit que l'action soit fondée en raison pour être qualifiée de rationnelle, Elster s'en tient plutôt à la rationalité instrumentale, augmentée de critères concernant la formation des désirs et des croyances. Donc, quant aux comportements non conséquentialistes qui ne sont pas *manifestement* irrationnels sous n'importe quelle définition, Boudon aura tendance à les classer comme rationnels, et Elster comme irrationnels. Bien que Boudon reconnaisse du bout des lèvres la nécessité d'une conception de l'irrationalité, il en minimise simultanément la pertinence : « Souligner la dimension cognitive de la rationalité [...] [c]'est aussi se donner les moyens d'échapper aux délimitations arbitraires entre rationalité et irrationalité, comme celle qui oppose l'irrationalité de l'*explication par les normes* à une rationalité réduite à la rationalité instrumentale [...] » (Boudon, 2003a : 160 ; en italique dans l'original). L'exemple qu'offre Boudon est frappant, car il s'agit, sans la nommer, de la position d'Elster !

L'avantage principal, selon nous, de la théorie d'Elster est qu'elle nous fournit une caractérisation d'un type de comportement non conséquentialiste, celui impliquant les émotions, qui se situe conceptuellement en porte-à-faux avec l'hypothèse de rationalité, puisque cette caractérisation remet en question la capacité de l'agent à agir de façon pleinement rationnelle. En fondant la propriété motivationnelle des valeurs sur les émotions, on peut éviter de parler « d'impératifs » suivi d'exceptions, comme le fait Etzioni, et parler plutôt d'un amalgame émotion-rationalité où les proportions peuvent varier selon la situation. La caractérisation des comportements non conséquentialistes chez Boudon nous apparaît confuse. Comme argument soutenant sa théorie en opposition à la théorie du choix rationnel, Boudon propose son propre modèle du comportement des étudiants dans le système scolaire, qu'il introduit ainsi : « *Far from making the individual educational decisions a mere effect of cost-benefit calculations, I introduced the idea that they derive, rather, from a system of*

contextualized arguments » (Boudon, 1998 : 193). Il nous présente ensuite le processus décisionnel de trois étudiants idéaux typiques, qui varie selon leur croyance en leurs habiletés scolaires (croyances cognitives) et leurs préférences pour un statut social élevé (croyances axiologiques). Il apparaît clairement, toutefois, que leur décision de poursuivre ou non leurs études est de nature conséquentialiste : à partir d'une préférence pour un niveau donné de statut social et d'une croyance sur ses habiletés scolaires, l'agent choisit le moyen (poursuivre ou non ses études) qui maximisera la préférence. Pourquoi alors Boudon prétend-il qu'il s'agit d'une critique du choix rationnel ? Au moins deux réponses sont possibles. D'abord, il s'en prend peut-être au postulat de l'égoïsme (P5), car pour lui le statut social n'est pas nécessairement une affaire d'intérêt personnel ; l'agent peut se révéler sensible à la réputation familiale, par exemple. Cette critique est invalide, car, dans la théorie du choix rationnel, ce à quoi l'on réfère souvent comme « égoïsme » ne signifie ni plus ni moins que l'utilité propre à l'agent ; cela n'exclut pas qu'il puisse tirer son utilité de la satisfaction d'autrui. Une seconde réponse serait qu'il ne situe pas sa critique au niveau de la décision finale, mais exclusivement au niveau des états mentaux : ces croyances ne sont pas « instrumentales », car elles concernent, respectivement, une réalité empirique et un jugement de valeur. Cette conception des rationalités cognitive et axiologique revient souvent chez Boudon ; elle demeure extrêmement problématique notamment, car, comme nous l'avons relevé plus haut, la notion de « croyance instrumentale » n'a pas de sens en dehors du *wishful thinking* et ne fait certainement pas partie de l'hypothèse de rationalité instrumentale.

Bien que Elster nous offre une distinction beaucoup plus nette entre comportements conséquentialistes et non conséquentialistes, une grande ambiguïté demeure. Ses concepts de désirs « autonomes » et de croyances issues d'un « jugement » au sein de la théorie large de la rationalité demeurent vagues et intuitifs, et sa définition des émotions ne va pas beaucoup plus loin qu'un exposé phénoménal. Néanmoins, son opposition conceptuelle entre émotions et rationalité ouvre la voie à une possibilité de modélisation combinant la rigueur de la théorie du choix rationnel et certains mécanismes sociaux traitant des motivations non conséquentialistes.

1.4 Conclusion

En étudiant les modélisations rationnelles du conformisme aux valeurs et aux normes, nous en sommes venu au constat qu'au sein de la théorie du choix rationnel on n'a d'autre choix que de réduire les motivations non conséquentialistes aux motivations conséquentialistes. Si l'on adopte la stratégie de la rationalité postulée, on se prive de comportements réels pouvant avoir un effet significatif sur le phénomène collectif à l'étude, alors que si l'on adopte la stratégie dite « non standard », on a tendance à affaiblir la théorie en l'affligeant d'hypothèses auxiliaires sur l'agent. D'où l'intérêt principal de restreindre la portée de la théorie du choix rationnel aux champs et sous-champs où les

agents se montrent conséquentialistes : on préserve ainsi pleinement l'hypothèse de rationalité, la pierre de touche de la théorie, tout en produisant des modèles plausibles d'action collective.

Loin de nous l'idée de vouloir diviser les sciences sociales en champs « rationnels » et « non rationnels » et de confiner la théorie du choix rationnel à ces premiers. Nous préconisons plutôt des approches où l'explication d'un phénomène collectif peut comporter à la fois des aspects rationnels et non rationnels. C'est précisément ce qu'ont tenté d'accomplir Boudon et Elster, dans des optiques différentes. De telles approches exigent, par contre, que l'on doive abandonner la rigueur et la simplicité des modèles rationnels. En définitive, nous avons retenu la théorie des motivations mixtes d'Elster, non pas pour sa rigueur, car elle nous semble encore à un stade précoce, mais parce que ses propositions sur l'irrationalité et les émotions nous semblent une voie prometteuse pour une théorie qui permettra de traiter à la fois des comportements conséquentialistes et non conséquentialistes.

Nous pourrions faire le même constat à propos d'autres théories sociales, comme, par exemple, la théorie de la démocratie délibérative (Gagné, 2002). Ici, les citoyens sont supposés ne pas chercher à maximiser leur intérêt personnel, et ils sont également supposés être ouverts aux arguments d'autrui en toute bonne foi ; c'est ce que nous pourrions nommer « l'hypothèse de l'agent raisonnable », parallèlement à l'hypothèse de rationalité. Mais que faire lorsque les agents réels ne s'avèrent pas raisonnables ? Encore une fois, au lieu de chercher à « sauver » la théorie en excluant certains comportements et en en modifiant d'autres, il vaudrait mieux, selon nous, adopter une approche mixte de motivations, en invoquant les modèles délibératifs là où ils se révèlent vraiment pertinents et en n'hésitant pas à faire appel à la rationalité dans d'autres cas.

Bibliographie

- Akerlof, George A. « Loyalty Filters », *The American Economic Review*, vol. 73, 1983, p. 54-63.
- Arnsperger, Christian. « Engagement moral et optimisation individuelle », in Mahieu, F.-R. ; Rapoport, H. (dir), *Altruisme. Analyses économiques*, Paris, Economica, 1998, p. 191-214.
- Becker, Gary. « The Economic Approach to Human Behavior », in Elster, J. (dir), *Rational Choice*, New York Univ. Press, 1986, p. 108-122.
- Bernheim, B. Douglas. « A Theory of Conformity », *Journal of Political Economy*, vol. 102, 1994, p. 841-877.
- Boudon, Raymond. « Social Mechanisms Without Black Boxes », in Hedström, P. ; Swedberg, R. (dir), *Social Mechanisms*, New York, Cambridge Univ. Press, 1998, p. 172-203.
- . *Raison, bonnes raisons*, Paris, PUF, 2003a.
- . « Rationalité », in Boudon, R. et al. (dir.), *Dictionnaire de sociologie*, Paris, Larousse, 2003b, p. 194-195.

- Bowles, Samuel ; Gintis, Herbert. « How Communities Govern : The Structural Basis of Prosocial Norms », in Ben-Ner, A. ; Putterman, L. (dir), *Economics, Values, and Organization*, New York, Cambridge Univ. Press, 1998, p. 206-230.
- Brennan, Geoffrey ; Pettit, Philip. « The Hidden Economy of Esteem », *Economics and Philosophy*, vol. 16, 2000, p. 77-98.
- Brennan, Timothy J. « A Methodological Assessment of Multiple Utility Frameworks », *Economics and Philosophy*, vol. 5, 1989, p. 189-208.
- . « The Trouble with Norms », in Koford, K. J. ; Miller, J. B. (dir), *Social Norms and Economic Institutions*, Ann Arbor, Univ. of Michigan Press, 1991, p. 85-94.
- Danielson, Peter. « Rationality and Evolution », in Mele, A. R. (dir), *The Oxford Handbook of Rationality*, Oxford Univ. Press, 2004, p. 417-437.
- Demeulenaere, Pierre. *Homo oeconomicus*, Paris, PUF, 1996.
- Elster, Jon. *Sour Grapes*, New York, Cambridge Univ. Press, 1983.
- . « Introduction », in Elster, J. (dir), *Rational Choice*, New York Univ. Press, 1986, p. 1-33.
- . *Solomonic Judgments*, New York, Cambridge Univ. Press, 1989.
- . « A Plea for Mechanisms », in Hedström, P. ; Swedberg, R. (dir), *Social Mechanisms*, New York, Cambridge Univ. Press, 1998, p. 45-73.
- . *Alchemies of the Mind*, New York, Cambridge Univ. Press, 1999.
- . « Costs and Constraints in the Economy of the Mind », in Brocas, I. ; Carrillo, J. D. (dir), *The Psychology of Economic Decisions*, vol. 2, *Reasons and Choices*, New York, Oxford Univ. Press, 2004, p. 3-14.
- Etzioni, Amitai. « The Case for a Multiple-Utility Conception », *Economics and Philosophy*, vol. 2, 1986, p. 159-183.
- Etzioni, Amitai. *The Moral Dimension*, New York, Free Press, 1988.
- Ferejohn, John. « Rational Choice Theory and Social Explanation », *Economics and Philosophy*, vol. 18, 2002, p. 211-234.
- Follesdal, Dagfinn. « The Status of Rationality Assumptions in Interpretation and in the Explanation of Action », *Dialectica*, vol. 36, 1982, p. 301-316.
- Frank, Robert H. « Social Norms as Positional Arms Control Agreements », in Ben-Ner, A. ; Putterman, L. (dir), *Economics, Values, and Organization*, New York, Cambridge Univ. Press, 1998, p. 275-295.
- Gagné, Leary. « La délibération circonstancielle en théorie démocratique », *Philosophiques*, vol. 29, 2002, p. 327-350.
- Goldthorpe, John H. *On Sociology*, New York, Oxford Univ. Press, 2000.
- Goodin, Robert E. « Laundering Preferences », in Elster, J. ; Hylland A. (dir), *Foundations of Social Choice Theory*, New York, Cambridge Univ. Press, 1986, p. 75-102.
- . « Institutions and Their Design », in Goodin, R.E. (dir), *The Theory of Institutional Design*, New York, Cambridge Univ. Press, 1996, p. 1-53.
- Hardin, Russel. *One for All : The Logic of Group Conflict*, Princeton Univ. Press, 1995.
- Hausman, Daniel M. ; McPherson, Michael S. « Taking Ethics Seriously », *Journal of Economic Literature*, vol. 31, 1993, p. 671-731.
- . *The Inexact and Separate Science of Economics*, Cambridge (UK), Cambridge Univ. Press, 1992.
- Heath, Joseph. « Rational Choice with Deontic Constraints », *Canadian Journal of Philosophy*, vol. 31, 2001, p. 361-388.

- Höllander, Heinz. « A Social Exchange Approach to Voluntary Cooperation », *The American Economic Review*, vol. 80, 1990, p. 1157-1167.
- Isaac, Alan G. « Morality, Maximization, and Economic Behavior », *Southern Economic Journal*, vol. 63, 1997, p. 559-570.
- Kuran, Timur. « Moral Overload and its Alleviation », in Ben-Ner, A. ; Putterman, L. (dir), *Economics, Values, and Organization*, New York, Cambridge Univ. Press, p. 231-266, 1998.
- Levitt, Steven D. ; Dubner, Stephen J. *Freakonomics*, New York, William Morrow, 2005.
- Nee, Victor ; Ingram, Paul. « Embeddedness and Beyond : Institutions, Exchange, and Social Structure », in Nee, V. ; Brinton, M. C. (dir), *The New Institutionalism in Sociology*, New York, Russel Sage Foundation, 1998, p. 19-45.
- Ostrom, Elinor. « An Agenda for the Study of Institutions », *Public Choice*, vol. 48, 1986, p. 3-25.
- Pettit, Philip. « *Virtus Normativa* : Rational Perspectives », *Ethics*, vol. 100, 1990, p. 725-755.
- . « The Virtual Reality of *Homo Economicus* », *The Monist*, vol. 78, 1995, p. 308-329.
- Riker, William H. « Political Trust as Rational Choice », in Lewin, L. ; Vedung, E. (dir), *Politics as Rational Action*, Dordrecht, D. Reidel, 1980.
- Satz, Debra ; Ferejohn, John. « Rational Choice and Social Theory », *Journal of Philosophy*, vol. 91, 1994, p. 71-87.
- Sen, Amartya. « Rational Fools », in Sen, A. (dir), *Choice, Welfare, and Measurement*, Cambridge (MA), MIT Press, 1982, p. 84-106.
- Shepsle, Kenneth A. « Studying Institutions : Some Lessons From the Rational Choice Approach », *Journal of Theoretical Politics*, vol. 1, n° 2, 1989, p. 131-47.
- Weibull, Jörgen W. « Evolution, Rationality and Equilibrium in Games », *European Economic Review*, vol. 42, 1998, p. 641-649.
- Wolfelsperger, Alain. « La modélisation économique de la rationalité axiologique. Des sentiments moraux aux mécanismes sociaux de la moralité », in Boudon, R. ; Demeulenaere, P. ; Viale, R. (dir), *L'explication des normes sociales*, Paris, PUF, 2001, p. 63-92.

Précis de *Reconstructing Reason and Representation*

Murray Clarke

Volume 34, numéro 2, automne 2007

URI : <https://id.erudit.org/iderudit/016997ar>

DOI : <https://doi.org/10.7202/016997ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Société de philosophie du Québec

ISSN

0316-2923 (imprimé)

1492-1391 (numérique)

[Découvrir la revue](#)

Citer ce document

Clarke, M. (2007). Précis de *Reconstructing Reason and Representation*. *Philosophiques*, 34 (2), 353–362. <https://doi.org/10.7202/016997ar>

Précis de *Reconstructing Reason and Representation*

MURRAY CLARKE
Université Concordia

J'avais deux objectifs à l'esprit quand j'ai écrit *Reconstructing Reason and Representation*. Premièrement, j'ai voulu clarifier et évaluer les mérites empiriques et conceptuels de la psychologie évolutionniste. Deuxièmement, j'ai voulu évaluer les implications de la psychologie évolutionniste pour certains thèmes de l'épistémologie, la philosophie des sciences et la philosophie de l'esprit. Je voulais offrir, durant le processus, une nouvelle approche de l'épistémologie naturaliste. L'idée était que nous devions permettre à l'épistémologie de devenir modulaire et la percevoir comme un ensemble de genres naturels. Ma conviction était qu'après avoir fait ce changement nous pourrions libérer l'épistémologie de la stagnation scolastique de l'analyse conceptuelle. J'ai essayé de poser les bases pour ce projet en élucidant les idées centrales qui constituent la version de la psychologie évolutionniste de Leda Cosmides et John Tooby comme suit. Cosmides et Tooby ont adopté l'affirmation méthodologique de la cohérence trans-théorique ou, comme l'appelle Jerry Fodor : « la pertinence explicative réciproque ». Par exemple, la biologie de l'évolution et la psychologie cognitive se contraignent mutuellement pour l'acceptabilité des théories que chacune peut intégrer. Elles souscrivent aussi à l'idée de nature humaine universelle quant aux mécanismes psychologiques complexes, ou modules darwiniens, et non quant à l'expression des comportements culturels. Deuxièmement, ces mécanismes psychologiques complexes sont des adaptations qui ont été élaborées par la sélection naturelle au cours d'une certaine période évolutionnaire. Par conséquent, les modules darwiniens sont des structures cognitives innées dont les propriétés principales sont largement déterminées par des facteurs génétiques. Une troisième idée est que la structure complexe de l'esprit humain est adaptée au mode de vie des chasseurs-cueilleurs du Pléistocène, et pas nécessairement à nos circonstances modernes. Ce que nous pensons de l'histoire humaine récente, c.-à-d. les deux derniers millénaires, n'a pas grand-chose à voir avec la formation de l'esprit humain, parce que nos esprits ont été largement formés par les deux derniers millions d'années comme chasseurs-cueilleurs du Pléistocène. Une quatrième idée est que les modules darwiniens ont un domaine spécifique. Selon Cosmides et Tooby, nos esprits consistent largement en « une constellation de mécanismes spécifiques possédant des procédures propres à leur domaine spécifique, qui agissent sur les représentations d'un domaine spécifique, ou les deux » (Cosmides et Tooby, 1994, p. 94). Cela signifie qu'une structure cognitive innée qui est spécifique ne répondra seulement qu'à un *input* représentationnel d'un type particulier. La cinquième idée, c'est que les modules darwiniens sont des mécanismes computationnels. Comme l'affirment Cosmides et Tooby :

« notre architecture ressemble à une confédération de centaines de milliers d'ordinateurs fonctionnellement dédiés (souvent appelés modules) [...] » (Tooby et Cosmides, 1995, p. xiii). On doit à Turing l'idée qui sous-tend cette conception. C'est lui qui introduisit pour la première fois l'idée que les processus mentaux sont des computations. De tels dispositifs computationnels sont des ordinateurs classiques. Par conséquent, les processus mentaux cognitifs sont des opérations formelles définies par des représentations mentales structurées syntaxiquement qui sont similaires à des phrases. Une computation est un processus causal qui est conduit syntaxiquement. Finalement, certaines motivations menant à croire à l'existence des modules darwiniens viennent de l'idée de Chomsky selon laquelle l'argument de la pauvreté du stimulus détermine l'information qu'un esprit doit avoir de façon innée. Il faut soustraire l'information provenant de l'environnement de l'information requise à un enfant pour atteindre la maîtrise de la langue. Ce qui en résulte est ce à quoi contribuent les connaissances innées de l'enfant à la procédure d'acquisition du langage. Pour les psychologues évolutionnistes, ce qui en résulte et qui est inné est considérable : un esprit qui est largement, mais pas complètement, constitué par des centaines de milliers d'ordinateurs fonctionnellement dédiés. Les modules darwiniens, en somme, qui sont innés, sélectionnés par la nature, fonctionnellement spécifiques, universels, constituent des mécanismes computationnels à la Turing, lesquels, souvent, s'activent de concert avec des ensembles de données spécifiques ou des représentations. Appelons cela « l'Hypothèse de la modularité massive » ou HMM. Cosmides et Tooby ne se sont pas prononcés sur la question de la connexion entre les processeurs computationnels qui ont un domaine spécifique et les ensembles d'informations qui ont des domaines spécifiques. Il semble naturel de joindre les processeurs darwiniens qui ont un domaine spécifique avec les ensembles chomskyens d'information qui ont un domaine spécifique. Appelons cela les modules darwiniens/chomskyens (MDC). J'ai eu de la difficulté à faire ressortir que Cosmides et Tooby n'endossent pas cette position. C'est-à-dire qu'ils pensent qu'il existe des MDC mais qu'il doit bien aussi y avoir d'autres aspects de notre architecture cognitive.

Cosmides et Tooby, en fait, ne s'engagent pas envers la notion selon laquelle l'esprit doit être entièrement modulaire. Pour une première approximation, il faut penser aux possibilités selon quatre choix :

	Processeurs ayant un domaine spécifique (Modules darwiniens)	Processeurs ayant un domaine général
Ensemble de données ayant un domaine spécifique (Modules chomskyens)	A	B
Ensemble de données ayant un domaine général	C	D

N.B. : le choix A, par exemple, représente la conjonction de processeurs ayant un domaine spécifique, ou un module darwinien, avec un ensemble de données ayant un domaine spécifique, ou un module chomskyen.

Cosmides et Tooby suggèrent que les choix A, B, C et D sont tous des options viables et que ces quatre options ne sont pas mutuellement exclusives (Cosmides et Tooby, 1994, p. 104). Voici ce qu'ils disent à leur propos : « Chacune de ces possibilités peut être correcte. En effet, toutes peuvent être correctes, bien que dans différents domaines » (Cosmides et Tooby, 1994, p. 104). Ils laissent entrevoir la possibilité que ces options soient vraies, mais dans différents domaines. Pour Cosmides et Tooby, l'esprit n'a pas à être seulement modulaire (ou seulement massivement modulaire). L'esprit, dans leur perspective, peut aussi être non modulaire en certains aspects. Cosmides et Tooby acceptent la représentation massivement modulaire et le modèle processuaire de la cognition qui pose que :

Le principe RMMP

L'esprit est largement composé d'une vaste assemblée de modules darwiniens/chomskyens. Avertissement : certains choix B, C ou D doivent être des composants de l'esprit.

Ni Cosmides et Tooby ni moi ne souscrivent à l'idée d'un esprit entièrement modulaire. En même temps, il faut souligner que Cosmides et Tooby souscrivent certainement à la position selon laquelle les modules darwiniens/chomskyens dominent l'architecture mentale.

Une partie significative du travail de Cosmides et Tooby en psychologie évolutionniste a été dédiée à l'étude du raisonnement humain. En particulier, les effets liés au contenu lors de la tâche de sélection de Wason (1968), ainsi que Wason et Philip Johnson-Laird (1972), ont démontré que l'exécution du raisonnement, quant à différentes tâches qui requièrent l'usage d'une seule règle d'inférence déductive, varie en fonction du contenu en jeu dans les règles d'inférences. Cela viole l'idée la plus fondamentale de la logique formelle, c'est-à-dire que les arguments sont valides purement en fonction de leur forme abstraite, indépendamment de leur contenu. Que les humains échouent systématiquement à observer l'aspect de la neutralité du contenu quant aux tâches de raisonnements déductifs est apparu comme une surprise considérable. Par exemple, si l'on considère la tâche de sélection de Wason suivante (Johnson-Laird, 1983, p. 30). Un expérimentateur dépose quatre cartes en face d'un sujet avec les symboles suivants :

E K 4 7

Le sujet sait que chaque carte possède un chiffre sur un côté et une lettre sur l'autre. L'expérimentateur présente ensuite au sujet les généralisations suivantes : si une carte possède une voyelle sur un côté alors il y a un chiffre pair sur l'autre côté.

La tâche du sujet est de retourner seulement les cartes qui ont besoin d'être observées pour vérifier si la généralisation est vraie ou fausse. L'ordre

dans lequel les cartes sont retournées n'est pas un problème. Cette tâche apparemment simple s'est révélée être difficile à résoudre. La plupart des gens conçoivent que la carte avec la voyelle doit être retournée. La généralisation demeure valide si, une fois tournée, le chiffre est pair. Si la carte représente un chiffre impair, alors la généralisation doit être fausse. De façon similaire, la plupart des sujets réalisent que la carte avec une consonne ne doit pas être touchée puisque la règle ne dit rien à propos des consonnes. Certains sujets tournent la carte avec le chiffre pair, d'autres non. Mais la carte avec un chiffre pair n'a pas besoin d'être tournée puisque, qu'il y ait une voyelle ou non sur l'autre côté, la conditionnelle sera vraie.

Le problème central concerne la carte avec le chiffre impair. Cette carte doit être tournée puisque, si elle contient une voyelle, alors la généralisation est fausse. Mais peu de sujets insistent sur le fait de la retourner quoique la raison de la vérifier soit exactement identique à la raison poussant à vérifier la carte avec la voyelle : la généralisation est clairement fausse lorsqu'une voyelle et un chiffre impair sont sur les deux côtés de la carte. Wason et Johnson-Laird ont tenté de nombreuses variantes quant à la procédure et au matériel dans le but d'améliorer les performances jusqu'à ce qu'ils trouvent une modification simple ayant des effets frappants (Johnson-Laird, 1983, p. 31). Lorsque les sujets se font présenter quatre cartes représentant des voyages, c.-à-d., avec une destination sur un côté et un mode de transport de l'autre, les résultats se révèlent notablement accrus. Donc, avec les cartes Manchester, Sheffield, Train et Automobile ainsi que la règle générale :

À chaque fois que je vais à Manchester, je voyage en train,

plus de 60 % des sujets comprennent qu'ils doivent retourner la carte indiquant « Automobile ». Cela représente un net contraste avec l'exemple précédent où presque personne n'a compris — dans le cas parallèle où les chiffres impairs devaient être retournés. Et dans le groupe contrôle, à peine plus de 12 % des sujets ont fait le choix équivalent lorsque du matériel abstrait était en jeu. Ce qui suggère que les matériaux réalistes ou familiers produisent de bien meilleurs résultats que les matériaux abstraits ou peu familiers, indépendamment du fait que des expériences distinctes emploient des généralisations avec la même forme logique et les mêmes conditions de vérité. De nombreuses autres variations sur ces expériences ont été effectuées qui semblent confirmer ces résultats.

Toutefois, c'est le travail révolutionnaire de Cosmides et Tooby expliquant les effets liés au contenu lors des tâches de sélection de Wason qui, en partie, a préparé le terrain pour le projet plus général de la psychologie évolutionniste. Cosmides et Tooby expliquaient les effets liés au contenu en invoquant la présence ou l'absence d'un « contrat social » dans la tâche de sélection et non pas la notion associationniste de « familiarité » de Wason et Johnson-Laird. Mais l'idée d'un esprit inné et massivement modulaire se trouvait clairement

en arrière-plan. La stratégie de « l'ingénierie à rebours » a guidé la méthodologie de ces études. « L'ingénierie à rebours » se rapporte à une stratégie de design expérimental où l'on doit tenter de déterminer les problèmes adaptatifs auxquels les chasseurs-cueilleurs du Pléistocène ont fait face, et, alors, désigner des expériences qui auraient fait ressortir la perspicacité des adaptations fonctionnelles qui sont apparues en réponse à ces problèmes adaptatifs. Il faut comprendre les caractéristiques des conceptions causées par la sélection naturelle afin de clarifier de si complexes adaptations fonctionnelles. Comme le notent Cosmides et Tooby en ce qui concerne le rôle de la chance dans l'évolution :

Des allées et venues aléatoires ne construisent pas systématiquement des arrangements fonctionnels improbables et complexes comme le système visuel, la faculté de langage ou le contrôle moteur. La seule explication connue pour l'existence de conceptions fonctionnelles complexes dans les systèmes organiques est la sélection naturelle (Cosmides et Tooby, 1994, p. 86).

Ce point vaut aussi bien pour la capacité à raisonner chez les humains. En tant que telle, la méthodologie de l'ingénierie à rebours a guidé les études de Cosmides et Tooby sur la tâche de sélection de Wason.

Pour les philosophes, le résultat peut-être le plus stupéfiant de la psychologie empirique des trois dernières décennies a été la déconstruction de la notion qu'il existe une psycho-logique russellienne sous-jacente qui guide l'esprit et le langage. En suivant scrupuleusement les standards de la logique, les humains apparaissent être sauvagement irrationnels dans le sens où il n'y a pas de systèmes logiques employés par les humains qui préservent la vérité, la neutralité quant au contenu et la généralité à travers les domaines d'application. Nous faisons des erreurs inductives et déductives qui sont répandues et influentes. Les résultats expérimentaux soutenant ces affirmations sont robustes et reproductibles. Nous commettons le sophisme de la conjonction, nous sommes coupables de la négligence du taux de base et de l'excès de confiance. Dans les chapitres quatre et cinq, je passe en revue certaines données empiriques concernant ces études. Comme nous l'avons noté auparavant, Cosmides et Tooby ont essayé de renverser les interprétations des théories associationnistes de la disponibilité [*associationism-based availability theorists*], comme celles de Wason, Kahneman et Tversky, Nisbett et Ross, et d'autres, concernant les effets liés au contenu dans les tâches de sélection de Wason. Les associationnistes argumentent que la familiarité avec les données (ou l'expérience différentielle) explique le fait que les sujets vont raisonner en accord avec, par exemple, le *modus tollens* dans certains cas, mais violeront le *modus tollens* dans d'autres. La supposition des associationnistes était que les humains possèdent une capacité de raisonner qui a un domaine général, et indépendante du contenu. Plus tard, les théoriciens du raisonnement pratique, tels que Cheng et Holyoak, ont argumenté, dans le même esprit, en faveur d'une capacité de raisonner qui aurait un domaine général, mais ils suggèrent que les humains emploient des « schémas de raisonnement pratique » qui ont

été induits par l'expérience récurrente à l'intérieur de domaines ayant des buts définis. Les schémas étaient envisagés comme étant dépendants du contexte, même si les processus cognitifs inductifs engendrant ces schémas étaient pensés comme étant indépendants du contenu. Les expériences différentielles expliquent quels schémas étaient construits et pourquoi certains autres ne l'étaient pas. Toutefois, Cosmides a tenté de montrer que ni les théories associationnistes de la disponibilité ni les théories du raisonnement pratique basées sur l'induction ne peuvent expliquer les effets liés au contenu dans la tâche de sélection de Wason. Elle argumente que la présence d'un contrat social intégré dans la tâche de sélection de Wason explique les effets liés au contenu. S'il existe un contrat social en jeu dans cette tâche, même s'il n'est pas familier aux sujets, alors leurs résultats seront meilleurs. La prédiction de Cosmides a été confirmée par ses expériences. Visiblement, ces résultats empiriques défient la vérité des théories de la disponibilité.

Cheng et Holyoak affirment que les schémas de raisonnements pratiques peuvent expliquer les effets liés au contenu lors de la tâche de sélection de Wason dans le sens où lorsqu'un schéma est actif, les sujets raisonnent en accord avec le calcul des propositions. Inversement, là où aucun schéma n'est visible, les sujets échouent à raisonner en accord avec le calcul des propositions. Cosmides a noté que les règles de contrat social impliquent des règles de permission (ou schémas), mais les règles de permission n'impliquent pas toutes un contrat social. Il en est ainsi à cause de la déclaration du contrat social : « Si le bénéfice revient à un seul, alors il en doit en payer le prix » entraîne la règle de permission « Si quelqu'un entreprend l'action A, alors il doit satisfaire la condition P ». Mais l'inverse ne tient pas. Tous les bénéfices acquis sont des actions qui ont été prises, mais toutes les actions entreprises ne sont pas des bénéfices acquis. Comme Cosmides l'a noté : « Une règle de permission est aussi une règle de contrat social seulement lorsque le sujet interprète "l'action à entreprendre" comme un bénéfice rationné, et la "condition à satisfaire" comme un coût requis » (Cosmides, 1989, p. 237). Il en résulte que le domaine des schémas de permission est plus large que les algorithmes des contrats sociaux.

Ultérieurement, Gigerenzer et Hug ont reproduit les résultats de Cosmides mais ont offert un amendement amical à ces résultats. Ils s'entendent sur le fait que les associationnismes familiers et les hypothèses concernant les schémas de raisonnements pratiques sont faux (ou semblent faux). Mais, ils ont démontré que, lorsque les contrats sociaux qui impliquent des algorithmes de détection de tricheur sont en jeu, les effets liés au contenu apparaissent. En revanche, lorsque seuls les contrats sociaux sont en jeu et aucun algorithme de détection des tricheurs, aucun effet lié au contenu n'est visible. Ensuite, Cosmides et Tooby ont argumenté que les humains pouvaient être de bien meilleurs raisonneurs intuitivement et inductivement qu'ils ne peuvent être de bons raisonneurs intuitivement et déductivement. Une série d'études de Cosmides et Tooby où la négligence du taux de base est en jeu ont montré que si les sujets se font présenter les problèmes comme impliquant des fréquences

relatives, plutôt que des probabilités appliquées au cas par cas, les résultats sont considérablement meilleurs. Dans le même ordre d'idée, d'autres chercheurs ont obtenu des résultats concordants quant à la fréquence relative appliquée à la négligence du taux de base, le sophisme de la conjonction et le biais de l'excès de confiance. Ces données peignent une image bien plus positive des capacités humaines à inférer inductivement que celle présentée par les théoriciens du raisonnement associationnistes et pragmatiques en 1970 et 1980, ou même par Cosmides et Tooby dans les années 1980. Cependant, le statut du raisonnement déductif reste plutôt faible par rapport aux stricts standards logiques. Avec le rapide survol de la position de Cosmides et Tooby concernant cette partie de la psychologie évolutionniste qui s'occupe du raisonnement humain en question, j'ai essayé de recenser ce que je crois être les implications philosophiques de ces résultats.

Dans le deuxième chapitre, je m'occupe des inquiétudes fondamentales que Jerry Fodor a exprimées contre le projet de la modularité massive. Dans son livre, *The Mind Doesn't Work That Way*, Fodor argumente que la Nouvelle Synthèse, c.-à-d. la psychologie évolutionniste de Cosmides et Tooby, ou de Pinker et Plotkin, était vraiment mal avisée. Fodor pense que ce qu'il appelle le « problème de l'*input* », un problème a priori, immobilise la Nouvelle Synthèse dans son cheminement. Il offre aussi plusieurs objections à celle-ci. Il affirme en particulier que la psychologie évolutionniste ne peut intégrer les inférences globales et abductives dans le contexte de processeurs locaux et computationnels. J'ai soutenu qu'il interprète mal la position de Cosmides et Tooby et charge donc, à maintes reprises, un homme de paille. De plus, j'ai soutenu que l'hypothèse de la modularité massive peut être étendue de façon à expliquer ce que j'appelle « l'abduction évolutionniste non explicite ». Évaluer la critique de Fodor était important parce qu'il nous permet de soulever des préoccupations que de nombreux cognitivistes peuvent entretenir à propos de la psychologie évolutionniste et parce que cela nous permettra une compréhension plus fine de la façon dont Cosmides et Tooby conçoivent leur propre projet. Il existe en particulier une rumeur qui circule selon laquelle Cosmides et Tooby souscrivent à la notion d'esprit complètement modulaire. J'espère avoir fait disparaître cette idée complètement erronée. Dans le troisième chapitre, j'ai prospecté des solutions concernant les erreurs de représentation afin de résoudre le problème de la disjonction. La théorie causale rudimentaire de la représentation pose qu'une instanciation de « D » est causée, de façon fiable, par D. Cela établit le fait que les conditions gouvernant ce que signifie pour D d'être représenté par « D » est identique à la condition qu'un tel particulier soit vrai. Il est impossible, comme tel, d'introduire la fausseté dans cette perspective. Il est possible de penser que D cause des instanciations de « D » est vrai et que E cause des instanciations de « D » est faux. Mais cela ne marchera pas. Les « D » sont causés de manière fiable par les propriétés disjonctives d'être (D ou E). Comme telles, les instances de E-cause- « D » sont vraies parce qu'elles sont causées de façon fiable par

(D ou E), et nous n'avons pas de théorie de l'erreur de représentation. Voilà le problème de la disjonction. Je soutiens qu'il est possible de le résoudre en utilisant les ressources de l'hypothèse de la modularité massive de Cosmides et Tooby, combinée avec une contribution étiologique et fiabiliste de la fonction propre des processus computationnels qui constituent de tels modules. Le résultat est une solution aux erreurs de représentations entièrement applicable aux parties de l'esprit qui sont modulaires. Dans les grandes lignes, si le fossé entre le domaine propre (dans lequel le module a été sélectionné) et le domaine actuel (dans lequel nous existons actuellement) est assez large, alors la fonction propre du module faillira. Les erreurs de représentation surviennent quand le fossé entre le domaine propre et le domaine actuel est significatif et résulte en un mauvais fonctionnement des modules. De plus, les types de croyances erronées étaient endémiques chez nos ancêtres si une mauvaise identification des problèmes adaptatifs significatifs en résultait. Les erreurs de représentation sont une partie du processus qui mène au mauvais fonctionnement des modules. Les croyances erronées concernant les problèmes adaptatifs déclenchent des réponses inappropriées des modules. Les erreurs de représentation déclenchent un mauvais fonctionnement, et ce mauvais fonctionnement résulte en une erreur biologique. Inversement, pour nos ancêtres, les types de croyances vraies étaient importants à l'identification des problèmes adaptatifs. Cette contribution ne surestime pas les capacités représentationnelles des modules tout en réussissant à biologiser la contribution de Fodor à la dépendance asymétrique des erreurs de représentations. Je me suis débarrassé de la notion métaphysique d'une « *representation consumer* » tout en préservant les avantages principaux de la contribution de Millikan quant aux fonctions propres. Cette contribution montre que la signification importe parce que la vérité et la fausseté importent. Dans le chapitre quatre, j'ai soutenu que des représentations lexicales précises ont été cruciales pour la survivance et la reproduction d'homo sapiens sapiens. J'ai soutenu spécifiquement que des processus fiables ont été sélectionnés en raison de leur connexion indirecte, mais étroite, avec les croyances vraies durant la période chasseur-cueilleur du Pléistocène de notre histoire ancestrale. Les croyances vraies ne sont pas héréditaires, les processus fiables le sont. Ces processus fiables qui sont connectés avec le raisonnement prennent la forme d'algorithmes darwiniens : une pléthore de règles inférentielles spécialisées et spécifiques à un domaine qui sont constituées pour résoudre des problèmes spécifiques, récurrents et adaptatifs dans les contextes d'échanges sociaux. Les humains, ai-je soutenu, ne raisonnent pas logiquement mais adaptativement. Dans le chapitre cinq, j'ai d'abord souligné que la conception fiabiliste et externaliste de la justification épistémique et de la connaissance fait face à deux groupes de critiques : les épistémologues analytiques et les philosophes des sciences. Le premier groupe reproche aux épistémologues naturalistes de « détourner le sujet » et donc de manquer les débats classiques de l'épistémologie traditionnelle et autres problèmes importants. Le second groupe, ironiquement, reproche aux épistémologues

naturalistes de « ne pas parvenir à changer de sujet », à tel point que leurs travaux sont connectés à l'épistémologie analytique traditionnelle, une forme de recherche dépassée et utopique. Les philosophes des sciences pensent que l'épistémologie traditionnelle est utopique parce que c'est une tentative de réponse aux standards déraisonnablement élevés que les sceptiques souhaitent imposer à toute contribution adéquate de la connaissance. Selon les philosophes des sciences, le secret est de rejeter ces standards sceptiques déraisonnablement élevés. Il est clair qu'aucun des deux groupes ne peut avoir raison. J'ai endossé une position en faveur du projet naturaliste en dessinant d'abord une distinction entre les sens mélioratif et non mélioratif de la justification. J'ai utilisé cette distinction pour soutenir que les deux groupes de critiques ont manqué la cible du projet naturaliste mais pour différentes raisons. Ensuite, j'ai passé quelque peu en revue la littérature sur la rationalité humaine du point de vue de la psychologie empirique et suggéré comment une épistémologie naturalisée, qui prend au sérieux ces résultats empiriques et qui est sensible à la distinction mélioratif/non mélioratif, peut être développée. Tout au long de mon cheminement, j'ai essayé de reconnecter l'épistémologie analytique avec la philosophie des sciences en montrant comment les projets naturalistes sont liés dans les deux camps. Parce qu'il me semble qu'il est grand temps que les philosophes reconnaissent que l'épistémologie standard du sujet connaissant est en continuité avec l'épistémologie du scientifique. Quine, selon moi, avait raison : la science est seulement une sophistication du sens commun. Les deux voies de la philosophie analytique, symbolisée par l'analyse informelle mooréenne d'un côté, et l'analyse formelle russellienne de l'autre, ont finalement besoin d'être jointes encore une fois. La façon de le faire, je pense, est de montrer comment une épistémologie naturalisée peut informer et être sensible à une philosophie naturaliste des sciences. Les deux solitudes doivent se fondre l'une dans l'autre. Le chapitre six représente le commencement, à partir des deux chapitres précédents, en ce que cette discussion est directement centrée sur la littérature de la modularité massive et non sur une théorie de la rationalité. Il débute avec une discussion de l'objection, par les non-naturalistes, que la science n'a pas de place dans les problèmes épistémiques. Des non-naturalistes, comme Richard Feldman et Richard Fumerton, ont soutenu que les débats épistémiques sont normatifs ou évaluatifs par nature. Si tel est le cas, il doit être inutile d'employer des ressources factuelles ou descriptives de la science pour reconstruire des notions normatives ou épistémiques comme la justification ou la connaissance. J'ai soutenu que les non-naturalistes font simplement fausse route en cette matière puisque la connaissance est en fait un ensemble de genres naturels qui devraient être étudiés empiriquement, et non un genre conceptuel à étudier par un appel a priori de la raison. J'ai proposé l'argument de la connaissance comme genre naturel (ACGN) pour soutenir ma perspective. En me positionnant de cette façon, je ne me distingue pas seulement du principe central de l'épistémologie analytique, je me distingue aussi de la forme d'analyse de la justification et de la connaissance adoptée dans

les chapitres quatre et cinq. J'argumenterai pourtant que même sur le terrain ouvert par l'analyse conceptuelle ma perspective est supérieure aux conceptions internalistes de la connaissance et aux autres conceptions externalistes de la connaissance. Par le fait de voir la connaissance comme un ensemble de genres naturels, la mission d'une épistémologie complètement naturalisée, pour la première fois dans l'histoire de l'épistémologie, est-elle finalement offerte ? De ma perspective, il s'agit d'une transformation méthodologique nouvelle et importante qui prépare le terrain à ma contribution modulaire de la connaissance. En somme, ma contribution concernant la connaissance dépend des idées jumelles que la connaissance est un ensemble de genres naturels (et non un genre conceptuel) et qu'une telle connaissance se loge à l'intérieur d'une vaste assemblée de sous-ensembles appropriés de modules RMMP. La connaissance empirique est le résultat des modules darwiniens qui convertissent les inputs sensoriels, si bien que la connaissance a priori est le résultat des modules darwiniens/chomskyens. Une telle connaissance modulaire est instrumentalement importante puisqu'elle résulte, une fois attachée correctement à nos désirs, en la satisfaction de nos besoins biologiques et d'autres buts que nous avons. Il en résulte que la connaissance est composée d'un ensemble de genres naturels logeant dans des modules distincts de l'esprit. Steve Stich a argumenté en faveur du relativisme épistémique dans son classique de 1990, *The Fragmentation of Reason*. Ce que je crois, c'est qu'il existe une bonne raison de penser que la littérature la plus récente sur la rationalité et la modularité massive se présente à nous avec des assises pour une image différente de la raison et de la représentation. La nouvelle image est, je crois, décidément plus optimiste que celle présentée par Stich. Ma conception modulaire de la connaissance est une espèce de fondationnalisme, bien qu'elle ne conçoive pas la connaissance comme un genre conceptuel mais comme un ensemble de genres naturels. Comme telle, ma perspective de la connaissance et de la justification est incompatible avec le relativisme épistémique, même si elle soutient une fragmentation de la connaissance. Je me distingue d'une autre façon du relativisme de Stich en pensant qu'il existe des assises évolutionnistes pour penser que la vérité a contribué et contribue encore à notre adaptativité en tant qu'espèce.

Améliorer les théories de la fiabilité

Robert Hudson

Volume 34, numéro 2, automne 2007

URI : <https://id.erudit.org/iderudit/016999ar>

DOI : <https://doi.org/10.7202/016999ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Société de philosophie du Québec

ISSN

0316-2923 (imprimé)

1492-1391 (numérique)

[Découvrir la revue](#)

Citer ce document

Hudson, R. (2007). Améliorer les théories de la fiabilité. *Philosophiques*, 34 (2), 363–366. <https://doi.org/10.7202/016999ar>

Améliorer les théories de la fiabilité

ROBERT HUDSON

Université de Saskatchewan

Dans *Reconstructing Reason and Representation*, Murray Clarke explore la pertinence de la recherche empirique pour la philosophie de l'esprit et l'épistémologie. Mon intérêt porte, dans ce commentaire, sur la solution naturaliste qu'il cherche à fournir pour un programme particulièrement dense en faveur d'une épistémologie fiabiliste, à savoir, le problème de la généralité. Le cœur de sa solution se trouve dans son déploiement d'une « théorie de la modularité massive » qui décrit l'esprit comme étant composé d'un ensemble étendu de « processeurs computationnels spécialisés, ou modules darwiniens » (p. 6 ; toutes les références se rapportent au livre de Clarke) forgés dans l'histoire évolutive de l'*homo sapiens*. Clarke pense que sa théorie de la modularité procure une solution effective au problème de la généralité. Je rejeterai cette affirmation et offrirai une solution de rechange qui satisfait mieux un des objectifs avoués de Clarke dans ce livre — incorporer une perspective externaliste (dans les termes de Clarke : non méliorative) à une perspective internaliste (dans les termes de Clarke : méliorative) aux théories de la justification épistémique.

Clarke est, sans équivoque, un fiabiliste quant à la justification épistémique et il appuie clairement son fiabilisme sur un naturalisme, fondé sur l'évolution : pour l'essentiel, les organismes qui survivent et vivent pour transmettre leur bagage génétique aux générations futures sont celles qui entretiennent des croyances vraies sur le monde ; et la façon de garantir qu'on entretient des croyances vraies sur le monde est de posséder des mécanismes cognitifs qui aboutissent à des croyances vraies, c'est-à-dire des mécanismes cognitifs qui sont épistémologiquement fiables (voir p. 88). À cet égard, il offre une histoire naturaliste en désaccord avec les histoires présentées par Ruth Millikan (p. 65-66) et Stephen Stich (p. 77), qui affirment tous deux que la production fiable de croyances vraies n'est pas essentielle à la survivance (en bref, posséder des croyances faussement positives peut être plus adaptatif). Le débat opposant Clarke à Millikan et Stich sur ce point n'a pas à être clarifié ici. Pour les besoins de l'argumentation, nous accordons à Clarke l'avantage évolutionnaire de la fiabilité, des processus engendrant les croyances, et nous nous concentrerons sur son autre affirmation selon laquelle lorsque ces processus impliquent des raisonnements, ceux-ci doivent être implantés dans des modules darwiniens spécifiques. L'existence de tels modules a été originellement montrée, selon Clarke, par un ensemble d'expériences sur la tâche de sélection de Wason faites par Leda Cosmides et John Tooby. En supposant toujours, pour les besoins de l'argumentation, que de tels modules existent et qu'ils permettent la fiabilité de la cognition, l'affirmation de Clarke qui nous importe est que l'existence de tels

modules résout l'épineux problème de la généralité qui empoisonne toutes les théories fiabilistes de la justification.

Essentiellement, le problème de la généralité se présente comme suit : il n'est pas évident de savoir à quel niveau de description nous devrions caractériser les processus (de formation de croyances) afin de parvenir à déterminer s'ils sont fiables ou non. Si nous décrivons ces processus trop précisément, disons, avec seulement une instance (pour l'esprit, le particulier en question), alors tout processus qui culmine en une croyance vraie sera considéré comme fiable. De l'autre côté, si nous caractérisons un processus trop largement, alors nous pourrions manquer le détail des informations cruciales au processus en question — qui auraient pu, potentiellement, nous fournir une évaluation plus précise de sa fiabilité. Pour l'internalisme épistémologique, l'intractabilité du problème de la généralité annonce l'incohérence fondamentale des théories fiabilistes, et, par conséquent, il est important que les fiabilistes aient une réponse. La réponse de Clarke est la suivante :

Ma perspective ne souffre pas du problème de la généralité puisque la méthode utilisée pour déterminer la fiabilité modulaire ne repose pas sur des types qui sont pris a priori. Les types de modules sont plutôt découverts empiriquement. Nous ne créons pas les modules, nous les découvrons. La fiabilité de tels types de modules est expliquée par le fait qu'ils ont été sélectionnés précisément parce qu'ils étaient fiables dans l'EAE [environnement de l'adaptation évolutive]. Dans le EAE, ces modules sont une réussite parce qu'ils sont fiables. Par conséquent, nous ne sommes pas en position de devoir fournir une explication des types de modules qui soit trop étroite ou trop large. La sélection naturelle a fait le travail à notre place. Tout ce que nous avons à faire est de découvrir les modules en question. Il n'existe tout simplement pas de problème de la généralité pour déterminer les types de modules (p. 65).

Cette réponse aurait été excellente s'il y avait eu une voie empirique toute tracée pour isoler les modules — mais il n'y en a pas. Pour identifier les modules, il faut identifier les comportements propices à l'évolution qui sont observables chez les organismes et inférer alors l'existence d'un module qui permet, de par sa structure, ce comportement. À cet égard, Clarke cite l'exemple des castors qui claquent leur queue à la surface de l'eau pour avertir d'un danger (p. 60-61 et p. 66). Millikan note la tendance des castors à faire des claquements faussement positifs, et incrimine donc la fiabilité de cette tendance. Ici, Clarke affirme la stupidité évolutionnaire de tels claquements erronés s'ils ont été développés dans le EAE comme étant un dispositif fiable de défense par avertissement. Et il cite d'autres exemples de processus fiables, comme la tendance humaine de se recroqueviller en présence de serpents (même inoffensifs), et aussi cette étude minutieusement examinée de la réussite des sujets à la tâche de sélection de Wason, dans les cas où cette tâche introduit des conditions d'échanges sociaux tout « comme ceux [...] que nos chasseurs-cueilleurs du Pléistocène ont dû partager eux-mêmes à l'époque » (p. 98). En bref, il est convaincu d'identifier les processus fiables engendrant les croyances (ou, plutôt, il est convaincu

d'hypostasier de tels processus dans notre passé évolutionnaire) et d'affirmer alors l'existence des modules sous-jacents à ces processus. De cette façon, les modules sont présumés empiriquement discernables, dépassant ainsi le bournier analytique du problème de la généralité.

Cependant, je voudrais soutenir que cela revient à présumer une solution au problème de la généralité, non pas à y répondre. Parce que le problème précis qui nous intéresse est la mesure par laquelle nous pouvons déterminer quand un processus engendrant des croyances est fiable. Le problème de la généralité suggère que cette évaluation de la fiabilité doit attendre une détermination antérieure du niveau approprié de description pour ce processus — et rien de ce qu'a dit Clarke n'exclut ce réquisit. Il est vrai que Clarke s'intéresse aux avantages évolutifs d'un processus pour déterminer sa fiabilité ; mais il existe aussi un problème de la généralité analogue avec les avantages évolutifs — la question de savoir si un processus (engendrant des croyances) est évolutivement avantageux dépend elle-même de la façon dont il est décrit. Par exemple, quant aux claquements des castors déjà mentionnés, la question de savoir s'ils ont un avantage évolutif dépend de la façon dont nous décrivons la situation dans laquelle ils s'effectuent. Donc, pour formuler brièvement l'objection, les modules empiriquement isolés (de la façon suggérée par Clarke) ne peuvent résoudre le problème de la généralité puisque des tels isolements supposent que nous sommes capables d'identifier les processus engendrant des croyances qui sont fiables (ce qui signifie évolutivement avantageux), et une telle identification antérieure est problématique précisément en raison du problème de la généralité.

Comme solution de rechange, laissez-moi explorer comment il est possible de résoudre le problème de la généralité d'une manière cohérente avec un des buts avoués par Clarke dans son livre, qui consiste à montrer comment une perspective externaliste (non méliorative) de la justification épistémique peut et doit informer la perspective internaliste (méliorative) (voir Clarke, p. 101-114). Essentiellement, en déterminant le niveau de description approprié pour un processus engendrant des croyances, je soutiens qu'il faut solliciter la description de l'agent qui croit et qui décrira ce processus d'une certaine façon. Alors, sa croyance en tant qu'engendrée par certains processus — tels que décrits — est-elle justifiée ? Seulement si, en premier lieu, sa description est précise : il ne peut être *méliorativement* (de façon internaliste) justifié d'accepter une croyance s'il l'énonce de façon confuse. De plus, le processus engendrant les croyances telles que décrites doit être fiable en engendrant les croyances vraies. Par exemple, il doit être *suffisamment* fiable, ce qui peut ne pas être le cas si la description de l'agent est trop large. Donc, dans l'ensemble, la croyance d'une personne ne sera pas *épistémiquement* justifiée si elle décrit erronément la façon dont elle arrive (du point de vue externaliste) à sa croyance, ou si elle décrit précisément le processus par lequel elle arrive à sa croyance (peu importe le niveau de description) mais que le processus résultant, tel que décrit, n'engendre pas, de manière fiable, de croyances vraies. Dans le cas où les deux conditions suggérées sont

satisfaites, je soumets que nous avons une forme de justification qui combine tout à la fois la perspective méliorative (puisque la description même de l'agent joue un rôle central) et la perspective non méliorative (puisque la fiabilité externe est requise), une forme qui résout tout aussi bien le problème de la généralité. Elle est en mesure de résoudre le problème de la généralité puisque le niveau de description est établi de façon déterminée par la manière dont l'agent décrit le processus par lequel il est arrivé à sa croyance.

Pour donner un bref exemple afin d'illustrer mon propos, supposez un agent affirmant qu'il sait qu'un certain objet observé est un chêne. Dans le fiabilisme, nous demandons si le processus par lequel il est arrivé à sa croyance est fiable. Mais comment devrions-nous décrire ce processus ? La description est déterminée par la façon dont l'agent lui-même décrit le processus par lequel il est arrivé (du point de vue externaliste) à sa croyance. S'il affirme être arrivé à sa croyance qu'un objet observé est un chêne en donnant un certain crédit à des conseillers extra-terrestres, alors sa description erronée du processus d'acquisition des croyances exclut la possibilité qu'il soit justifié de croire que c'est un chêne. Inversement, s'il décrit le processus comme « une observation faite lors d'un après-midi clair, dans un état psychologique normal », si le processus par lequel il est arrivé à sa croyance peut être décrit (précisément) de la même façon, et si nous sommes assurés que cela est vraiment une méthode fiable à l'acquisition de croyances vraies à propos des chênes, alors sa croyance est épistémiquement justifiée. Je soutiens que ce type de perspective combine effectivement tout à la fois les aspects mélioratifs et non mélioratifs de la justification, une tâche que Clarke approuve. De plus, nous réussissons à faire cela tout en évitant le problème de la généralité, et tout en évitant également (et en ce qu'elle considère le problème de la généralité comme improductif) un détour controversé par les contributions psychologiques de la modularité.

Notice critique : *Reconstructing Reason and Representation* — Murray Clarke

Michael Bishop

Volume 34, numéro 2, automne 2007

URI : <https://id.erudit.org/iderudit/017425ar>

DOI : <https://doi.org/10.7202/017425ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Société de philosophie du Québec

ISSN

0316-2923 (imprimé)

1492-1391 (numérique)

[Découvrir la revue](#)

Citer ce document

Bishop, M. (2007). Notice critique : *Reconstructing Reason and Representation* — Murray Clarke. *Philosophiques*, 34 (2), 367–374. <https://doi.org/10.7202/017425ar>

Notice critique : *Reconstructing Reason and Representation* — Murray Clarke

MICHAEL BISHOP
Université de la Floride

Nul ne peut raisonnablement nier que nos esprits sont le produit de l'évolution. Le problème va au-delà de cette généralisation. La psychologie évolutionniste (PE) gère le problème en proposant une histoire très spécifique de la façon dont nos esprits ont été formés par l'évolution : quelque part au cours du Pléistocène, un événement génétique s'est produit et a occasionné, chez certains de nos prédécesseurs chasseurs-cueilleurs, un module mental (un processeur computationnel) qui a efficacement contribué à relever les défis quant à leur capacité à survivre et se reproduire. Conséquemment, le module acquis par les chasseurs-cueilleurs leur a donné un avantage reproductif sur leurs contemporains qui en étaient privés, et donc, après plusieurs générations, il est devenu prévalent parmi nos ancêtres. Ce processus s'est répété de nombreuses fois pour produire chez nos ancêtres beaucoup (peut-être des milliers) de modules innés et spécialisés (Barkow, Cosmides & Tooby 1992). La sélection naturelle nous a légué des esprits massivement modulaires.

Depuis environ une décennie, de sérieuses inquiétudes sur les fondements biologiques de la PE ont été soulevées (voir Buller, 2005). Le livre ambitieux et intéressant de Murray Clarke, *Reconstructing Reason and Representation*, ne soulève pas ces soucis à leur propos. Il vise plutôt à extraire de la PE un ensemble complet de perspectives sur la connaissance et l'esprit humain. Ce compte rendu mettra l'accent sur trois des thèmes principaux de Clarke : sa défense de la modularité massive, sa solution partielle au problème de la disjonction (ou erreur de représentation), et sa proposition selon laquelle la connaissance est fragmentée.

1. La modularité massive

Le cas standard de « l'effet lié au contenu » de la modularité massive commence avec la découverte empirique selon laquelle la valeur du raisonnement des agents cognitifs dépend du contenu de ses conditionnelles (Cosmides, 1989). Considérez ces deux règles conditionnelles (Clarke, 2004, 92) :

- a. Si une personne a la note « D », alors son document doit être identifié code « 3 ».
- b. Si une personne boit de la bière, alors elle doit avoir plus de 20 ans.

En dépit du fait qu'elles partagent la même structure logique, les gens trouvent plus facile de raisonner correctement à propos de *b* que de *a*. Qu'est-ce qui explique cet effet lié au contenu ? Une hypothèse veut que les humains aient

un « module de détection des tricheurs », lequel est activé à chaque fois qu'un sujet identifie une situation sociale d'échange (Gigerenzer et Hug, 1992). Dans l'exemple ci-dessus, *b* déclenche notre module de détection des tricheurs (donc, nous savons qu'il faut vérifier l'âge du consommateur de bière et la boisson du mineur) mais *a* ne le fait pas (Clarke, 2004, 7-12, 90-98). L'objection la plus sérieuse à cet argument pose une hypothèse rivale : l'effet lié au contenu peut être parfaitement bien expliqué par des mécanismes généraux opérant sur des connaissances innées. En conséquence, la découverte de l'effet lié au contenu ne nous donne pas de bonnes raisons d'accepter que l'esprit soit massivement modulaire. Je crois que cette objection a été proposée explicitement pour la première fois par Richard Samuels (1998), bien que, comme celui-ci le note, elle soit implicite dans le travail de Carey et Spelke (1994). (Pour d'autres versions de l'argument, voir Fodor, 2000, et Buller, 2005.)

Clarke défend que le problème de cette objection est qu'elle présuppose une version invraisemblablement forte de la modularité massive — nommément, que l'esprit est *entièrement* modulaire. Mais il insiste sur l'idée que la modularité massive est une hypothèse inclusive. Il maintient que l'esprit est « dominé » par les modules, mais il reconnaît la possibilité que certaines pensées d'ordre supérieur peuvent être le résultat d'un processeur général agissant sur des ensembles spécifiques ou généraux de connaissance (2004, 7, 36, 28-29, 138-140). Étant donné cette perspective sur la situation dialectique, Clarke souligne que l'hypothèse proposée par Samuels (comme Fodor et les autres) n'est pas incompatible avec l'hypothèse de la modularité massive. Elle laisse donc la modularité massive indemne. Le problème est que dans sa réponse, Clarke se représente mal l'objection. Quand Samuels suggère que l'effet lié au contenu peut être expliqué par une hypothèse rivale, il conclut explicitement que les partisans de la PE ont failli « à nous fournir de bonnes raisons de *préférer* [la modularité massive] à une image *rivale* de l'esprit », et il en résulte que l'hypothèse de la modularité massive « est à la fois sans garantie et immotivée » (Samuels, 1998, 575, soulignement ajouté). Le point est que les partisans défendant la PE sur la base de données psychologiques, comme Clarke, ne nous donnent apparemment aucune raison de croire que l'esprit est dominé par des modules spécialisés *plutôt que* par des processus généraux agissant sur des connaissances spécifiques. Ces données peuvent être parfaitement expliquées sans qu'on fasse appel à aucun module.

À certaines occasions, Clarke saisit la portée réelle de l'objection. Par exemple, lorsqu'il caractérise la version de Fodor comme suit : « En effet, [Fodor] affirme qu'il est possible d'employer un mécanisme d'apprentissage général non modulaire à côté de toutes sortes d'informations innées » (2004, 28). Clarke esquive cette objection, mais seulement en évoquant l'hypothèse de la modularité massive. En répondant à Fodor, il fait endosser aux partisans de la modularité massive l'affirmation qu'il est impossible pour l'esprit d'être seulement constitué d'un « mécanisme général d'apprentissage non modulaire *couplé uniquement avec de l'information perceptuelle* » (2004, 29, soulignement

ajouté). Mais cette conclusion n'est pas équivalente à la version canonique de Clarke concernant l'hypothèse de la modularité massive (qui, je le rappelle, soutient que l'esprit est dominé par des modules spécialisés mais ne s'épuise pas dans ceux-ci). Pour nous en rendre compte, prenons un nativiste comme Fodor. Il sera certainement d'accord avec le fait que la cognition supérieure n'implique pas de mécanismes généraux d'apprentissage agissant *seulement* sur de l'information perceptuelle. Mais nous pouvons néanmoins nier, en restant cohérent, que l'esprit est dominé par des processeurs spécifiques.

Un défi sérieux pour les partisans de la PE est d'expliquer pourquoi il est plausible de croire que l'esprit est dominé par des modules spécialisés plutôt que par des processeurs généraux agissant sur des connaissances spécifiques. D'après ce que j'ai pu constater, ce n'est pas un défi que Clarke a relevé avec succès.

2. Le problème de la disjonction (ou des erreurs de représentation)

Supposez qu'un exemplaire d'un type d'état « R » apparaisse presque toujours en présence de chevaux mais apparaisse plutôt occasionnellement en présence d'autres mammifères quadrupèdes. Que représente « R » ? Intuitivement, il semble évident qu'il représente les chevaux. Auquel cas les exemplaires de « R », en présence d'autres quadrupèdes, constituent une erreur de représentation. Mais peut-être « R » représente-t-il des chevaux ou des mules ? Auquel cas les exemplaires de « R » sont des erreurs de représentation uniquement quand la chose représentée est autre chose qu'un cheval ou une mule. Ou peut-être « R » représente-t-il tous les objets, et seulement ceux qui tendent à susciter les exemplaires de « R » ? Auquel cas les exemplaires de « R » ne peuvent pas être des erreurs de représentation. Le problème de la disjonction (ou de l'erreur de représentation) est donc ceci : en vertu de quel fait les exemplaires de « R » occasionnent-ils des *erreurs* de représentation ?

La solution partielle de Clarke au problème de la disjonction fait appel à la différence entre le *domaine propre* d'un module (l'environnement dans lequel l'évolution a sélectionné ce module) et le *domaine actuel* de ce module (l'environnement dans lequel le module est actif actuellement). La « théorie du fossé » de Clarke explique les erreurs de représentation comme le résultat de la présence d'un fossé entre le domaine propre et le domaine actuel du module. Supposez que les castors possèdent un module cognitif qui tend à signaler « X » quand un prédateur est à proximité ; ce signal amène un castor en train de nager à claquer sa queue sur la surface de l'eau. Ce module est sélectionné par l'évolution parce qu'il signale les prédateurs, et cela, pour Clarke, est ce que le type d'état mental « X » représente. Mais prenez les castors vivant au zoo de San Diego. Puisqu'il n'y a plus de prédateurs dans leur environnement, tout exemplaire de l'état mental « X » est une erreur de représentation. Toutefois, cette erreur de représentation est le résultat du mode « propre » de fonctionnement du module dans un environnement différent de celui pour lequel le module a été adapté. « Les erreurs de représentations ne concernent pas l'esprit » (Clarke, 2004, 60). Elles adviennent parce que l'environnement

dans lequel l'évolution a sélectionné le module associé avec l'état mental est différent de l'environnement dans lequel le module agit actuellement. (Clarke reconnaît que cette solution est partielle, puisque « dans la mesure où cette partie de l'esprit est susceptible d'être non modulaire, une analyse différente serait requise pour rendre compte de ces cas » [2004, 58]).

La théorie du fossé de Clarke soulève une inquiétude manifeste : un état mental est-il une erreur de représentation lorsque son module cognitif associé est dans son domaine propre (quand il n'y a pas de fossé entre le domaine propre et le domaine actuel) ? Prenez un environnement où l'évolution a sélectionné le module chez les prédécesseurs des castors qui les a menés à claquer leur queue. Pendant les (présumées) milliers d'années durant lesquelles le module a été sélectionné, il s'est sûrement déclenché en présence d'un animal qui n'était pas un prédateur, disons un ongulé végétarien. Ce cas est-il un exemple d'erreur de représentation ? Il semble que Clarke soit forcé à le nier puisqu'il n'y a pas de fossé entre le domaine propre et le domaine actuel. Mais Clarke rejette l'implication : « Bien sûr, je ne voulais pas suggérer qu'aucune erreur n'advient dans le [domaine propre] ; le Pléistocène n'était pas parfait » (2004, 60). Mais si l'erreur de représentation est le résultat d'un module qui n'est pas dans son domaine propre, comment les erreurs de représentations peuvent-elles advenir quand ce module est dans son domaine propre ? Voilà ce qu'en dit Clarke :

Il n'y a aucun domaine dont le domaine propre n'est pas contextuel... Donc, par exemple, la période du Pléistocène était un domaine propre relativement à nos domaines actuels (contemporains). Mais, la période du Pléistocène était aussi un domaine actuel relativement au domaine propre du Miocène. Penser autrement est supposer que la sélection naturelle n'est apparue pour la première fois que durant la période du Pléistocène (2004, 65).

Dans ce passage, Clarke suggère que les erreurs de représentation adviennent parce que le domaine propre du Pléistocène possède un domaine propre antérieur dans le Miocène. Mais cela est confus. Les domaines propres sont définis relativement à des modules particuliers : le domaine propre d'un module est l'environnement dans lequel l'évolution a sélectionné ce module. Si le module a été sélectionné durant une certaine période du Pléistocène, alors, par définition, c'était son domaine propre. Il est incompréhensible de se prononcer sur le « domaine propre » pour « nos domaines actuels (contemporains) » ou pour « la période du Pléistocène ». Bien entendu, la théorie du fossé peut expliquer comment les erreurs de représentation advenaient durant le Pléistocène — si l'esprit de nos ancêtres du Pléistocène consistait en modules dont le domaine propre précédait le Pléistocène. Mais cela ne résout pas l'énigme de la façon dont la théorie du fossé explique les erreurs de représentation quand le module associé avec l'état mental qui constitue une erreur de représentation est dans son domaine propre.

3. Modules fiables et connaissance

Clarke défend que l'évolution a sélectionné des modules cognitifs hautement fiables. Bien qu'il admette que nos mécanismes judicatifs ne sont pas toujours parfaitement fiables (2004, 78), il ne dit pas jusqu'à quel point la sélection naturelle a été efficace dans le façonnage des mécanismes judicatifs fiables. *Tous* nos modules cognitifs sont-ils hautement fiables — au moins dans leur domaine propre ? Ou seulement la plupart d'entre eux ? Ce n'est pas clair, mais Clarke donne le ton de sa discussion au début du chapitre quatre : « Il y a peut-être [...] des cas où "l'ignorance est source de bonheur" ou "peu de connaissance est une chose dangereuse", mais généralement les ensembles de fausses croyances ne peuvent tout simplement pas, le plus souvent, être un avantage pour les humains. En fait, je suis presque embarrassé de devoir défendre l'affirmation que les fausses croyances diminuent l'adaptabilité de l'organisme [*inclusive fitness*], donc sa négation me semble entièrement absurde » (2004, 75).

À qui Clarke attribue-t-il cette affirmation « absurde » ? Il n'est pas du tout évident que Stephen Stich, sa principale cible dans ce chapitre, l'affirme (voir 1990, 121). Ce que défend Stich, c'est que l'évolution a pu sélectionner des mécanismes inférentiels faillibles qui commettent beaucoup d'erreurs mineures peu coûteuses en échange de ne (presque) jamais faire d'erreurs majeures très coûteuses. Il décrit un scénario dans lequel un organisme emploie une stratégie d'évitement des erreurs majeures pour identifier les aliments toxiques. L'organisme en vient parfois à considérer erronément qu'un aliment est toxique mais (presque) jamais faussement qu'un aliment n'est pas toxique.

La stratégie d'évitement des risques, qui est faillible et prompte à l'erreur, a bien pu être favorisée par la sélection naturelle. Parce que la sélection naturelle est indifférente à la vérité ; elle s'occupe seulement du succès reproductif. Et du point de vue du succès reproductif, il est souvent préférable d'être sauf (ou trompé) que désolé (Stich, 1990, 62).

L'argument de Stich montre qu'il est possible pour l'évolution de sélectionner des mécanismes faillibles de formation des croyances, mais Clarke défend « que les arguments conceptuels à propos de la simple possibilité physique font peu de travail empirique » (2004, 77). L'ironie, c'est que pendant que Clarke rejette l'argument de Stich parce qu'il met l'accent uniquement sur « ce qui "pourrait bien" être le cas » (2004, 77), il procède immédiatement à la défense du fait que les croyances vraies sont instrumentalement bénéfiques, « les processus fiables *auraient bien pu* être sélectionnés » (2004, 78, soulignement ajouté). Équité oblige, Clarke présente aussi deux autres arguments. Le premier repose sur l'hypothèse que « l'on pourrait avoir autant de raison de croire que le raisonnement fiable a été sélectionné pour ce propos que de croire que des capacités visuelles faibles ont été sélectionnées pour celui-ci » (2004, 90). Le second argument fait appel au fait que nous raisonnons de façon assez fiable dans les cas de détection des tricheurs, et de la même manière lorsque les problèmes sont formulés en termes de fréquences plutôt que de probabilités

(2004, 90-98, 118-119). Comme Clarke le reconnaît, aucun de ces arguments ne nous donne de raisons fortes de croire que *tous* les milliers de modules cognitifs que l'évolution est censée avoir sélectionné sont fiables — ou étaient fiables dans leurs domaines propres (2004, 99).

Sommes-nous dans une impasse opposant deux arguments aussi plausibles ? Peut-être que non. Il existe des appuis empiriques qui représentent un sérieux défi à l'optimisme évolutionnaire de Clarke concernant nos habiletés cognitives. Shelley Taylor et ses collègues ont affirmé que « l'optimisme irréaliste à propos de l'avenir permet généralement une meilleure adaptabilité en ce qu'il promeut [...] des sentiments d'amour propre, la capacité à et envers les autres, la persistance et la créativité dans la poursuite d'objectifs, et la capacité à gérer efficacement le stress » (Taylor et al., 1992, 460). Les recherches suggèrent que les optimistes minimalement irréalistes s'en tirent mieux quant aux problèmes de santé que les réalistes, ils tendent à guérir plus rapidement que les réalistes et ils ont des taux de rémission plus élevés que les réalistes (Scheier et al., 1989 ; Fitzgerald et al., 1993). Alors que l'optimisme irréaliste semble avoir des effets salutaires sur la santé et la survie, la plupart d'entre nous semblent posséder une surprenante et persistante capacité à adopter et protéger de l'infirmité les croyances trop optimistes nous concernant (Taylor, 1989, chapitre 4).

Nous sommes enclins à croire que nous possédons une foule de caractéristiques socialement désirables, et que nous sommes exempts de la plupart de celles qui sont socialement indésirables. Par exemple, une large majorité de la population générale pense qu'elle est plus intelligente, plus encline à l'équité, moins préjudiciable et plus habile derrière un volant d'automobile que la personne moyenne. Ce phénomène est si récurrent et omniprésent qu'il est désormais connu sous le nom de « l'effet Lac Wobegon », à la suite de la communauté fictive de Garrison Keillor où « les femmes sont fortes, les hommes beaux et tous les enfants sont supérieurs à la moyenne ». Une enquête établie sur un million d'élèves sortants du secondaire montre que 70 % d'entre eux pensent qu'ils se situent au-dessus de la moyenne quant à leur habileté de meneur et seulement 2 % pensent qu'ils sont en dessous de la moyenne. En termes de socialisation avec les pairs, tous les étudiants pensent qu'ils sont au-dessus de la moyenne, 60 % pensent qu'ils sont parmi les premiers 10 % et 25 % pensent se trouver dans le premier 1 % ! De peur que d'aucuns croient qu'une autoévaluation si enflée n'a cours que dans les esprits des étudiants inexpérimentés du secondaire, il faudrait souligner qu'une enquête sur les professeurs d'université montre que 94 % d'entre eux se pensent plus compétents dans leur travail que la moyenne de leurs collègues (Gilovich, 1991, 77).

Si la PE est sur la bonne voie, il n'est pas « absurde » de supposer que l'évolution a pu favoriser au moins quelques modules cognitifs du type Lac Wobegon.

Clarke défend une conception largement fiabiliste de la connaissance (où celle-ci est une croyance produite qui est vraie et fiable). Les représentations vraies qui sont « abritées » dans les modules darwiniens innés et fiables sont

des connaissances innées (2004, 140-141) ; et les représentations vraies qui sont les *outputs* de modules darwiniens fiables sont des connaissances empiriques (2004, 141-142). Clarke identifie un avantage important de coupler le fiabilisme avec la psychologie évolutionnaire : la promesse d'une solution graduelle au problème de la généralité, au problème de trouver une solution principielle à l'identification des processus mentaux dont la fiabilité fixe le statut justificatif des croyances empiriques. Une croyance empirique est une connaissance si elle est une croyance vraie produite par un module darwinien fiable et inné (2004, 64). Jusque-là, tout va bien.

Le problème, de mon point de vue, vient de ce que Clarke défend une conception fragmentée de la connaissance. Sa défense s'effectue en deux parties. Premièrement, il offre les raisons suivantes de croire que la connaissance est un genre naturel (2004, 133-136) : la connaissance existe ; ce n'est pas un « genre conceptuel » parce que si c'en était un, elle serait relative culturellement (ce qu'elle n'est pas). Alors, en conséquence, elle doit être un genre naturel. Clarke soutient l'affirmation que la connaissance doit être un genre ou n'être rien du tout sur la base que si d'aucuns revendiquent qu'elle est autre chose, « [il] commence à perdre prise sur ce qui constitue l'objet de la discussion. Appelez cela la conception « "mystérieuse" de la connaissance » (2004, 135). Mais, assurément, cet argument est trop permissif. Supposez que nous nommions la chose dans mon bureau « la chose du bureau de Mike ». Nous pouvons employer le cadre des arguments de Clarke pour montrer que, parce que la chose du bureau de Mike existe et n'est pas culturellement relative, ce doit être un genre naturel. Mais c'est entièrement faux. Clarke ne reconnaît pas la possibilité des *salmigondis*. Peut-être ce que nous nommons « connaissance » est-il un assemblage hétéroclite d'états qui ne peuvent pas jouer un rôle prééminent dans les lois causales. En fait, Clarke se trouve sur le point d'embrasser cette perspective dans la seconde partie de sa défense d'une conception fragmentée de la connaissance. Il affirme qu'« il n'existe pas de genre naturel unique qui se trouve être la connaissance » (2004, 145). Cela parce que chaque module cognitif possède sa propre histoire causale, et que les genres naturels sont définis (en partie) en termes de lois causales responsables de leur unité. Mais au lieu de conclure que la connaissance est un *salmigondis* fragmenté, il affirme que la connaissance est un « ensemble de genres naturels » fragmentés (2004, 128, 146).

La conception de la connaissance fragmentée de Clarke est intrigante mais sommaire. En fait, l'acquisition de la connaissance pour la PE impliquerait la coopération de multiples modules différents. Par exemple, supposez que j'aie appris pour la première fois que *p* par témoignage. Cela implique la mémoire à court terme, plusieurs processeurs linguistiques et des processeurs d'information de nature sociale (par exemple, le locuteur est-il digne de confiance ? est-ce qu'il s'exprime littéralement ?). Ensuite, mes connaissances ont été enrichies par l'expérience (ce qui implique aussi la coopération de plusieurs modules différents), et c'est à présent par la mémoire à long terme que tout

cela est maintenu. À chaque fois que mes connaissances sont étayées par l'opération d'un autre module, il semblerait, dans la perspective de Clarke, que mes connaissances deviennent une instanciation d'un nouveau genre naturel. Si cela est acceptable et s'il existe des milliers de modules, alors il pourrait facilement y avoir des millions de combinaisons différentes de modules qui ont pris part à l'établissement et au maintien des différentes instanciations des connaissances. La connaissance, selon la perspective de Clarke, semble être fractionnée en millions de fragments. Mais quel intérêt théorique ou explicatif peut nous rapporter une telle sur-taxonomisation ? J. D. Trout et moi-même (2005) avons argumenté que l'épistémologie est obsédée, à tort, par la catégorie de la connaissance. Je suggère que la perspective de Clarke donne raison à notre dérogation face à la connaissance, puisqu'elle mène à une conception de celle-ci trop éclatée pour être théoriquement utile.

Bibliographie

- Barkow, J., L. Cosmides et J. Tooby. *The Adapted Mind : Evolutionary Psychology and the Generation of Culture*, Oxford, Oxford University Press, 1992.
- Bishop, M. et J. D. Trout, *Epistemology and the Psychology of Human Judgment*, New York, Oxford University Press, 2005.
- Buller, D. *Adapting Minds : Evolutionary Psychology and the Persistent Quest for Human Nature*, Cambridge, MA, The MIT Press, 2005.
- Carey, S. et E. Spelke. *Domains-Specific Knowledge and Conceptual Change*, in *Mapping the Mind : Domain-specificity in Cognition and Culture*, L. Hirshfield et S. Gelman (dir.), Cambridge, Cambridge University Press, 1994.
- Clarke, M. *Reconstructing Reason and Representation*, Cambridge, MA, The MIT Press, 2004.
- Cosmides, L. *The Logic of Social Exchange : Has Natural Selection Shaped How Humans Reason*, *Ethology and Sociobiology*, 10, 51-97, 1989.
- Fitzgerald, T., H Tennen, G. Affleck, G. Pransky. *The relative importance of Dispositional Optimism and Control Appraisals in Quality of Life after Coronary Bypass surgery*, in *Journal of Behavioral Medicine*, 1993, 16, 25-43.
- Fodor, J. *The Mind Doesn't Work That Way*, Cambridge, MA, The MIT Press, 2000.
- Gigerenzer, G. et K. Hug. *Domain-specific Reasoning : Social Contracts, Cheating and Perspective Change*, *Cognition*, 1992, 43, 127-171.
- Gilovich, T. *How We Know What Isn't So*, New York, The Free Press, 1991.
- Samuels, R. *Evolutionary Psychology and the Massive Modularity Hypothesis*, *British Journal for the Philosophy of Science*, 1998, 49, 575-602.
- Scheier, M., K. Matthews, J. Owens, G. Magovern, R. Lefebvre, R. Abbott et C. Carver. *Dispositional optimism and Recovery from Coronary Artery Bypass Surgery : The Beneficial Effects on Physical and Psychological Well-Being*, in *Journal of Personality and Social Psychology*, 1989, 57, 1024-1040.
- Stich, S. *The Fragmentation of Reason*, Cambridge, MIT Press, 1990.
- Taylor, S. *Positive Illusions : Creative Self-Deception and the Healthy Mind*, New York, Basic Books, 1989.

Faraway, so close?

Luc Faucher

Volume 34, numéro 2, automne 2007

URI : <https://id.erudit.org/iderudit/017426ar>

DOI : <https://doi.org/10.7202/017426ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Société de philosophie du Québec

ISSN

0316-2923 (imprimé)

1492-1391 (numérique)

[Découvrir la revue](#)

Citer ce document

Faucher, L. (2007). *Faraway, so close?*. *Philosophiques*, 34 (2), 375–383.
<https://doi.org/10.7202/017426ar>

Faraway, so close?

LUC FAUCHER

Département de philosophie UQAM

Dans *Reconstructing Reason and Representation*, Murray Clarke se propose d'utiliser une version de la psychologie évolutionniste (celle défendue entre autres par Cosmides et Tooby, Pinker et Buss) pour jeter une lumière nouvelle sur certains problèmes de la philosophie de l'esprit, de l'épistémologie et de la philosophie des sciences. Bien que je sois sympathique à ce projet, j'aimerais cependant défendre l'idée que la position de Clarke est, par certains côtés, trop éloignée et, par d'autres, trop près de la psychologie évolutionniste.

Trop loin

À de nombreuses reprises, et en réponse aux arguments de philosophes comme Stich, Clarke soutient que les théories évolutionnistes nous donnent des raisons de penser que les mécanismes psychologiques hérités de nos ancêtres chasseurs-cueilleurs sont fiables, c'est-à-dire qu'ils produisent des croyances vraies (ou des représentations indexicales précises [*accurate*]). Il écrit par exemple :

Reliable processes (such as good reasoning, perceptual capacities, and other sensory modalities) being traits, were selected for because they produce something that confers a tremendous evolutionary advantage: accurate indexical representations. True beliefs are the product of reliable processes... (80 ; je souligne)

Je veux soutenir qu'une lecture attentive de ce que les psychologues évolutionnistes affirment montre deux choses qui vont à l'encontre de la position de Clarke : 1) les psychologues évolutionnistes sont des pragmatistes et proposent une version très affaiblie (sinon tout à fait différente de ce que Clarke défend) de ce qu'est une « croyance vraie » ; 2) il n'est pas du tout clair que les processus fiables produits par l'évolution ont comme *output* des croyances vraies.

Tout d'abord, pour les psychologues évolutionnistes, la « Vérité » n'est pas importante pour la survie ; ce qui l'est, c'est l'action correcte (celle-ci pouvant être générée par des croyances fausses). En ce sens, les évolutionnistes sont des pragmatistes :

For animals, knowledge only exists because ancestrally its production served as a means to correct action. [...] The usual move is to argue that successful action self-evidently seems to depend on the attainment of true belief, so that the primary functional identity of the brain must be as a knower, a reasoner, and an acquirer of truths (Cosmides et al., 2005, p. 328).

Mais, comme ils le notent :

There have been a series of negative consequences for cognitive science that stem from its primary emphasis on knowledge acquisition.[...] [parmi celles-là, le fait

que], it ignores the many causal pathways whereby our evolved architecture should have been designed to manufacture, store, communicate, and act on the basis of representations that would not qualify as a rational architecture's efficient attempt at constructing true beliefs (p. 329 ; je souligne).

Ils expliquent qu'il existe une kyrielle de raisons allant à l'encontre de l'idée selon laquelle le fait qu'un processus est l'objet de la sélection naturelle implique qu'il devrait produire des croyances vraies :

Leaving aside the necessary coparticipation of value in the definition of truth, the existence of conflicts of interest in social life constitutes the source of many other deviations from truth-seeking as an engineered goal of all cognitive mechanisms. The system may be required to reason about value, and there is no truth of the matter for valuation. Individuals may adopt beliefs (e.g. God is three in one; Darwinism is irrelevant) because they socially coordinate them with others. The recomputation required to adopt the true belief may be too costly, at least for a period of time, so that temporary denial (as in grief) may be functional. The introduction of true information may be too disruptive to successful functioning, as when you choose not to look down when climbing a cliff face. To the extent a data store is computed for communication to others rather than to be acted on by oneself, then the optimal impact on others will be the criterion and not truth value (note 11, p. 329).

Pour donner un autre exemple d'un mécanisme psychologique sélectionné qui ne produit pas nécessairement de croyances vraies, prenons le mécanisme responsable de l'évitement de l'inceste. Les travaux empiriques sur la question confirment une hypothèse, connue depuis déjà une trentaine d'années, selon laquelle un des indices utilisés dans l'évitement de l'inceste entre frères et sœurs est la co-socialisation dans les premières années de la vie. Lieberman (Lieberman, et al., 2003 ; voir Faucher, 2007, pour un résumé de ces travaux) suppose que la période de co-socialisation est utilisée pour déterminer qui sont les apparentés d'un sujet. Cette conclusion concernant les apparentés peut ensuite être acheminée à différents modules, comme celui en charge de la coopération parentale et celui en charge des préférences sexuelles. Comme le remarquent Cosmides, Tooby et Barrett, les deux systèmes peuvent avoir des seuils d'activation différents :

[...] [B]ecause the scope and fitness consequences for helping and for incest avoidance are different, the brain may represent two different (but related) values for genetic relatedness between a given pair of individuals. Each is "true" (functionally calibrated), with respect to the action-regulatory system it inhabits, but they may be different. Females may, for example, represent individuals as more highly related for purposes of incest avoidance than as objects of altruism, because the asymmetric consequences of a miss versus a false alarm are different for incest avoidance and kin assistance (2005, p. 332).

Non seulement est-il possible que dans une situation une femme considère que quelqu'un est son apparenté alors que dans une autre non, mais les différences existent également entre sexes. Toujours selon Lieberman (mais voir

également Fessler et Navarrette, 2004), les femmes concluent plus rapidement à l'apparement que les hommes, c'est-à-dire que moins d'années de co-socialisation sont nécessaires pour conclure à l'apparement pour les femmes que pour les hommes. Donc, par exemple, une femme pourrait considérer un homme X comme un apparement alors que ce même X, lui, ne la considérerait pas comme son apparement (bien entendu, nous ne parlons pas ici de représentations conscientes et/ou culturelles). Qui plus est, cette femme pourrait considérer ce même X comme un apparement dans le cadre de ses préférences sexuelles et comme un non-apparement dans le cadre de ses préférences altruistes.

Je crois que des cas comme ceux de l'inceste (mais on aurait pu en donner d'autres, voir entre autres Haselton et al., 2005) montrent deux choses : en supposant quelque chose comme le concept de vérité-correspondance et que cette conception dépende de l'idée que le monde auquel nos énoncés ou pensées réfère existe indépendamment de nous (réalisme), alors : 1) ce que la cognition piste (*track*) peut varier en fonction de l'âge, du sexe, de la position hiérarchique, de l'état émotionnel et motivationnel de l'individu, etc. (les systèmes régulateurs d'action dont parlent Cosmides et Tooby tiennent compte de ces variables dans leur évaluation des états du monde) ; donc : 2) la cognition ne piste (*track*) pas la vérité¹.

Trop proche

Comme je le mentionnais au début de ce texte, Clarke utilise une version particulière de la psychologie évolutionniste pour explorer certaines questions philosophiques, comme celle de la nature de la connaissance. Selon plusieurs, cette version « modulo-centriste » de la psychologie tend à considérer la culture plus ou moins comme un vernis sur un noyau de mécanismes universels et spécifiques à l'espèce. L'acceptation de ce paradigme se traduit chez Clarke par une attention exclusive aux mécanismes de raisonnement psychologique évolués :

What I mean by “evolutionary reasoning” is that those reliable processes connected with reasoning take the form of Darwinian algorithms : a plethora of specialized, domain-specific inference rules designed to solve recurrent, adaptive problems in the social exchange contexts that our ancestors, that is, the Pleistocene hunter-gatherer, faced. In effect, my claim will be that there are no universally reliable reasoning processes, but only “reliable-in-a-social-exchange-context” inference rules (89).

1. Clarke pourrait rétorquer contre (1) que son argument ne vise qu'à montrer qu'il est possible de concevoir que certains mécanismes peuvent « pister » la vérité et que les mécanismes de raisonnement sont de ceux-là. Dans son livre, il utilise une analogie avec la vision pour montrer que tous les mécanismes ne peuvent pas être conçus comme produisant surtout des croyances fausses. Je n'avais pas l'espace ici pour défendre ma position contre cette réponse, mais je crois que des phénomènes comme la vision colorée (voir Hardin, 1990, p. 556) : « [...] *the facts about chromatic phenomena make it very hard to construe colors as properties of physical objects or processes outside the body of the perceiver* ») le remplissage (*filling-in*), la plasticité diachronique de la perception (McCauley et Heinrich, 2006), la vision ego-centrée, etc., montrent (au mieux) que l'analogie utilisée par Clarke ne réussit pas à établir ce qu'il voulait.

Elle le conduit également à considérer le raisonnement abstrait comme un parasite (*free rider*) :

The sort of abstract reasoning involved in science, mathematics, and logic is the result of a free rider effect. But free rider effects confer no evolutionary advantage (82).

Dans ce qui suit, je ne m'attarderai pas à examiner la valeur des travaux démontrant l'existence de ces modules (pour un résumé des discussions à ce sujet, voir Buller, 2005). Ce que je ferai plutôt est de présenter une solution de rechange à la conception étroite du rapport entre culture et cognition que soutient Clarke, celle défendue entre autres par les partisans de la conception de l'esprit étendue (mais aussi de la construction de niche et de la co-évolution gène-culture ; voir Odling-Smee et al., 2003, pour la première conception, et Boyd et Richerson, 2004, pour la seconde). J'aime penser à ce débat comme celui entre (Murray) Clarke et (Andy) Clark.

Dans un article récent (2001), Andy Clark défend l'idée selon laquelle nous sommes des « *natural born cyborg* ». Ce qu'il entend par là est que :

We — far more than any other creatures on the planet — deploy non-biological elements (instruments, media, notations) to complement our basic biological modes of processing, creating extended cognitive systems whose computational and problem-solving profiles are quite different from those of the naked brain (134).

Il y a deux façons, selon Clark, de concevoir cette extension du cerveau (les deux ne sont pas incompatibles et co-existent assurément) : une première (préférée par Dennett, voir également Griffiths et Stoltz, 2000²) consiste à voir les technologies culturelles comme reprogrammant le cerveau ; la seconde (préférée par Clark) consiste à voir les technologies culturelles comme des additions aux pouvoirs du cerveau nu qui ont pour fonction d'étendre ses capacités (ce qu'il nomme l'échafaudage externe [*external scaffolding*]).

Cette conception de l'esprit étendu et situé socialement, si elle est exacte (et je crois qu'elle l'est), pose au moins trois problèmes pour Clarke.

Le premier de ces problèmes est la remise en cause de l'universalité de l'organisation cognitive humaine. Pour prendre un exemple bien connu, le psychologue Richard Nisbett et ses collègues ont montré dans leurs travaux qu'il existe des « styles » de raisonnement différents selon les cultures, plus précisément que :

- 1) les jugements des orientaux sont plus dépendants du contexte que ceux des occidentaux ;

2. "These 'biological' features represent our true nature — the naked ape stripped of its cultural clothes. It seems to us that this traditional project is as misguided as seeking to investigate the true nature of an ant by removing the distorting influence of the nest ! [...] Human nature must inevitably be a product of a developmental matrix which includes a great deal of cultural scaffolding" (Griffiths et Stoltz, 2000, p. 45).

- 2) les occidentaux sont plus susceptibles de commettre l'erreur d'attribution fondamentale que les orientaux ;
- 3) les orientaux groupent en général les objets en fonction de leurs relations et similarités, alors que les occidentaux les groupent en fonction de leur appartenance à une catégorie ;
- 4) Les occidentaux décontextualisent plus facilement le contenu et la structure que les orientaux et font donc plus facilement des inférences à propos de la validité des arguments.

Ces résultats rendent plausible l'idée

qu'un nombre indéfiniment grand de processus cognitifs « fondamentaux » sont hautement malléables [et que] même si toutes les cultures possèdent essentiellement les mêmes processus cognitifs de base comme outils, leurs outils de choix pour le même problème sont habituellement très différents. Les gens peuvent différer de façon marquée dans leurs croyances concernant la question de savoir si un problème requiert une clé à mollette ou une pince, dans leur habilité à utiliser les deux types d'outil et dans la localisation des outils particuliers sur le dessous ou sur le dessus de leur boîte à outils. De plus, les membres de différentes cultures peuvent ne pas voir la même situation comme un problème nécessitant une réparation. Une apparence de contradiction est un problème pour un occidental alors qu'elle peut bien ne pas l'être pour un oriental. [...] Une autre façon dont les processus cognitifs peuvent différer est que les cultures peuvent construire des outils cognitifs composites à partir de la même trousse à outils (*toolkit*) universels de base, accomplissant ainsi des actes d'ingénierie cognitive élaborés [...] comme les statistiques modernes, les règles méthodologiques, les analyses coût-bénéfice [...] (Nisbett et al, 2001, 306).

Si l'analyse des résultats de Nisbett et ses collègues est exacte, il existerait donc plusieurs (au moins deux) styles cognitifs différents de raisonnement. On peut se demander comment l'épistémologie fiabiliste de Clarke pourrait être transformée pour accommoder ces données. Peut-être Clarke croit-il que le raisonnement en contexte d'échange social est immunisé contre ce genre de variation. Heinrich et Heinrich (2006) montrent cependant qu'il y a de bonnes raisons de penser que ce n'est pas le cas. Ils argumentent que les variations dans certains paramètres de l'environnement coopératif humain (grandeur des groupes, durée des interactions, qualité des informations concernant les comportements passés des interacteurs, écologie des stratégies sociales) impliquent que « l'évolution génétique de la coopération par la réciprocité directe *n'est pas* une solution robuste pour la plupart des animaux dans la plupart des contextes d'échange complexe [...] [il n'est donc pas possible] de construire une machine à réciprocité tout usage » (2006, p. 234). Aussi, suggèrent-ils que l'apprentissage culturel est une composante essentielle du mécanisme d'adaptation à l'échange social. Mère Nature nous dote bien sûr d'une « psychologie de la réciprocité », mais, selon eux, cette machinerie ne contient pas de règles comme le donnant-donnant (*tit-for-tat*), mais plutôt une « machinerie

pour apprendre rapidement les règles de réciprocité efficaces localement [*locally-successful rules of reciprocity*] » (idem, p. 235). Si Heinrich et Heinrich ont raison, il pourrait bien se faire que la fiabilité des règles de réciprocité ne doive pas être évaluée en fonction de notre environnement évolutionniste d'adaptation (EEA), mais plutôt en fonction des contextes adaptatifs particuliers (CAP).

Un second problème, découlant lui aussi de l'idée de re-programmation culturelle de l'esprit, concerne une des prémisses utilisées par Clarke dans son argument en faveur de l'idée que la connaissance est une espèce naturelle (NKA), qui me semble fausse, soit la prémisse 9 selon laquelle le relativisme épistémique est faux. La prémisse est importante non seulement dans le contexte de l'argument de la connaissance comme espèce naturelle, mais également dans la solution qu'offre Clarke aux disputes entre philosophes analytiques et philosophes naturalistes au sujet des notions de justification. Si je comprends bien, sa solution au dernier problème consiste en une division des tâches où les philosophes analytiques produiraient une analyse du concept de justification ou de connaissance et où les philosophes naturalistes étudieraient plutôt les conditions dans lesquelles les mécanismes de production de croyances peuvent être considérés comme fiables. Au chapitre V, Clarke pose qu'il est « crucial » de voir les deux projets comme connectés (p. 105) : l'analyse non méliorative du concept de justification ou de connaissance contribuant à définir ce que doivent produire les processus qui sont étudiés par les scientifiques (au chapitre suivant, Clarke semble abandonner le projet de collaboration).

Un problème potentiel pour la prémisse (mais également pour la division des tâches) a été mis au jour par Stich et ses collègues (Nichols et al., 2003 ; pour un résumé voir Stich, 2004) est que, contrairement à ce qu'affirme Clarke, le concept de justification ou de connaissance analysé par les philosophes analytiques semble dépendre de la culture exactement de la façon dont les constructionnistes sociaux le supposent. Ce que montrent les études de Stich est que les intuitions épistémiques des sujets varient selon les cultures auxquelles ils appartiennent. Prenons l'exemple suivant :

Bob a une amie, Jill, qui conduit une Buick depuis des années. Bob pense donc que Jill conduit une voiture américaine. Il n'est cependant pas averti du fait que la Buick de Jill a été récemment volée, ni que celle-ci l'a remplacée par une Pontiac, soit une autre marque de voiture américaine. Bob sait-il vraiment que Jill conduit une voiture américaine, ou le croit-il simplement ?

Pour prendre un exemple, si la plupart des occidentaux (74 %) sont tentés de dire que dans la situation décrite par la vignette précédente, Bob ne possède pas une connaissance, mais seulement une croyance, 57 % des orientaux croient, eux, qu'il a une connaissance (le pourcentage passe à 61 % quand on considère les sujets indiens, pakistanais et bengalis).

Il semble donc que des gens appartenant à des cultures différentes aient des intuitions différentes concernant ce qu'est la connaissance. Il semble donc que

la prémisse 9 est fausse. Les travaux de Stich ont également un impact sur la compréhension de ce à quoi se livrent les philosophes analytiques : lorsqu'ils font une analyse conceptuelle de la justification ou de la connaissance, leur analyse est limitée à leur culture (et même à leur classe sociale). Pour le dire autrement, « la réponse correcte aux questions normatives sera à chaque fois différente pour chaque groupe ayant des intuitions différentes » (Stich, 2004, p. 69).

Je crois que le problème n'est pas dramatique pour le projet de Clarke, bien qu'il doive abandonner l'idée d'une continuité entre le projet classique de l'épistémologie et sa version naturalisée. Si l'on accepte ce relativisme, il faut je crois également accepter d'une part que les résultats de l'épistémologie classique ont une valeur ethnographique, mais surtout que la notion de connaissance avec laquelle on décidera de travailler (par exemple, si on est évolutionniste) peut être assez radicalement différente de celle du sens commun. C'est justement, il me semble, le projet de Clarke au chapitre VI.

Le troisième problème posé par la culture dans l'approche de Clarke est le fait de considérer que les technologies culturelles ne sont que des parasites (*free riders*). Je suis bien sûr d'accord avec lui pour dire que nous ne sommes pas préparés par mère Nature à suivre les canons de la logique, mais il ne faudrait pas négliger le rôle adaptatif de la culture. Comme le notent les anthropologues Boyd et Richerson, ce qui explique que les humains ont pu occuper des environnements si différents et parfois si hostiles n'est probablement pas la possession de modules darwiniens universels, mais plutôt des capacités psychologiques relevant davantage du domaine général, comme l'apprentissage social (l'imitation). C'est que :

l'apprentissage social permet aux populations humaines d'accumuler des réservoirs d'informations adaptatives au cours de plusieurs générations, ce qui mène à l'évolution culturelle cumulative d'institutions et de technologies hautement adaptatives (manuscrit).

Sans tomber dans le fonctionnalisme à tout crin, on peut affirmer que certains produits culturels, comme la logique ou les mathématiques, sont de telles technologies sociales adaptatives. À ma connaissance, il n'existe pas de données concernant l'adaptativité de ces systèmes symboliques culturels, mais on peut l'inférer de certaines caractéristiques de l'épidémiologie de ces systèmes. Par exemple, on possède une histoire assez complète (voir Dahaene, 1997, chapitre IV) de la façon dont certains systèmes de représentations numériques ont été « culturellement » sélectionnés parce qu'ils offraient des avantages sur les autres (c'est-à-dire qu'ils permettaient de faire plus facilement certaines opérations que leurs compétiteurs. Par exemple, la représentation en chiffres arabes rend la multiplication beaucoup plus facile que la représentation en chiffres romains, permettant à certaines activités humaines liées au commerce de se développer et de s'étendre). On pourrait avancer un argument similaire concernant le raisonnement logique abstrait dont le

développement semble lié aux pratiques de débat, de relations contractuelles, de science théorique, et à la formalisation de la connaissance.

Je crois que si l'on adopte la perspective de Boyd et Richerson, l'existence d'outils culturels comme la logique abstraite, même s'ils sont difficiles à manipuler, peut changer la donne épistémique. Elle nous force entre autres à abandonner la perspective étroite des modules et des processus psychologiques internes qui étaient certes adaptatifs dans l'environnement de nos ancêtres pour considérer l'esprit comme fonctionnant à l'intérieur d'un filet d'institutions et de technologies culturelles. Il est possible que nous ne soyons pas préparés par mère Nature à raisonner de façon fiable dans tous les domaines, mais qu'à cela ne tienne, il existe une technologie pour le faire ! Bien sûr, cette technologie est difficile à maîtriser, mais qui a dit qu'être rationnel ne demandait pas d'effort !³ Pour reprendre un exemple utilisé parfois par Boyd et Richerson, nous n'avons sûrement pas de module dédié à la construction de kayak, mais si nous habitons le Grand-Nord, il est utile de faire des efforts pour apprendre à construire cette technologie si nous voulons survivre.

Une véritable épistémologie naturaliste ne sera donc pas étroitement modulaire, elle laissera également une place pour la culture. Comme le rappellent Boyd et Richerson, le fait que l'approche évolutionniste et l'approche culturaliste négligent chacune les apports de l'autre

revient à couper les gens en deux et à n'étudier qu'une moitié ou l'autre. Les humains ne peuvent survivre sans leur culture et leurs gènes intacts. N'étudier que des moitiés est comme étudier des corps mutilés plutôt que des humains vivants. Nous avons besoin d'une théorie dans laquelle la culture est importante et imbriquée de façon intime avec les autres aspects de la biologie humaine (Boyd and Richerson, version manuscrite du premier chapitre de leur livre de 2004).

Bibliographie

- Boyd, R. et P. J. Richerson. *Not by Genes Alone : How Culture Transformed Human Evolution*, Chicago, University of Chicago Press, 2004.
- Buller, D. J. *Adapting Mind : Evolutionary Psychology and the Persistent Quest for Human Nature*, Cambridge, MIT Press, 2005.
- Clark, A. « Reason, Robots and the Extended Mind », *Mind and Language*, vol. 16, n° 2, 2001, p. 121-145.
- Cosmides, L., J. Tooby et C. Barrett. « Resolving the Debate on Innate Ideas ». In *The Innate Mind : Structure and Content*, sous la direction de P. Carruthers, S. Laurence et S. Stich, 2005, p. 305-337.
- Dehaene, S. *La bosse des maths*, Paris, Odile Jacob, 1997.
- Faucher, L. « Les émotions morales à la lumière de la psychologie évolutionniste : le cas de l'évitement de l'inceste ». In *Morale et évolution biologique : entre*

3. Les travaux de psychologues comme Moutier et Houdé (2003) montrent qu'en augmentant la motivation et le contrôle dans la production de réponses, le taux de succès aux tâches comme celles de Wason augmentait au même niveau que lorsque celles-ci sont formulées en termes sociaux. Il y a donc de l'espoir !

- déterminisme et liberté, sous la direction de C. Clavien, Lausanne, Presses polytechniques universitaires romandes, 2007.
- Fessler, D. et C. D. Navarrette. « Third-Party Attitudes Toward Sibling Incest Evidence for Westermarck's Hypotheses », *Evolution and Human Behavior*, 25, 2004, pp. 277-294.
- Griffiths, P. E. et K. Stotz. « How the Mind Grows : A Developmental Perspective on the Biology of Cognition », *Synthese* 122, n° 1-2, 2000, p. 29-51.
- Haselton, M. G., D. Nettle et P. W. Andrews. « The Evolution of Cognitive Biases ». In *The Handbook of Evolutionary Psychology*, sous la direction de D. Buss, Hoboken, N.J., John Wiley and Sons, 2005, p. 724-746.
- Hardin, C. L. « Color and Illusion ». In *Mind and Cognition : A Reader*, sous la direction de W. Lycan, Cambridge, Blackwell, 1990, p. 555-566.
- Heinrich, J. et N. Heinrich. « Culture, Evolution and the Puzzle of Human Cooperation », *Cognitive Systems Research*, 7, 2006, 220-245.
- Lieberman, D., J. Tooby et L. Cosmides. « Does Morality Have a Biological Basis ? An Empirical Test of the Factors Governing Moral Sentiments Relating to Incest », *Proceedings of the Royal Society of London*, série B, 270, 2003, pp. 819-82.
- McCauley, Robert et Joseph Henrich. « Susceptibility to the Muller-Lyer Illusion, Theory-Neutral Observation, and the Diachronic Penetrability of the Visual Input System », *Philosophical Psychology*, vol. 19, n° 1, 2006, 1-23.
- Moutier, S. et O. Houdé. « Judgment Under Uncertainty and Conjunction-Fallacy Inhibition Training », *Thinking & Reasoning*, 9, 2003, 185-201.
- Nichols, S., S. Stich et J. M. Weinberg. [En ligne] [Meta-Skepticism : Meditations in Ethno-Epistemology]. In *The Skeptics*, sous la direction de S. Luper, U. K Aldershot, Ashgate Publishing, 2003, p. 227-247.
- Nisbett, R. E., K. Peng, I. Choi et A. Norenzayan. « Culture and Systems of Thought : Holistic vs. Analytic Cognition », *Psychological Review*, 108, 2001, 291-310.
- Odling-Smee F. J., K. N. Laland et M. W. Feldman. *Niche Construction. The Neglected Process in Evolution*, Monographs in Population Biology, Princeton University Press, 2003.
- Stich, S. « Philosophie et psychologie cognitive ». In *La philosophie cognitive*, sous la direction de E. Pacherie et J. Proust, Paris, Éditions Ophrys, 2004, p. 55-70.

Réponses à mes critiques

Murray Clarke

Volume 34, numéro 2, automne 2007

URI : <https://id.erudit.org/iderudit/017427ar>

DOI : <https://doi.org/10.7202/017427ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Société de philosophie du Québec

ISSN

0316-2923 (imprimé)

1492-1391 (numérique)

[Découvrir la revue](#)

Citer ce document

Clarke, M. (2007). Réponses à mes critiques. *Philosophiques*, 34 (2), 385–402.
<https://doi.org/10.7202/017427ar>

Réponses à mes critiques

MURRAY CLARKE

Université Concordia

Robert Hudson, Michael Bishop et Luc Faucher ont trouvé plusieurs points de discussion intéressants et offert de nombreux éclaircissements dans leurs commentaires sur mon livre *Reconstructing Reason and Representation* (MIT, 2004). Je leur suis redevable ainsi qu'au directeur de *Philosophiques* de cette occasion de clarifier ma perspective. Je répondrai séparément à chaque auteur. Je commence avec Robert Hudson.

A. Hudson

La discussion éclairante de Hudson est centrée sur l'aspect fiabiliste de la justification épistémique. Il commence en affirmant que je désire intégrer une dimension externaliste aux théories internalistes de la justification épistémique. Regrettablement, je ne sais pas où il a trouvé cette idée dans mon livre. J'ai offert une perspective sur la connaissance purement externaliste, fiabiliste et modulaire dans le chapitre six du livre. J'ai aussi offert une analyse de la connexion entre les aspects externaliste et internaliste de la justification épistémique dans le chapitre cinq. Dans ce chapitre, je suggère que le projet mélioratif d'offrir des conseils épistémiques à l'agent cognitif, sur la base d'une mise à jour rationnelle de ses croyances et à la lumière de données nouvelles, peut être mené, d'un côté, en étant conscient de la littérature empirique sur le raisonnement humain en psychologie et, de l'autre côté, en ayant clairement à l'esprit un aspect non mélioratif de la justification et de la connaissance. La justification non méliorative concerne les caractéristiques permettant aux croyances d'être de bonnes croyances et qui leur fournissent la justification épistémique, le type de justification qui fait que de telles croyances sont probablement vraies. Un tel aspect non mélioratif de la justification devrait, ai-je soutenu, être externaliste par nature. Le point n'était pas d'intégrer un élément externaliste dans un aspect internaliste de la justification mais de clarifier les connexions conceptuelles entre ce que je crois être deux types discrets de justification. Mon avis est qu'un aspect externaliste de la justification épistémique est nécessaire pour la connaissance alors qu'un aspect internaliste de la justification épistémique n'a pas de place dans une contribution à la connaissance. Le sens internaliste de la justification est plutôt, en fait, porté vers la poursuite de l'enquête. Dans cette perspective, l'enquête n'a pas à voir avec la connaissance *per se* mais avec les normes publiques de satisfaction, c.-à-d. les standards disciplinaires que les pairs partagent collectivement mais implicitement. Par exemple, les standards implicites adoptés par les journaux scientifiques pour l'acceptation des articles constituent une exemplification de telles normes publiques et illustrent les standards internalistes de justification épistémique. Dans ma perspective, c'est

une confusion de suggérer que nous devrions intégrer des éléments externalistes à un aspect internaliste de la justification parce que les racines des projets internaliste et externaliste sont dirigées vers des questions épistémiques différentes. Ayant ces rappels à l'esprit, je vais maintenant m'attarder aux affirmations plus spécifiques que Hudson a formulées sur le problème de la généralité.

Le problème de la généralité concerne le degré d'ouverture des types de processus en question. Comme je l'ai noté dans le livre :

Goldman affirme qu'une croyance est justifiée si et seulement si elle est causée par un processus fiable. Un processus fiable est un processus qui tend à produire la vérité. Le problème de la généralité peut être posé comme un dilemme. Si nous interprétons les types de processus trop étroitement alors un processus fiable risque d'avoir une seule instance ; toutes les croyances vraies seront formées de façon fiable et toutes les croyances fausses ne seront pas formées de façon fiable. Telle est la facette du « cas singulier » du dilemme. D'un autre côté, si nous interprétons le type de processus trop largement, alors la facette « sans distinction » du dilemme surgit. Par exemple, la perception visuelle doit être pensée comme un type de processus. Mais les croyances à propos des vaches qui surviennent lorsque celles-ci sont à dix pieds semblent être bien plus justifiées que les croyances à propos des vaches qui surviennent à un demi-mile de distance. Toutefois, la contribution du processus fiable semble n'avoir aucun moyen de marquer cette importante distinction. Apparemment, Goldman et les autres théoriciens du processus fiable n'ont aucune façon de spécifier, en principe, le degré de généralité des types de processus (Conee et Feldman, 1998 ; Clarke, 2004, 85-86).

J'ai suggéré dans le livre que ma perspective ne rencontre pas le problème de la généralité a priori et le problème conceptuel parce que les types de processus sont déterminés empiriquement par des psychologues évolutionnistes et d'autres scientifiques. Les modules cognitifs représentent les types de processus qui donnent naissance aux croyances bien formées. J'ai pointé les modules darwiniens de raisonnement comme exemple central de tels types de processus. Hudson pense que ce transfert de la responsabilité conceptuelle aux psychologues évolutionnistes ne marchera pas. Comme il le note :

Cette réponse aurait été excellente s'il existait une voie empirique toute tracée pour isoler les modules — mais il n'y en a pas. Pour identifier les modules, il faut identifier les comportements propices à l'évolution qui sont observables chez les organismes et inférer alors l'existence d'un module qui permet, de par sa structure, ce comportement.

Toutefois, il poursuit :

Je voudrais soutenir que cela revient à présumer une solution au problème de la généralité, non pas à y répondre. Parce que le problème précis qui nous intéresse est la mesure par laquelle nous pouvons déterminer quand un processus générant des croyances est fiable. Le problème de la généralité suggère que cette évaluation de la fiabilité doit attendre une détermination antérieure du niveau approprié de description pour ce processus — et rien de ce qu'a dit Clarke ne supprime ce réquisit.

En d'autres mots, Hudson affirme ici que, par définition, déterminer le degré de généralité d'un type de processus doit être accompli a priori. D'où ce qui compte comme un type de processus (qui, dans ce cas, équivaut à un type de module) peut uniquement être déterminé par une analyse conceptuelle a priori. Puisque c'est ainsi, nous allons informer les psychologues évolutionnistes que les philosophes doivent d'abord déterminer ce qui peut compter comme un type de module a priori avant qu'ils commencent leurs expérimentations sur le raisonnement humain. Tout le travail conceptuel doit être effectué prioritairement à l'accomplissement de l'enquête empirique sous peine de circularité. Je ne trouve pas cette approche de l'investigation empirique plausible. Les scientifiques devraient attendre le travail conceptuel des philosophes chargés de les éclairer afin que leur travail puisse procéder rapidement. Je doute qu'une science ayant connu un certain succès ait déjà procédé de cette façon ou qu'elle procède jamais de cette façon ou, encore, que si elle procède effectivement de cette façon, le succès de cette science soit redevable à un tel travail.

Naturaliser l'épistémologie requiert une méthode qui diffère de l'analyse conceptuelle traditionnelle. Exposer une telle méthode était un objectif central de mon livre. Je suis satisfait d'adopter une telle approche novatrice en philosophie mais je comprends que Hudson ne soit pas aussi satisfait que moi. Pour voir le problème avec cette approche, demandez-vous, est-ce que Cosmides et Tooby, ou d'autres psychologues évolutionnistes, ont produit les résultats importants et intéressants qui continuent de marquer leur travail s'ils avaient adopté la méthode conceptuelle a priori de Hudson ? La réponse, je crois, est un « NON » retentissant et résonnant. À mon avis, c'est une bénédiction que les philosophes ne dictent pas les méthodes appropriées aux scientifiques. Je tiens pour acquis que discerner ce qui constitue un module ou un type de processus est difficile et impossible sans le type de travail continu et soigneux qui caractérise la recherche de scientifiques comme Cosmides et Tooby. Après tout, ce n'est pas comme s'ils affirmaient allègrement l'existence de nouveaux modules tous les jours de la semaine. Mais imaginer que de tels types de modules puissent être déterminés a priori, ou déterminer le degré de généralité que de tels types peuvent posséder a priori tend vers la crédulité. Aucune réussite scientifique n'a, à ma connaissance, été précédée de façon notable du travail conceptuel des philosophes, et en reposant d'abord sur lui. Comme je la comprends, la tâche du philosophe des sciences qui est intéressé par la méthodologie est de tenter de comprendre la méthodologie scientifique, et non pas de prescrire une telle méthodologie. Je suggérerai que le taux de réussite des philosophes tentant de pourvoir des solutions prescriptives aux scientifiques est comparable avec le taux de réussite de ceux qui prédisent que le marché boursier sera favorable.

Dans ce contexte, la suggestion de Conee et Feldman selon laquelle une chose comme des conditions nécessaires et suffisantes est requise pour spécifier l'importance des types de processus est problématique. Une telle requête

découle d'une conception des concepts, c.-à-d. le définitionnisme, qui a été rejetée il y a longtemps par la psychologie (voir Fodor, 1998 ; Margolis et Laurence, 1999). La chasse à de telles conditions nécessaires et suffisantes représente le désir utopique de précision là où une telle précision ne peut exister et n'est même pas nécessaire à une bonne science. Dans leur article *The Generality Problem for Reliabilism*, Conee et Feldman mènent ce désir presque obsédant de trouver de telles conditions à sa conclusion logique — qui est à couper le souffle. Après avoir rejeté une variété de solutions scientifiques au problème de la généralité, Conee et Feldman concluent que « bien que la science fournisse des outils pour sélectionner les candidats aux types pertinents, il n'existe pas de bonnes raisons de penser que les classifications scientifiques fournissent des outils pour résoudre le problème de la généralité » (Conee et Feldman, 1998). Il en est ainsi parce que, même pour le cas simple de la perception visuelle, ce que sont les types d'explications psychologiques pertinents n'est pas clair, et la question de savoir comment ces types peuvent être utiles pour les théories fiabilistes de la justification ne l'est pas non plus. En général, Conee et Feldman doutent qu'il n'existe qu'une seule et unique explication psychologique pour chaque croyance. Les tentatives d'Alston, Goldman et Baergen pour spécifier ce qui compte comme type de processus demeurent en attente de solutions. Les deux auteurs concluent que les théories fiabilistes de la justification et de la connaissance semblent sans espoir puisque le problème de la généralité est insoluble. Je devrais ajouter que les explications psychologiques ont l'air extrêmement difficiles, sinon impossibles à réaliser de leur perspective et, pour cette raison, la science est rendue inapte à répondre à ses propres questions. Mais, à coup sûr, une erreur s'est glissée ici. Comme je l'ai suggéré plus tôt, le type de solution définitionniste au problème de la généralité que Conee et Feldman recherchent est utopique. Étant donné les standards irréalistes qu'ils proposent pour solutionner adéquatement le problème de la généralité, il n'est pas surprenant que toutes les propositions soient considérées comme des échecs et que la science elle-même demeure en attente de solutions. Mais pourquoi suivre ceux qui trouvent que ces standards, imposés par Conee et Feldman ou ceux qui sont prêts à rejeter des parties significatives de la science dans l'optique de satisfaire ces standards apparemment irraisonnables, sont raisonnables ? Ce sont exactement de tels standards (ou ce que j'appelle plus généralement dans mon livre des « moulins à vent scolastique ») qui ont occasionné une quasi-absence de progrès en épistémologie. Le résultat est que le travail des épistémologues traditionnels, aussi légitime qu'il puisse paraître, est à mon avis largement improductif et franchement inutile à la fois à la communauté scientifique et à de larges sections de la communauté philosophique. C'est sur ce point que j'ai mis l'accent dans mon ouvrage. En tous cas, je m'occupe à présent des commentaires incisifs de Michael Bishop sur le livre.

B. Bishop

Bishop a soulevé, de manière articulée, trois problématiques dans sa revue critique : ma défense de la modularité massive, ma solution partielle au problème de la disjonction, et ma proposition selon laquelle la connaissance est fragmentée. Je le suivrai en m'occupant à mon tour de ces problèmes.

1. La modularité massive

Bishop a correctement remarqué que, dans le livre, j'ai soutenu que l'hypothèse de la modularité massive est liée à l'idée que l'esprit consiste essentiellement en processeurs darwiniens possédant un domaine spécifique qui agit sur des ensembles de données chomskyens dont le domaine est spécifique. C'est pourquoi il doit exister certains processeurs dont le domaine général agit sur des ensembles de données qui ont un domaine général ou spécifique, ou des processeurs darwiniens qui ont un domaine spécifique agissant sur un ensemble de données qui ont un domaine général. Je réfère à ces trois dernières possibilités comme étant les cas D, B et C dans la page six du livre. Le reproche de Bishop, emprunté à Richard Samuels, est qu'il n'y a pas de raison de préférer l'idée que l'esprit consiste essentiellement en modules ayant des domaines spécifiques plutôt qu'en processeurs ayant des domaines généraux agissant sur des connaissances dont le domaine est spécifique, étant donné les effets liés au contenu dans la tâche de sélection de Wason.

D'où même si Cosmides et Tooby ont raison de penser que l'esprit ne peut pas consister seulement en processeurs ayant un domaine général qui agissent seulement sur l'information perceptuelle, ils ont tort de penser que l'esprit ne peut pas consister seulement en processeurs ayant un domaine général qui agissent sur des ensembles de données dont le domaine est spécifique (ou ce que Samuels appelle le « Modèle bibliothèque de la cognition »). Cosmides et Tooby n'ont pas réussi à évacuer cette dernière possibilité. Dans mon livre, au deuxième chapitre, j'ai soulevé les inquiétudes de Samuels et Fodor. Dans ce chapitre, j'ai présenté trois arguments de Cosmides et Tooby provenant de leur article de 1994, *Origins of Domain-Specificity: The Evolution of Functional Organization*. J'ai soutenu que les critiques de Fodor concernant ces trois arguments n'ont pas réussi à atteindre Cosmides et Tooby principalement parce qu'il interprète leurs arguments de façon systématiquement erronée. Je ne peux pas répéter tout ce que j'ai dit dans le deuxième chapitre, mais je dirai les choses suivantes. Cosmides et Tooby affirment que leurs trois arguments sont un moyen de soutenir l'idée qu'il est impossible en principe que la psychologie humaine qui ne contient rien d'autre que des mécanismes ayant un domaine général ait pu évoluer parce que :

- a. Ce qui est considéré comme un comportement adaptatif diffère de domaine en domaine ; il n'existe donc pas de critères du succès du domaine général qui soit corrélé avec l'adaptativité.

- b. Les trajectoires d'adaptation des actions ne peuvent être déduites ni apprises par des critères généraux, parce qu'elles dépendent de relations statistiques entre les traits caractéristiques de l'environnement, le comportement et l'adaptativité qui émergent durant de nombreuses générations et sont, par conséquent, inobservables pendant la durée d'une seule vie.
- c. L'explosion combinatoire paralyse tout système ayant véritablement un domaine général lorsqu'il est confronté à la complexité de la réalité (Cosmides et Tooby, 1994, 91).

La possibilité que Bishop met de l'avant, c.-à-d. des processeurs ayant un domaine général qui agissent sur le modèle bibliothèque de l'esprit évite ces objections parce que les ensembles de données innées qui ont un domaine spécifique peuvent aider à résoudre les questions spécifiques du problème de la réussite, le problème de la relation statistique transgénérationnelle et le problème de la complexité du monde réel. Cependant, Fodor s'objecte au deuxième argument dans ses propres termes, et ses objections, comme je l'ai souligné dans le livre, sont erronées. Aussi ai-je souligné dans mon livre que Fodor a raison de penser que la troisième objection n'est pas satisfaisante si l'on adopte la perspective de Bishop-Samuels sur la cognition. Bien sûr, Cosmides et Tooby ont un objectif différent en ce qui concerne le troisième argument : l'idée que seuls les processeurs ayant un domaine général se combinent avec l'information perceptuelle. En général, ils pensent qu'il est invraisemblable que l'évolution ait produit des processeurs ayant un domaine général parce que l'évolution est, approximativement, une opération de bricolage. C'est-à-dire qu'effectivement l'évolution agit sur les matériaux dont elle dispose : de manière récurrente, nous obtenons des solutions locales à des problèmes immédiats. C'est pourquoi Cosmides et Tooby insistent sur l'idée de l'ingénierie à rebours. Il faut essayer d'imaginer les tâches auxquelles les chasseurs-cueilleurs du Pléistocène ont été confrontés afin de comprendre la fonction des modules actuels. Cette approche est parfaitement raisonnable, même évidente, compte tenu de l'historique des discussions sur l'évolution en philosophie de la biologie. À l'inverse, l'idée de processeurs globaux ayant des domaines généraux qui auraient été sélectionnés afin de résoudre n'importe quel problème serait peu vraisemblable, même panglossienne, étant donné tout ce que l'on sait sur l'évolution. Donc, même s'il est impossible d'écarter a priori, sur une base logique, l'idée que seuls les processeurs ayant un domaine général agissent sur un ensemble inné de données ayant un domaine spécifique, ce n'est pas, en général, ce que les biologistes qui étudient l'évolution pensent être le cas. C'est parce que la bibliothèque innée de connaissances de Fodor-Samuels, qui a un domaine spécifique, aurait à encoder une masse colossale de données. Les exigences de stockage seraient, proportionnellement, monstrueuses. Et, par conséquent, les considérations de simplicité militent contre la probabilité que cette perspective soit vraie. En fait, nous avons constaté le même

problème avec la perspective du concept radical de nativisme chez Fodor (1975) où quelque 500 000 concepts étaient déclarés comme innés. Fodor a peut-être raison à propos de l'innéisme de nos concepts, mais je n'en attendrai certainement pas la confirmation et presque personne dans la communauté des sciences cognitives non plus. (Bien sûr, Fodor a sagement réorienté cette perspective dans son livre de 1998, *Concepts*). Donc, il existe une raison générale et importante, bien que n'étant pas décisive, de préférer l'idée de processeurs darwiniens ayant un domaine spécifique. D'autres raisons convaincantes de préférer l'image de l'esprit selon Cosmides et Tooby se trouvent dans les résultats complexes et détaillés concernant les effets liés au contenu dans la tâche de sélection de Wason. Cosmides est justement reconnu pour avoir soutenu ses opinions contre l'explication en terme de domaine général de la tâche de sélection de Wason. J'explique de nombreux exemples de son travail dans le livre et je ne reviendrai donc pas sur les données ici. Il est certainement quelque peu étrange de suggérer, comme Bishop le fait, qu'il n'y a aucune raison de penser que les processeurs darwiniens ayant un domaine spécifique existent et dominent peut-être l'esprit à la lumière de l'ensemble substantiel d'expériences menées par les psychologues évolutionnistes durant ces années. Pour être convaincant, Bishop aurait besoin de fournir des informations empiriquement détaillées qui contrediraient le travail de Cosmides et Tooby sur la tâche de sélection de Wason. À ma connaissance, Samuels et Fodor n'ont effectué aucun travail expérimental sur la tâche de sélection de Wason ! Brièvement, alors que le modèle bibliothèque demeure une option logiquement cohérente, cette tâche requiert plus de soutien empirique avant de provoquer notre assentiment. Il devrait être aussi mentionné que Fodor, dans son livre *The Mind Doesn't Work That Way*, a accusé Cosmides et Tooby d'argumenter que l'esprit consiste seulement en processeurs darwiniens ayant un domaine spécifique. Comme je l'ai souligné dans mon ouvrage, il s'agit clairement d'une mauvaise interprétation de leur pensée.

2. Le problème de la disjonction

Bishop affirme que la théorie du fossé ne peut expliquer comment un module qui a été sélectionné peut faire une erreur de représentation quant à l'état du monde lorsqu'il fonctionne dans son domaine propre. Comme il le note :

Considérez un environnement où l'évolution a sélectionné chez les prédécesseurs des castors le module qui mène ces prédécesseurs à claquer leur queue. Pendant les (présumées) milliers d'années durant lesquelles le module a été sélectionné, il s'est sûrement déclenché en présence d'un animal qui n'était pas un prédateur, disons un ongulé végétarien. Ce cas est-il un exemple d'erreur de représentation ? Il semble que Clarke soit forcé à le nier puisqu'il n'y a pas de fossé entre le domaine propre et actuel.

Bishop, dans ce passage, néglige un point crucial à propos des modules ; nommément, l'idée d'un module sélectionné résulte du fait que les modules sont fiables, non pas infaillibles. Ma définition de la fiabilité des modules (FM)

dit que : « Un module est fiable si et seulement s'il produit habituellement et de façon réussie des résultats dans le domaine actuel [...] et il devrait habituellement produire des résultats satisfaisants dans des domaines possibles [...] qui sont similaires, de façon pertinente, au domaine actuel. » Il s'ensuit que les erreurs de représentation peuvent surgir même quand le domaine propre est le domaine actuel. Pourquoi ? Les erreurs de ce type ne sont pas expliquées par la théorie du « fossé » mais simplement par le fait que les modules qui sont sélectionnés ne sont pas infaillibles, seulement fiables. Si j'avais soutenu que la sélection naturelle a sélectionné des modules infaillibles dans leur domaine propre, alors Bishop aurait émis une bonne remarque. Mais, manifestement, personne ne voudrait défendre une telle conception panglossienne de l'évolution. Cela ne constitue pas une surprise pour les théoriciens de l'évolution que celle-ci produit des solutions qui ne sont pas parfaites. Prenez la perception visuelle. La plupart des cognitivistes s'entendraient sur l'aspect modulaire et fiable de la vision. La perception visuelle mène-t-elle à des erreurs de représentations à l'intérieur de son domaine propre ? La réponse est « oui ». Penser autrement serait peu plausible. Évidemment, la théorie du fossé n'est pas une tentative d'explication de ces cas ni une tentative d'explication des erreurs non modulaires. Mais cela n'est pas surprenant puisqu'elle n'a pas été constituée pour expliquer ces types d'erreurs. Visiblement, la notion de domaine propre a quelque chose de l'idéalisation du point de vue évolutionniste. Elle implique ces conditions où une fonction propre est sélectionnée, où la fonction propre d'un module est, typiquement, accomplie avec succès. Mais l'exécution réussie d'une fonction propre n'arrive pas toujours dans le EAE. Parfois, les conditions de l'EAE empêchent la fonction propre d'un module de se réaliser. Lorsqu'il n'y a pas de fossé entre le domaine propre et le domaine actuel, il est presque tautologique de souligner que la « théorie du fossé » ne sera pas à même d'expliquer comment les erreurs de représentation peuvent advenir. En bref, j'accepte l'affirmation de Bishop selon laquelle les erreurs de représentation surviennent même à l'intérieur du domaine propre (en fait, je l'ai affirmé dans le livre) mais je refuse que cela constitue une objection à la théorie du « fossé » étant donné que les modules sélectionnés ne sont pas infaillibles.

3. Les modules fiables et la connaissance

Bishop dédie sa troisième section à ce que j'ai dit à propos des croyances vraies et de la connaissance. Je commençais mon quatrième chapitre en soulignant que je trouve absurde de supposer qu'un ensemble de croyances généralement fausses constitue un avantage pour les humains ou que les croyances vraies puissent amoindrir l'adaptativité inclusive des organismes. Bishop doute que Stich souscrive à de telles affirmations. Comme il le note : « Ce que défend Stich est que l'évolution a pu sélectionner des mécanismes inférentiels faillibles qui commettent beaucoup d'erreurs mineures peu coûteuses en échange de ne (presque) jamais faire d'erreurs majeures très coûteuses. » J'ai explicitement accepté ce point soulevé par Stich lorsque j'ai noté dans le livre que : « Le fait

de croire beaucoup de faux positifs peut être plus sûr que de croire beaucoup de faux négatifs, ce qui rendrait possible que certains systèmes fiables d'inférence ne sont pas favorisés par la sélection naturelle. La question intéressante, toutefois, est de savoir si une telle possibilité est susceptible d'advenir. C'est-à-dire : étant donné tout ce que nous savons à propos de la cognition humaine, est-il physiquement possible que l'évolution ait produit des systèmes fiables d'inférence ? (Clarke, 76-77). Mon affirmation était que Stich se concentrait sur la mauvaise question en focalisant sur ce qui « aurait bien pu » arriver et, par conséquent, j'ai essayé d'orienter la discipline dans une direction plus productive. Bishop a remarqué qu'un de mes arguments était structuré de manière similaire quant à la possibilité de processus fiables qui « auraient pu » être sélectionnés et il trouve cela ironique. Cela dit, est-ce que je ne commets pas l'erreur même que j'accuse Stich de commettre ? Ma réponse est « Non ». En fait, l'ironie était intentionnelle. J'ai essayé de renvoyer à Stich le problème de son argument, d'une façon légèrement différente, en imitant la forme de son argument avec un argument de même forme.

Bishop indique qu'il existe un ensemble significatif de données produit par les psychologues évolutionnistes et suggérant que de nombreux modules pointent vers la production de croyances fausses. Ces modules produisent des croyances optimistes chez ceux qui en sont dotés et, apparemment, de telles croyances influencent positivement la santé, la survivance, et ne sont pas faciles à contrecarrer. Je ne doute pas de la vérité de ces affirmations, mais je doute que cette donnée « représent[e] un sérieux défi à l'optimisme évolutionnaire de Clarke concernant nos habiletés cognitives », comme Bishop l'a dit. Ma confiance en nos habiletés cognitives concerne les croyances indexicales à propos de l'espace et du temps où l'action est envisagée. Je pense qu'il était important pour nos ancêtres et pour nous de pouvoir reconnaître les prédateurs qui menacent notre existence, d'identifier la nourriture comestible, etc. Je ne doute pas que les humains sont capables de toutes sortes de fausses croyances, de second ordre, à propos d'eux-mêmes, mais cela n'est pas pertinent pour ce que je défends dans le livre. Ce point ressurgira dans la section sur Faucher.

Finalement, Bishop soutient que ma perspective faisant de la connaissance un ensemble de genres naturels intégrés dans les modules d'un esprit massivement modulaire mène à une conception de la connaissance qui semble inapte à être théoriquement utile. La raison qui le motive est que, puisque le module qui justifie une croyance peut changer au cours du temps, il s'ensuit que la connaissance comme genre naturel changera elle aussi. Cela parce que les genres sont individués selon les trois conditions de Richard Boyd : ils sont des ensembles de propriétés homéostatiques, ils affichent des effets standards observables, et des micropropriétés provoquent des propriétés observables à travers les lois causales. Puisque chaque genre sera le résultat de micropropriétés distinctes causant des effets observables à travers des lois causales distinctes, il en découle qu'une connaissance distincte, en termes de genres naturels, émergera. Bishop pense que les philosophes sont obsédés, de manière inappropriée,

par la connaissance, et il se réjouit donc de ce qu'il croit être ma dérogation à la connaissance puisque la fragmentation de la connaissance que je défends rend la connaissance trop éclatée pour être théoriquement utile. J'admets que la connaissance est fragmentée dans ma perspective et que cela rend les choses plus complexes qu'on pourrait le croire. Mais je doute que cela prouve qu'elle n'est pas théoriquement utile. Le fait que la connaissance soit fragmentée à travers une pléthore de modules cognitifs rend plutôt la connaissance susceptible d'être empiriquement investiguée. Nous pouvons à présent étudier la connaissance d'une manière véritablement scientifique, et non pas comme un concept à clarifier par une analyse *a priori*, mais comme un phénomène empiriquement découvert. Je trouve cette perspective stimulante, toutefois, elle requiert l'abandon du concept de connaissance. (Évidemment, le fait que Bishop pense que j'ai offert un concept de connaissance plutôt qu'une recherche d'un phénomène empirique indique une mauvaise compréhension de la nature de mon projet.) L'étude de la connaissance sera théoriquement aussi utile que l'est n'importe quelle étude des autres genres naturels. Cette étude ne manquera d'intérêt, ou ne sera théoriquement inutile que pour ceux qui veulent poursuivre avec l'analyse conceptuelle. Mais, étant donné l'approche naturaliste, avouée, de Bishop en philosophie, je suis surpris qu'il ne trouve pas la perspective de découvrir que la connaissance est un ensemble de genres naturels théoriquement utiles, et peut-être même philosophiquement libérateurs.

C. Faucher

Le commentaire de Luc Faucher est centré autour de la littérature récente en sciences cognitives. Il commence en affirmant que je vise à défendre la contribution de Cosmides et Tooby, Pinker et Buss à la psychologie évolutionniste. Bien sûr, j'affirme seulement que ma contribution reflète fidèlement la position de Cosmides et Tooby, pas celle de Pinker et Buss. Deuxièmement, ma position va au-delà de ce qu'affirment Cosmides et Tooby en ce que, par exemple, je soutiens que les inférences non modulaires et abductives existent. Cosmides et Tooby n'ont jamais abordé, à ma connaissance, les inférences abductives. En dépit de ces omissions, Faucher a raison d'associer ma contribution avec celle de Cosmides et Tooby. Il suggère que ma perspective sur la psychologie évolutionniste est, à différents égards, trop éloignée, et trop rapprochée, de la psychologie évolutionniste. Cela suggère que les psychologues évolutionnistes s'entendent sur l'objet, les théories et les méthodes de leur discipline. Je suis moins sûr que Faucher qu'un tel programme unifié appelé « psychologie évolutionniste » existe. Cela dit, pour ce qui suit, je présupposerai que cette affirmation est juste.

Faucher déclare que ma contribution est trop éloignée de la psychologie évolutionniste en ce que les psychologues évolutionnistes souscrivent à une perspective notablement plus faible et pragmatiste de ce qui constitue une « croyance vraie », et il déclare que la façon dont les processus fiables produits par l'évolution possèdent des croyances vraies comme output n'est pas

très claire. Dans de récents passages de Cosmides et Tooby cités par Faucher, il apparaît clairement qu'ils défendent à présent une conception de la vérité plus faible que celle qui était implicite dans leurs travaux sur la tâche de sélection de Wason. En fait, je ne suis pas certain qu'il existe une conception cohérente de la vérité qui puisse être tirée de ces citations. L'article en question de Cosmides et Tooby a été publié après mon livre, comme la plupart de la littérature citée par Faucher. Et par conséquent je ne peux être tenu responsable de n'avoir pas incorporé des informations qui n'étaient pas à ma disposition. La littérature sur le raisonnement que j'ai consultée ne contient pas de conception pragmatiste de la vérité. Je crois que quelque chose comme une théorie correspondantiste de la vérité est implicite dans la littérature sur la tâche de sélection de Wason, bien que ni Cosmides et Tooby ni moi-même aient discuté cette problématique. Un autre sujet important pour ma réponse à Faucher se rapporte à la question de savoir quelles classes de croyances sont considérées comme étant des output de processus fiables. La perspective que je défends est que les croyances indexicales précises ont souvent été le résultat de processus fiables, comme pour la perception visuelle et les bons raisonnements. Je pense que la perception visuelle a été sélectionnée chez nos ancêtres parce qu'elle rendait possible les croyances indexicales précises concernant (comme je l'ai indiqué plus tôt) ce qui devrait être légitimement considéré comme des proies et des prédateurs, ce qui constituerait des aliments nourrissants, et ainsi de suite. C'est-à-dire qu'il devait être important pour nos ancêtres de satisfaire les quatre types biologiques de besoins pour la nourriture, la fuite, le combat et la reproduction. Puisque la survivance est une condition nécessaire à la reproduction, et que la reproduction est un objectif central de l'évolution, il s'ensuit que les mécanismes qui facilitent la survivance, ou le comportement adéquat, seront indirectement sélectionnés. C'est-à-dire, indépendamment de ce que sont nos objectifs, évolutifs ou non évolutifs, connaître la vérité rendra la satisfaction de ces objectifs plus à même d'advenir. Et, par conséquent, quant à la vérité je suis un pragmatiste au moins en un sens : faire appel à la vérité est le moyen le plus utile pour garantir toutes les fins que nous désirons atteindre (voir Kornblith, 2002). (Cependant, je pense assurément qu'une certaine version de la théorie correspondantiste de la vérité est correcte.) Ce dont nous avons besoin, surtout, c'est une vérité fiable, ou une croyance vraie produite de manière fiable. Ma proposition est que les adaptations fonctionnelles complexes, ou modules, comme la perception visuelle, qui ont été sélectionnées présentent toutes une tendance à produire des croyances indexicales précises, c.-à-d. des croyances vraies à propos d'ici et maintenant. De cette façon, la connaissance perceptuelle est le résultat d'un module de perception visuelle fiable. C'est, bien sûr, entièrement possible et logiquement conséquent avec mon affirmation selon laquelle les autres croyances non indexicales devraient être systématiquement, ou non systématiquement fausses. Par exemple, il existe un effet de supériorité par rapport à la moyenne (Alicke 1985) en psychologie empirique suggérant que la plupart des humains surestiment les traits subjectifs

comme leur intelligence, leur honnêteté, leur persévérance, leur originalité, leur amabilité et leur fiabilité. Comme McKay et Dennett le suggèrent :

La plupart des étudiants au collège tendent à croire qu'ils vont avoir une espérance de vie supérieure à la moyenne, alors que la plupart des instructeurs au collège croient qu'ils sont des professeurs meilleurs que la moyenne (Cross, 1977). La plupart des gens tendent aussi à croire que leurs habiletés de conducteur sont supérieures à la moyenne — même ceux qui ont été hospitalisés à la suite d'un accident (exemple de McKenna, Stanier & Lewis, 1991 ; Williams, 2003). En fait, la plupart des gens se croient meilleurs que la moyenne dans presque toutes les dimensions qui sont subjectives et socialement désirables (Myers, 2002) [...] Les illusions positives peuvent bien être répandues, mais sont-elles adaptatives, évolutionnairement parlant ? Taylor et al. (2000) ont fait la conjecture que les illusions positives doivent exercer leurs bienfaits médicaux en régulant les réponses physiologique et neuroendocrinienne aux circonstances stressantes. Les activations induites par le stress sur le système nerveux autonome et l'axe hypothalamo-hypophysaire-surrénalien (HHS) facilitent les réponses de type « combat ou fuite » et sont donc adaptatives à court terme. Les activations chroniques ou récurrentes de ces systèmes, cependant, peuvent être préjudiciables à la santé (voir McEwen, 1998), donc, les mécanismes psychologiques qui contraignent l'activation de tels systèmes peuvent être bénéfiques. En accord avec l'hypothèse ci-dessus, Taylor et al. (2003) ont déterminé que les pensées auto-valorisantes chez les adultes sains ont été associées avec des réponses cardiovasculaires moins élevées au stress, un rétablissement cardiaque plus rapide et des niveaux de base de cortisol inférieurs (McKay and Dennett, 2007).

Mc Kay et Dennett suggèrent que les croyances erronées évoluent dans de tels cas. C'est-à-dire que généralement la croyance vraie est adaptative, mais, en de rares circonstances, les croyances systématiquement erronées peuvent être adaptatives. Il est important de voir que ma position dans le livre concerne les croyances indexicales, et non pas les croyances de second ordre à propos de soi-même auxquelles McKay et Dennett font référence. En ce qui concerne de telles croyances de second ordre, il existe exactement trois possibilités. Option un : la croyance systématiquement fausse est adaptative ; option deux : la croyance fausse est toujours non épistémique et non pertinente pour l'adaptation ; option trois : la croyance non systématiquement fausse peut apparaître et est adaptative. Je tiens l'option trois pour intenable parce que la croyance non systématiquement fausse ne peut être sélectionnée ; des effets aléatoires ne peuvent mener à la sélection de mécanismes qui produisent de tels effets. Au mieux, de tels effets seraient des « cavaliers seuls » [*free-riders*]. Ce qui laisse ouvertes la première et la seconde possibilité. La théorie de la gestion des erreurs (TGE) ; (Haselton à venir ; Haselton & Buss, 2000, 2003 ; Haselton et Nettle, 2005) vise à montrer que « les erreurs cognitives (y compris les fausses croyances) ne sont pas nécessairement des dysfonctionnements reflétant les limites du design évolutionnaire ; plutôt, de telles erreurs reflètent des biais systématiques judicieux qui maximisent l'adaptabilité en dépit de l'augmentation globale du taux d'erreurs » (McKay et Dennett, 2007). Et

par conséquent la TGE appuie l'option un. L'exemple de la nourriture toxique de Stich, provenant de *The Fragmentation of Reason*, constitue ce genre d'exemples où il existe des asymétries fiables dans le coût des erreurs, c.-à-d. quand un ensemble de croyances largement fausses améliorent davantage l'adaptabilité qu'un ensemble de croyances largement vraies, où les croyances fausses importent peu dans le premier système mais où elles sont fatales dans le deuxième. En général, la littérature sur les illusions positives, y compris l'effet de supériorité par rapport à la moyenne, appuie l'option un. L'exemple de l'évitement de l'inceste de Faucher tombe aussi dans cette catégorie. L'option deux, alternativement, devrait suggérer que les croyances fausses qui ne sont pas systématiques ne sont pas pertinentes pour l'évolution précisément parce qu'elles ne sont pas systématiques et, comme telles, ne peuvent pas être sélectionnées. Cette position est cohérente et pourrait même s'avérer juste. Les humains possèdent peut-être simplement tous les types de croyances fausses qui ont été obtenues de manière non systématique. De telles croyances devraient seulement constituer un bruit d'arrière-fond ou des façons plaisantes d'occuper la conscience de tout un chacun. Peut-être que de telles croyances sont simplement des parties d'un « spectacle en cours » pour un esprit quinqué : les croyances fausses, mais inoffensives, qui sont le résultats d'effets liés au cavalier seul. Alternativement, de telles fausses croyances doivent constituer des tentatives humaines sincères d'effectuer des contributions artistiques ou autres contributions culturelles à la société. Je ne pense pas que le verdict soit posé sur cette problématique. Surtout, indépendamment de la perspective appuyée par les données, c.-à-d., à savoir si l'option un ou deux est correcte, le résultat est conséquent avec la position que je défends dans mon livre. Cela est dû au fait que les croyances de l'option un et deux sont, largement, des croyances de second ordre sur soi-même ou sur des propriétés inobservables chez les autres. Les croyances en question ne sont pas les croyances indexicales, souvent perceptuelles, qui sont cruciales pour la survivance, le comportement adéquat et — souvent — l'évolution, auxquelles je fais référence dans mon livre.

Faucher consacre le reste de son commentaire à suggérer que ma perspective est trop étroitement liée aux vastes engagements métaphysique et méthodologique de la psychologie évolutionniste. Ces engagements résultent en une tendance à ignorer, ou minimiser, le rôle de la culture dans les explications de nos croyances par les psychologues évolutionnistes et moi-même. Évidemment, j'ai été très clair dans mon livre en disant que :

La vérité, dans cette perspective, est l'objectif central, bien que n'étant pas le seul objectif épistémique de la science et, en fait, de toutes les investigations qui sont guidées par la biologie. Bien sûr, les humains s'engagent dans de nombreuses poursuites qui ne sont pas basées sur la biologie mais sur des motivations culturelles. D'aucune façon je ne voudrais suggérer que ma contribution est une histoire exhaustive des humains ; mais j'espère qu'elle identifie une partie cruciale de cette histoire (Clarke, p. 150).

Et, par conséquent, je me réjouis des nouvelles et intéressantes discussions concernant la co-évolution gène-culture auxquelles réfère Faucher. Il déclare que je défends une perspective modulo-centriste de la psychologie où la culture ne serait que le vernis sur le cœur de mécanismes universels spécifiques à une espèce. De plus, je présente une « conception étroite des relations entre la culture et la cognition ». En acceptant ce paradigme, j'ai produit un livre qui résulte en une attention exclusivement portée vers les mécanismes psychologiques complexes du raisonnement, et mon « paradigme » me force à considérer le raisonnement abstrait comme un effet du cavalier seul.

Si la caractérisation que Faucher fait de mon livre était juste, alors je reconnaitrais avec lui que je suis en mauvaise posture. Heureusement, il se trompe. Premièrement, laissez-moi noter que le focus sur le raisonnement, par rapport aux nombreux autres sujets sur lesquels les psychologues évolutionnistes ont écrit, comme les stratégies d'accouplement et autres, est le résultat du fait que le sujet central de mon livre était la justification épistémique et la connaissance. Les résultats des études sur le raisonnement, à la fois évolutionniste et non évolutionniste, sont absolument cruciaux pour de telles problématiques épistémiques. Et, par conséquent, le focus sur les études de raisonnement n'avait rien à voir avec mon adoption de quelque « paradigme » kuhnien et tout à voir avec ce que je crois être la littérature pertinente pour s'occuper des problématiques centrales que je veux soulever ici. Cela dit, je pense certainement que les membres d'une espèce entreprennent leur vie avec un ensemble complexe d'adaptations fonctionnelles, comme un module pour la perception visuelle et une pléthore d'autres modules. Les humains sont dotés d'un module de réorientation spatiale et ils acquièrent, éventuellement, un module de théorie de l'esprit et beaucoup d'autres modules au cours du déploiement du programme développemental de l'individu. Avec Cosmides et Tooby, je souscris certainement à l'idée d'une nature humaine universelle basée sur l'existence de tels modules trans-spécifiques. Mais qui réfuterait sérieusement que la perception visuelle a été sélectionnée ? Ne dispose-t-on pas d'arguments solides pour appuyer l'idée d'un module de détection des tricheurs et un module de réorientation spatiale ? Je tiens pour acquis que Faucher ne cherche pas à remettre en cause l'idée que de tels modules existent. Et je peux accepter aisément l'idée que les humains emploient des éléments non biologiques pour compléter nos modes biologiques basiques de traitement. Je ne m'objecterai pas à appeler de tels éléments d'extension une extension de l'esprit pour autant qu'il s'agisse d'une métaphore puisque, manifestement, de tels ajouts ne sont pas littéralement une partie de notre machinerie initiale. À l'inverse, faire appel à des propriétés non géométriques de l'environnement, comme les couleurs d'un mur, les motifs sur une boîte ou les identités catégorielles d'un objet, dans le cas de la réorientation spatiale chez les humains, est un « ajout final » à notre module de réorientation spatiale et géométrique. Ce module est un mécanisme computationnel inné, informationnellement encapsulé, ayant un domaine spécifique, qui est modulaire et exclusivement

centré sur les propriétés spatiales des pièces, comme la longueur des murs et des angles (Hermer et Spelke, 1996). Comme Ridley le note : « La sélection naturelle tend à opérer non pas en effectuant des changements fondamentaux sur des traits adaptatifs préexistants mais en conservant ces traits et en construisant, par-dessus, de nouveaux processus » (Ridley, 1993). Les changements évolutifs tendent surtout à être implantés tardivement dans le développement d'un organisme, lorsqu'ils sont moins susceptibles de troubler les autres processus vitaux (Gould, 1992 ; Ridley, 1993 ; Clarke, 2004). De tels ajouts finals sont littéralement des parties des éléments innés de l'esprit. Il est extrêmement important de distinguer entre les capacités innées qui facilitent les ajouts finals résultant en l'appel aux propriétés non géométriques, dans le cas de la réorientation spatiale, et les ajouts culturels finals, non biologiques, comme les ordinateurs auxquels Andy Clark fait référence. Mieux vaut une distinction claire entre mon esprit et mon esprit utilisant l'ordinateur dans toute philosophie de l'esprit viable. Si je défends alors une conception étroite des relations culture/cognition, qu'il en soit ainsi. Il est préférable de défendre une conception étroite de l'esprit plutôt qu'une conception complètement fausse. Mais je suis content de laisser Fred Adams et Ken Aizawa défendre le genre de perspective que je défendrais contre Andy Clark, si j'en avais l'envie.

Les résultats de Nisbett et ses collègues suggèrent que les humains adoptent différents styles de raisonnement. Ces résultats minent-ils l'universalité de l'organisation de la cognition humaine ? Je ne suis pas sûr que Nisbett pense que ces résultats démontrent qu'il n'y a pas « d'organisation universelle de la cognition humaine ». Dans la citation rapportée par Faucher, il déclare que « toutes les cultures possèdent les mêmes processus cognitifs de base comme outils » mais « [...] leurs outils de choix pour le même problème sont habituellement différents ». Ce qui revient à dire que la niche écologique joue un rôle dans la manière dont les outils innés sont déployés. Mais qui refusera une telle affirmation ? Certainement pas Cosmides et Tooby. À l'appui, considérez cette citation issue de leur article de 1990 portant exactement sur ce sujet :

Si quelqu'un croit en une nature humaine universelle, comme nous le croyons, il faut observer diverses psychologies manifestes, traits ou comportements entre individus et à travers les cultures, et voir ceux-ci comme le produit d'une psychologie innée complexe, qui est commune et sous-jacente, opérant dans différentes circonstances (voir, par exemple, Daly, Wilson, et Weghorst, 1982). L'association entre l'inné et le manifeste fonctionne selon les principes de l'expression qui sont spécifiés dans les mécanismes de la psychologie innée ou dans des programmes développementaux innés qui informent les caractéristiques psychologiques ; ces expressions peuvent différer entre les individus quand différents inputs environnementaux fonctionnent selon la même procédure pour produire différents outputs manifestes (Cosmides et Tooby, 1987, 1989, 1990).

En bref, il n'est simplement pas évident que Tooby, Cosmides et Nisbett seraient en désaccord sur ces problématiques. Faucher rajoute quelques données provenant de Heinrich et Heinrich (1996), qui prétendent montrer que

les styles de raisonnement dans les contextes d'échanges sociaux varient au sein des groupes et des cultures. Certainement, les données exigent d'être examinées avec précaution, mais le fait est que les données en faveur d'un module de détection des tricheurs dans le contexte des échanges sociaux fait légion. C'est-à-dire que ces résultats sont robustes et ont été reproduits à travers plusieurs cultures et par de nombreux chercheurs comme Cosmides et Tooby, Gigerenzer et Hug, et d'autres. Il ne semble pas raisonnable, pour ma part, de rejeter simplement une quantité si importante de données, et reproduites plusieurs fois, que les psychologues évolutionnistes ont développé sur ces sujets.

Le second problème soulevé par Faucher concerne aussi la reprogrammation culturelle de l'esprit. Il affirme que la conséquence du fait que l'esprit est reprogrammé est la prémisse d'un de mes arguments en faveur de la conclusion, fausse, que la connaissance est un genre naturel. La prémisse affirme que le relativisme épistémique est faux. Voici ma réponse à ce sujet. La prémisse en question était une partie d'un argument destiné à montrer qu'en commençant avec les prémisses que les épistémologues traditionnels endossent habituellement, il est possible d'argumenter que la connaissance est un genre naturel. Une de ces prémisses est l'affirmation que le relativisme épistémique est faux. Mon affirmation était simplement que la plupart des épistémologues analytiques endossent cette prémisse, à tel point qu'ils croient que la connaissance est un concept unique et unifié. En d'autres mots, je n'ai pas besoin d'endosser **personnellement** cette prémisse. Tout ce que j'affirmais est que malgré les prémisses que les épistémologues analytiques trouvent typiquement plausibles, il **n'en reste pas moins** qu'il en découle que la connaissance est un genre naturel. Faucher a certainement raison d'indiquer que je change mes appuis entre le chapitre cinq et le chapitre six. Dans le chapitre cinq, je soutiens qu'il faut, entre autres, clarifier notre concept non mélioratif de justification et de connaissance dans l'optique de faire suivre une réponse à la question méliorative sur la façon dont les agents cognitifs devrait mettre à jour leurs croyances à la lumière de nouvelles données. Ce qui signifie que la poursuite de l'investigation sera améliorée en comprenant d'abord la nature de l'objet qui est poursuivi, c.-à-d., la justification épistémique non méliorative et la connaissance. Mais dans le chapitre six, je soutiens que les épistémologues devraient s'occuper du **phénomène** des connaissances comme genre naturel, et non pas du **concept** de connaissance tel que discuté dans le chapitre cinq. Faucher souligne avec justesse que la littérature empirique sur les intuitions épistémiques que Stich et ses collègues ont découvert indique qu'il n'existe pas d'intuitions partagées par tous les humains. Le résultat est qu'il est erroné de penser qu'il existe un **seul**, un unique **concept** de connaissance partagé par tous les humains.

Ce que je voudrais dire à ce propos est ce qui suit. Mon avis concernant l'épistémologie naturalisée continue d'être que nous devrions étudier la justification non méliorative et la connaissance dans l'optique de poursuivre le projet mélioratif de diviser une théorie qui offre des conseils à l'agent cognitif sur la façon dont il devrait mettre à jour rationnellement ses croyances à la

lumière de nouvelles données. Bien sûr, les résultats issus des sciences cognitives sont aussi essentiels à ce projet. La différence est que, dans le chapitre six, je suggère que la justification non méliorative et la connaissance sont constituées par un ensemble de connaissances qui sont des genres naturels. La notion de **concept** de connaissance avait par conséquent été rejetée. Cependant, la contribution au chapitre cinq et six à propos de la connaissance demeure entièrement externaliste et naturaliste en dépit des différences. J'aurai dit, aujourd'hui, qu'il faut étudier le phénomène de la connaissance comme genre naturel tel qu'on le découvre dans la littérature, par exemple, de l'éthologie cognitive. C'est-à-dire que diverses espèces possèdent des connaissances et que ces connaissances jouent un rôle dans une construction théorique réussie en éthologie cognitive. Hilary Kornblith construit un cas très plausible pour sa perspective dans son *Knowledge and Its Place in Nature* (2002). La différence entre ma perspective et celle de Kornblith est, évidemment, que je conçois la connaissance comme un ensemble de genres naturels alors qu'il pense que la connaissance est un seul genre. Je devrais ajouter que, dans ma perspective, une telle connaissance en tant que genres naturels ne constitue pas un argument en faveur du relativisme épistémique parce que je soutiens que tous les humains entretiennent, par exemple, des connaissances mathématiques et empiriques. Ce sont différents genres de connaissance, mais tous les humains (et certains animaux non humains) affichent de telles connaissances. Pour ces raisons, Faucher a raison de souligner que ma position est considérablement différente de celles des épistémologues classiques. Finalement, il note que, selon les travaux effectués par Boyd et Richerson, la capacité humaine d'occuper de nombreux environnements différents, qui sont souvent hostiles, suggère que ce ne sont pas des modules darwiniens universels que facilitent l'adaptation mais les adaptations culturelles à ces environnements. Donc, par exemple, l'habileté humaine de faire des mathématiques et de la logique d'une façon particulière a dû être « culturellement » sélectionnée en raison des avantages de cette façon de faire sur les autres façons de faire des mathématiques et de la logique. Comme Faucher le suggère : « La représentation en chiffres arabes rend les multiplications plus simples qu'avec la représentation en chiffres romains, permettant à certaines activités reliées au commerce de se développer et de s'étendre. » Mon avis, ici, est que de telles sélections « culturelles » doivent être vues comme des métaphores plutôt que comme des invocations littérales à la notion biologique de sélection naturelle pour les raisons qui suivent. Des résultats particuliers ne peuvent être sélectionnés, à part les mécanismes qui permettent à ces résultats de devenir manifestes. Et, par conséquent, il peut y avoir sélection pour une habileté à engager un raisonnement mathématique ; un mécanisme physique peut être sélectionné. Mais il ne peut y avoir de sélection pour certains résultats de ces mécanismes. Pour cette raison, parler de sélection « culturelle » ne peut être qu'une métaphore. Par analogie, il ne peut y avoir de sélection des croyances concernant la perception visuelle mais seulement une sélection des mécanismes ou modules qui sont des perceptions visuelles.

Il est certainement vrai que l'esprit fonctionne à l'intérieur d'un réseau d'institutions culturelles et de technologies, les biologistes appellent cela une « niche environnementale ». Les psychologues évolutionnistes, comme Cosmides et Tooby, seraient les derniers à nier ce fait ou à nier l'importance de telles niches pour l'évolution. Mais il ne faut pas confondre les niches environnementales dans lesquelles l'évolution de notre génotype s'effectue avec notre génotype. Cette erreur est semblable à confondre le bébé avec l'eau du bain ! Sous cet aspect, l'exhortation de Richerson et Boyd, que Faucher cite à la fin de son commentaire, selon laquelle que nous ne pouvons pas simplement étudier le génotype humain en l'absence de la culture est — bien que vraie — une platitude. Personne, à ma connaissance, n'oserait suggérer autre chose.

MURRAY CLARKE
Département de philosophie
Université Concordia

Bibliographie

- Boyd, R. et P. J. Richerson. *Not By Genes Alone : How Culture Transformed Human Evolution*, Chicago, University of Chicago Press, 2004.
- Clarke, Murray. *Reconstructing Reason and Representation*, Cambridge, Mass., The MIT Press, 2004.
- Conee, E. et R. Feldman. *The Generality Problem for Reliabilism*, *Philosophical Studies*, 89, 1998, 1, 1-29.
- Samuels, R. *Evolutionary Psychology and the Massive Modularity Hypothesis*, *British Journal for the Philosophy of Science*, 49, 1998, 575-602.
- Cosmides, L. et J. Tooby. *Origins of Domain-Specificity : The Evolution of Functional Organization*, in Hirschfield et Gelman, 1994, 85-116.
- Fodor, J. *Concepts : Where Cognitive Science Went Wrong*, Oxford, Clarendon Press, 1998.
- . *The Mind Doesn't Work That Way*, Cambridge, Mass., The MIT Press, 2000.
- Hirschfield, L. et S. Gelman. (dir.) *Mapping the Mind : Domain-Specificity in Cognition and Culture*, Cambridge, Cambridge University Press, 1994.
- Kornblith, H. *Knowledge and Its Place in Nature*, Oxford, Clarendon Press, 2002.
- Margolis, E. et S. Laurence. (dir.) *Concepts : Core Readings*, Cambridge, Mass., The MIT Press, 1999.
- McKay, R. et D. Dennett. *The Evolution of Misbelief*, in the *Adaptation and Representation Web Conference*, modéré par G. Origi, A. Reboul, et A. Wozniak, 2007. [En ligne] [www.interdisciplines.org/adaptation]
- Tooby, J. et L. Cosmides. *On the Universality of Human Nature and the Uniqueness of the Individual : The Role of Genetics and Adaptation*, *Journal of Personality*, 58, 1990, 1, 17-67.

Responsabilité et autonomie

Étude critique de : *Nomy Arpaly, Unprincipled Virtue. An Inquiry into Moral Agency*, New York, Oxford University Press, 2003, 208 p.

François Schroeter

Volume 34, numéro 2, automne 2007

URI : <https://id.erudit.org/iderudit/017428ar>

DOI : <https://doi.org/10.7202/017428ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Société de philosophie du Québec

ISSN

0316-2923 (imprimé)

1492-1391 (numérique)

[Découvrir la revue](#)

Citer cet article

Schroeter, F. (2007). Responsabilité et autonomie / Étude critique de : *Nomy Arpaly, Unprincipled Virtue. An Inquiry into Moral Agency*, New York, Oxford University Press, 2003, 208 p. *Philosophiques*, 34 (2), 403–418.
<https://doi.org/10.7202/017428ar>

Étude critique

Responsabilité et autonomie

Étude critique de: Nomy Arpaly, *Unprincipled Virtue. An Inquiry into Moral Agency*, New York, Oxford University Press, 2003, 208 p.

FRANÇOIS SCHROETER

Université de Melbourne

fschro@unimelb.edu.au

Selon Donald Davidson, il est possible d'expliquer l'agir humain en invoquant un modèle particulièrement simple : nos actions sont causées par nos désirs et nos croyances. Le modèle davidsonien (souvent qualifié de « *causal theory* » ou de « *standard theory of action* ») a exercé — et exerce aujourd'hui encore — une influence dominante dans la philosophie de l'action et la philosophie de l'esprit contemporaines. L'un des avantages les plus remarquables de ce modèle est bien sûr sa parcimonie psychologique. D'après Davidson et ses admirateurs, l'agir humain peut en effet s'expliquer sans que l'on doive invoquer « de mystérieux actes de la volonté, ou des formes de causalité étrangères à la science¹ ».

Durant ces quinze dernières années, le modèle davidsonien a été la cible d'un impressionnant barrage de critiques. Christine Korsgaard et David Velleman, en particulier, deux ardents défenseurs d'une approche néo-kantienne de l'agir humain, ont accusé Davidson d'avoir exclu l'agent humain de sa théorie de l'action. En effet, le simple bon sens semble nous indiquer que nos actions sont causées par nous, les agents, et non par nos désirs. Et c'est précisément cette capacité que l'agent a de causer et de contrôler ses actions — autrement dit, la capacité d'agir de façon autonome — qui différencie les êtres humains des animaux. Aux yeux de défenseurs de l'autonomie tels que Korsgaard et Velleman, le modèle davidsonien représente peut-être de façon satisfaisante le comportement animal, mais il ne dispose pas des outils conceptuels nécessaires à une représentation adéquate de l'autonomie caractéristique de notre agir². La tâche principale de la philosophie de l'action serait maintenant de combler cette lacune et d'enrichir le modèle davidsonien en y intégrant les éléments nécessaires à une approche adéquate de l'autonomie humaine.

Dans un important ouvrage récent³, Nomy Arpaly s'attaque aux défenseurs de l'autonomie, et en particulier à la résurgence néo-kantienne dans la philosophie de l'action contemporaine. Selon Arpaly, en effet, l'idée d'un

1. Cf. Davidson, 1980, 83, et plus généralement les essais 1-5 de ce volume. Le modèle davidsonien a bien sûr d'importantes racines historiques dans la philosophie de l'action d'Aristote, de Hobbes et de Hume. Cf. en particulier le chapitre 6 de la première partie du *Léviathan*.

2. Velleman, 1992, 2000, introduction. Cf. également Korsgaard, 1996, Lecture 9, et Korsgaard, 1997, section 2.

3. Arpaly, 2003. Toutes les références indiquées dans le texte sont des références à ce livre.

agent autonome contrôlant ses actions et choisissant lesquels de ses désirs il va suivre est un mythe. Arpaly fonde sa critique de l'autonomie sur une analyse originale des conditions de la responsabilité morale. Pour les défenseurs contemporains de l'autonomie comme pour Kant, c'est notre capacité d'agir de façon autonome qui fait de nous des êtres moralement responsables et qui nous distingue des créatures incapables d'agir moralement. Arpaly s'inscrit en faux contre cette prémisse centrale de l'éthique kantienne. À ses yeux, un examen minutieux de nos pratiques morales révèle que notre responsabilité morale, cette caractéristique distinctive et essentielle de l'agir humain, peut s'expliquer sans l'autonomie. En conséquence, Arpaly pense qu'il convient maintenant de nous débarrasser de notre obsession pour l'autonomie, et ce, tant dans le domaine de la philosophie de l'action que dans le domaine de la psychologie morale.

Les problèmes abordés par Arpaly sont à mon avis importants et méritent d'être pris au sérieux par les défenseurs de l'autonomie. Pourtant, je crois qu'il convient, en réponse à Arpaly, de repenser l'autonomie et non de rompre le lien que la tradition kantienne a établi entre l'autonomie et la responsabilité morale.

1. Émotions et raisons

La démarche d'Arpaly se fonde sur une critique radicale de l'intellectualisme, à ses yeux excessif, qui corrompt un important segment de la philosophie de l'action contemporaine. Selon la tradition kantienne, en particulier, ce sont nos capacités réflexives et discursives qui sont la marque distinctive des êtres humains et qui doivent jouer le rôle central dans une approche philosophique de l'autonomie. Ces capacités intellectuelles nous permettent de reconnaître les raisons pratiques qui s'appliquent à nous et de former des jugements normatifs sur ce qu'il faut faire. Pour les admirateurs contemporains de Kant, une action est autonome quand elle est causée par une telle reconnaissance de raisons issue de nos capacités intellectuelles. L'autonomie humaine est donc affaire de raison et non de désirs et inclinations⁴. Arpaly pense que la tradition intellectualiste kantienne défend une conception beaucoup trop étroite de la psychologie humaine. À son avis, une approche plus réaliste de l'agir humain doit en particulier accorder une place de choix à nos émotions et à la capacité qu'ont ces dernières de nous fournir un accès épistémologique aux raisons pratiques. Arpaly invoque toute une série d'exemples pour appuyer sa critique.

Huckleberry Finn, le héros de Mark Twain, est l'un des protagonistes principaux du livre d'Arpaly. Huck se lie d'amitié avec un esclave, Jim, qu'il aide à échapper à son maître. Pourtant, la conscience morale de Huck lui dit qu'il agit ainsi de manière immorale : comme tous les membres de sa communauté

4. Korsgaard et Velleman, deux des défenseurs les plus importants et originaux d'une approche néo-kantienne de l'agir et de l'autonomie, acceptent la dichotomie que Kant établit entre raison et inclinations. Même si leurs visions des capacités qui sous-tendent notre accès aux raisons pratiques sont sensiblement différentes, Korsgaard (1996) et Velleman (2000) admettent tous deux que les capacités intellectuelles réflexives jouent un rôle essentiel dans la reconnaissance de raisons pratiques et la formation des jugements normatifs qui guident nos actions autonomes.

il pense qu'aider un esclave à s'échapper, c'est le voler à son maître. Huck agit donc de manière acratique, contre le jugement réfléchi qu'il forme au sujet de ce qu'il doit faire. L'ensemble de la tradition philosophique considère de telles actions comme sérieusement déficientes — Huck agit par « pure inclination » d'après Kant, par « vertu naturelle » d'après Aristote. Arpaly pense que de telles critiques sont injustifiées. En agissant contre sa conscience, Huck agit de façon moralement bonne. De plus, son action est une réponse aux raisons morales qui prévalent dans sa situation (75-9). Lorsque Huck se trouve incapable de rendre l'esclave à son maître, il n'est pas la victime d'une simple impulsion irrationnelle. C'est parce qu'il s'est lié d'amitié avec Jim, parce qu'il a éprouvé de l'empathie à son égard et a reconnu que Jim était un être humain méritant le respect que Huck ne peut suivre les injonctions racistes de sa conscience morale. Les capacités émotionnelles de Huck lui permettent donc de détecter les raisons morales appropriées dans sa situation (l'humanité de Jim et le respect qu'il mérite), alors que son jugement moral explicite reflète simplement les convictions morales erronées de sa communauté. Selon Arpaly, les cas où ce sont nos émotions et non nos jugements réfléchis qui détectent les raisons appropriées ne sont pas rares. La tradition philosophique a trop souvent représenté les émotions comme de simples obstacles à la rationalité. Il est maintenant temps de mettre l'accent sur les cas où ce sont nos jugements réfléchis qui nous empêchent de suivre les raisons détectées par nos émotions.

La manière dont Arpaly discute l'exemple de Huck (et nombre d'autres exemples similaires) suggère que nous disposons, à ses yeux, de deux modes d'accès épistémologique aux raisons. Arpaly ne dément pas, bien entendu, l'importance de nos jugements réfléchis pour la connaissance des raisons pratiques. Nos capacités de réflexion sont particulièrement utiles quand nous sommes confrontés à des situations nouvelles ou spécialement complexes : dans de tels cas, ce n'est qu'après délibération réfléchie que nous pouvons discerner quelles sont les raisons qui peuvent justifier notre agir (63-5). Mais cette voie « intellectuelle » d'accès aux raisons n'est pas la seule dont nous disposons. Nous pouvons également reconnaître les raisons pratiques d'une manière plus « viscérale », parfaitement illustrée par l'exemple de Huck. Selon Arpaly, les animaux, même s'ils sont incapables de délibération réfléchie, disposent comme nous de cette voie d'accès viscérale aux raisons pratiques. Les chiens, par exemple, sont capables de conformer leur comportement à des raisons pratiques (63). Ce qu'Arpaly reproche à la tradition intellectualiste néo-kantienne, c'est de surestimer l'importance de nos capacités réflexives et de rejeter tout accès viscéral aux raisons. Bien souvent nos émotions nous mettent en contact avec des raisons que nous sommes incapables d'exprimer dans des jugements réfléchis.

L'idée d'une voie d'accès viscérale aux raisons n'est pas, bien sûr, sans détracteurs⁵. La nature des raisons est un sujet particulièrement controversé,

5. De même, l'idée parente selon laquelle les émotions constituent des perceptions de valeurs est également controversée (pour une illustration intéressante de cette controverse, cf. l'échange entre Dumouchel, 2002, et Tappolet, 2002).

et Arpaly ne propose aucune théorie métaéthique qui puisse dissiper les doutes de ses adversaires néo-kantiens. Velleman, par exemple, a récemment réaffirmé son opposition radicale à l'idée que de simples désirs ou émotions puissent nous donner accès aux raisons pratiques⁶. À ses yeux, une considération joue le rôle de raison pratique si elle contribue à justifier l'action dont elle est la raison. Avoir accès à une considération dans son rôle de raison, c'est donc avoir accès à la force justificatrice de cette considération. Selon Velleman, nos capacités intellectuelles réflexives (capacités conceptuelles et inférentielles) sont requises pour avoir accès au rôle justificateur des raisons. Désirs et émotions ne peuvent d'eux-mêmes nous donner accès à ce rôle justificateur.

Faut-il conclure que nous avons abouti ici à une impasse, les défenseurs du rôle épistémologique des émotions tels qu'Arpaly et leurs adversaires néo-kantiens assumant des approches de l'épistémologie des raisons radicalement opposées ? Je crois qu'il est fort heureusement possible d'éviter une telle impasse. On peut certainement reprocher à Arpaly de ne pas fournir une défense suffisante de l'accès viscéral que nos émotions nous donnent, selon elle, aux raisons pratiques. Une telle défense nécessiterait de toucher à des questions particulièrement épineuses, comme la nature des raisons et le contenu représentationnel des émotions. Arpaly n'aborde pas ces questions. Pourtant, cette lacune, même si elle est importante, ne semble pas en fin de compte porter à conséquence pour son argumentation principale.

Les raisons qui intéressent Arpaly sont en effet les raisons morales — les raisons appropriées à la question de la responsabilité morale. Et elle insiste sur le fait que de telles raisons ne sont pas suffisamment « simples » pour être reconnues par des créatures dénuées de capacités intellectuelles réflexives : en effet, nos capacités intellectuelles sont à ses yeux nécessaires pour reconnaître les raisons morales. La résistance viscérale de Huck, en particulier, est informée par un riche arrière-plan de croyances morales (par exemple la croyance qu'il faut aider ses amis) qui nécessitent le déploiement d'un répertoire conceptuel sophistiqué (146-8). Peu importe donc si la thèse plus ambitieuse suggérée par Arpaly — l'idée que de simples émotions procurent aux animaux ne disposant pas d'un répertoire conceptuel sophistiqué un accès viscéral non conceptuel aux raisons pratiques — résiste ou non aux doutes d'auteurs néo-kantiens tels que Velleman. La critique qu'Arpaly adresse ici à l'intellectualisme dominant dans la philosophie de l'action néo-kantienne contemporaine repose sur une thèse plus modeste. Ce que l'exemple de Huck démontre à ses yeux, c'est que nos émotions (informées par nos croyances morales implicites) peuvent dévoiler d'importantes raisons morales auxquelles nos pratiques délibératives et les jugements réfléchis qui en résultent demeurent aveugles. Cette thèse plus modeste ne repose sur aucune présupposition théorique sujette à controverse. Je crois qu'elle mérite d'être prise très au sérieux. Quand ils insistent sur le fait que la reconnaissance de raisons pratiques est la clé de l'autonomie, les

6. Velleman, 2000, introduction, et 2004.

auteurs néo-kantiens ont tort d'assumer l'existence d'une dichotomie rigide entre raison et inclination. Nos émotions, même quand elles vont à l'encontre de nos jugements réfléchis, peuvent nous donner accès à d'importantes raisons pratiques morales.

2. Valeur morale et autonomie

La revalorisation du rôle épistémologique des émotions constitue la première prémisse de la discussion qu'Arpaly consacre à la question de l'autonomie et de la responsabilité morale. La seconde prémisse de sa discussion réside dans la théorie originale de la « valeur morale » (*moral worth*) qu'elle propose. Les jugements que nous portons sur la valeur morale de nos actions (méritent-elles le blâme ou l'éloge [*praise*] ?) ont un lien très étroit avec la question de la responsabilité morale. Si nous estimons que la responsabilité d'un agent est diminuée, nous nous refusons en effet à le blâmer (ou, en tous cas, nous modérons notre blâme) pour ses actions immorales. Il est donc capital d'avoir une conception adéquate des jugements que nous portons sur la valeur morale de nos actions si nous voulons aborder la problématique de la responsabilité morale et déterminer le rôle qu'y joue l'autonomie de la volonté.

C'est dans le troisième chapitre de son livre qu'Arpaly présente sa théorie de la valeur morale. Dans son approche de la valeur morale de la bonne volonté, Kant insiste sur l'importance de la pureté des motifs : pour avoir valeur morale, une action doit être accomplie par devoir, et non seulement être en conformité avec le devoir. Arpaly concède à Kant qu'il n'est pas suffisant de conformer ses actions à la loi morale pour leur conférer une valeur morale, mais elle refuse de considérer l'agir par devoir comme une condition nécessaire à la valeur morale. À ses yeux, ce qui est requis pour la valeur morale, c'est que la volonté soit bonne, autrement dit que l'action soit accomplie *pour de bonnes raisons* (72). Ainsi, par exemple, si les actions du marchand de Kant n'ont pas de valeur morale, c'est parce que le marchand se conforme au devoir de ne pas voler ses clients pour de mauvaises raisons — pour des raisons d'intérêt personnel, et non pour des raisons morales. Selon Arpaly, il est absolument crucial d'éviter ici de tomber dans le préjugé intellectualiste qui a corrompu toute la tradition kantienne. Selon Arpaly, le cœur comme l'intellect peuvent assurer qu'une action est accomplie pour de bonnes raisons et qu'elle possède ainsi une valeur morale. Comme nous l'avons vu, l'action de Huck, par exemple, est accomplie pour de bonnes raisons morales : ce qui motive Huck, c'est sa reconnaissance de l'humanité de Jim et du respect que cette humanité mérite (73-9). Pourtant, l'action de Huck est accomplie contre son jugement moral explicite. Il faut donc rejeter la dichotomie kantienne traditionnelle qui oppose l'agir par devoir à l'agir par inclination. Il n'est pas nécessaire de suivre nos représentations explicites du devoir moral pour conférer une valeur morale à nos actions. Nos capacités émotionnelles peuvent, aux yeux d'Arpaly, conférer elles aussi une valeur morale à nos actions.

Pour compléter sa discussion de la valeur morale, Arpaly propose également une analyse de la valeur morale négative et des conditions qui justifient le blâme. Son analyse reflète ici de manière parfaitement symétrique les conditions qu'elle élabore pour la valeur morale positive. D'après Arpaly, une action mérite le blâme si elle a été causée en l'absence d'une bonne volonté ou par une volonté moralement mauvaise. En d'autres termes, une action mérite le blâme si l'agent n'a pas été suffisamment sensible à de bonnes raisons morales ou s'il a été sensible à des raisons moralement abjectes (79-83).

Armée de son analyse de la valeur morale, Arpaly peut se pencher sur la problématique de la responsabilité et de l'autonomie. Elle ne propose pas de théorie complète de la responsabilité morale. Pourtant, dans les deux derniers chapitres de son livre (chap. 4 et 5), elle adresse une critique radicale à toute la tradition kantienne qui a fait de l'autonomie de la volonté la pièce centrale de la responsabilité morale. En prélude à sa critique de l'autonomie, Arpaly dénonce l'ambiguïté qui accompagne de nos jours le terme « autonomie ». Ce terme est utilisé dans la littérature contemporaine pour signifier différentes choses, et une telle ambiguïté, qui n'est, de plus, pas toujours reconnue comme telle, peut bien sûr créer de nombreuses confusions. Je crois qu'Arpaly a raison de nous inciter à la prudence quand nous utilisons un terme associé à autant d'idées différentes⁷.

La critique principale d'Arpaly s'adresse à la composante fondamentale de l'autonomie, ce qu'elle appelle l'autonomie agentielle (*agent-autonomy*) et qu'elle définit de la façon suivante :

L'autonomie agentielle est la relation entre un agent et ses motivations, relation qui peut être caractérisée de façon approximative comme la capacité de l'agent de décider laquelle de ses motivations il va suivre : il s'agit d'un type d'autocontrôle ou d'autogouvernement que possèdent les personnes humaines et non les animaux non humains (118).

L'autonomie agentielle, telle que la définit Arpaly, est bien au cœur de la discussion philosophique contemporaine : c'est ce type d'autonomie que Harry Frankfurt, Christine Korsgaard, David Velleman, Michael Bratman, et bon nombre d'auteurs importants de la philosophie de l'action contemporaine

7. Arpaly distingue, dans la littérature philosophique contemporaine, les huit « significations » suivantes du terme « autonomie » : l'autocontrôle, l'autonomie comme indépendance générale, l'indépendance d'esprit, l'autonomie comme droit individuel, l'authenticité, l'auto-identification, l'autonomie comme idéal personnel, et l'autonomie comme capacité d'agir pour de bonnes raisons (117-30). Elle pense que la présence de ces huit significations différentes rend le terme « autonomie » peu utile dans une discussion philosophique sérieuse. Je crois qu'Arpaly exagère l'ambiguïté du terme « autonomie » et qu'une interprétation plus charitable du discours philosophique portant sur l'autonomie peut construire un modèle plus cohérent des multiples aspects de l'usage philosophique de ce terme. Mais il est possible d'éviter ici ces questions sémantiques. Comme je l'indique dans le texte, la critique fondamentale qu'Arpaly adresse aux défenseurs de l'autonomie est centrée sur une seule des « significations » du terme « autonomie » qu'elle distingue : l'autonomie comme autocontrôle ou l'autonomie agentielle. Tout risque d'ambiguïté peut donc être écarté ici.

considèrent comme une caractéristique essentielle de l'agir humain. Et c'est parce que l'agent humain a la capacité de choisir lequel de ses désirs il va suivre que, de l'avis des défenseurs de l'autonomie, nous pouvons lui attribuer la responsabilité morale de ses actions. Comme les animaux non humains n'ont pas la capacité de choisir lesquels de leurs désirs ils vont satisfaire, nous ne pouvons les tenir pour responsables de leurs comportements.

Selon Arpaly, l'importance attribuée à l'autonomie agentielle dans la tradition néo-kantienne et dans la philosophie de l'action contemporaine en général n'est qu'un autre symptôme des préjugés intellectualistes qui dominent notre tradition philosophique. Il n'est pas nécessaire de postuler un moi autonome et contrôleur pour expliquer ce qui fait de nous des agents responsables (132-9). Pour établir cette conclusion, Arpaly se fonde sur son analyse de la valeur morale (chap. 3). Il est communément reconnu que la responsabilité est une condition centrale de la valeur morale : nous ne considérons pas comme responsables les animaux ou les êtres humains sérieusement retardés, et de ce fait nous n'attribuons pas de valeur morale à leurs comportements. Au vu de l'approche de la valeur morale qu'elle développe dans son livre, il n'est pas difficile de voir pourquoi Arpaly pense que l'autonomie n'est pas nécessaire à l'explication de la responsabilité. Ce qui est crucial pour la valeur morale, c'est la capacité de suivre les raisons (*reason-responsiveness*), la capacité d'agir pour de bonnes raisons. C'est parce que Huck suit les raisons morales (l'humanité de Jim et le respect qu'elle mérite) que nous sommes prêts à attribuer une valeur morale à son action et à le considérer comme responsable. Par contre, la présence d'un moi contrôleur n'est, selon Arpaly, absolument pas requise pour la valeur morale et la responsabilité. L'action de Huck n'est certainement pas produite par un acte d'autocontrôle de la part d'un moi autonome. Huck est simplement envahi par le sentiment de sympathie qu'il éprouve pour Jim, et son jugement moral réfléchi ne réussit pas à guider son agir.

Arpaly, bien sûr, ne dément pas que l'autonomie agentielle soit une marque distinctive de l'être humain. Pourtant, ce n'est pas à son avis l'absence d'autonomie agentielle chez les animaux non humains qui explique pourquoi nous ne considérons pas ces derniers comme moralement responsables. Ce qui est essentiel à la responsabilité morale, c'est la capacité de suivre les raisons morales. Et c'est précisément cette capacité qui fait défaut aux animaux non humains. Même s'ils peuvent percevoir des raisons simples, les animaux non humains n'ont pas les capacités conceptuelles sophistiquées requises pour comprendre les exigences de la moralité et les raisons qui en découlent (146-7). C'est parce qu'ils disposent de ces capacités conceptuelles sophistiquées — et non parce qu'ils peuvent exercer une autonomie agentielle sur certaines de leurs actions — que les êtres humains sont moralement responsables de leurs actions.

3. Repenser l'autonomie

Quoi qu'en dise Arpaly, l'idée d'autonomie agentielle — l'idée d'un moi capable de résister à ses inclinations et de contrôler son agir — semble bien

occuper une place centrale dans notre conception de l'agir humain et de la responsabilité morale. Arpaly est consciente que l'attaque qu'elle dirige contre l'autonomie ne manquera pas d'engendrer un véritable barrage d'objections. Pour consolider sa position, elle se lance dans une discussion détaillée de toute une série d'exemples destinés à démontrer que l'idée d'autonomie ne joue pas dans l'agir humain le rôle fondamental qu'on lui attribue communément. Ironiquement, certains des exemples qu'elle examine me semblent bien plutôt suggérer que l'idée d'autonomie est indispensable à une approche de l'agir et de la responsabilité humaine.

Considérons en particulier une série importante de pathologies — telles que l'hyperactivité, le déficit attentionnel, ou le syndrome de Tourette — sur lesquelles Arpaly porte son attention. Arpaly examine le cas de John, un patient souffrant d'hyperactivité et de déficit attentionnel. Ses actions sont dictées par ses impulsions du moment. Après avoir remarqué qu'une maison est à vendre, il l'achète en quelques minutes, alors qu'il n'a nullement besoin d'un nouveau logement (151-3). Arpaly reconnaît que les défenseurs de l'autonomie de la volonté insisteront ici sur le fait que de tels exemples confirment l'importance de l'autonomie pour les questions de responsabilité. Ce qui fait défaut à John, c'est la capacité de résister à ses impulsions et de contrôler son agir de manière raisonnable — ce qui manque à John, c'est donc l'autonomie agentielle telle que la définit Arpaly. Du fait que cette capacité fondamentale d'autonomie fait défaut à John, nous estimons que sa responsabilité est sérieusement limitée. Faut-il en conclure que l'autonomie est après tout un élément central de l'agir et de la responsabilité humaine ?

La réponse d'Arpaly est malheureusement décevante à ce point. Elle concède que l'autonomie (la capacité de résister à ses impulsions et de contrôler son agir) a un rôle à jouer dans une théorie de la responsabilité. Pourtant, elle pense que l'exemple de John ne démontre pas que l'autonomie joue un rôle *privilégié* (153) dans les questions de responsabilité. À ses yeux, ce qui est central dans les questions de valeur morale et de responsabilité, c'est la qualité de la volonté de l'agent : si l'action indique que la volonté de l'agent est moralement mauvaise, elle mérite le blâme ; si elle indique au contraire que la volonté de l'agent est moralement bonne, elle mérite alors d'être louée. Parce que John souffre d'un déficit d'autonomie, ses actions impulsives ne nous permettent pas, selon Arpaly, de conclure que sa volonté est moralement mauvaise. De même, si John souffrait d'un déficit de la perception, le rendant par exemple incapable de reconnaître son épouse, il serait faux de conclure que sa volonté est moralement mauvaise quand, faute d'avoir reconnu son épouse, il ne lui procure pas l'aide dont elle a besoin. Arpaly conclut que, de même que le bon fonctionnement de nos organes de perception, l'autonomie de la volonté a peut-être un rôle indirect à jouer dans les questions de responsabilité : les déficits perceptuels, comme les déficits dans l'autonomie de la volonté, peuvent falsifier le lien que nous établissons d'ordinaire entre le comportement d'un agent et la qualité morale de sa volonté. Pourtant, et c'est là le point important à

ses yeux, il convient de ne pas perdre de vue le fait que c'est la question de la qualité morale de la volonté qui est décisive pour les questions de valeur morale et de responsabilité (152-3).

Une telle réponse a peu de chances de convaincre les défenseurs de l'autonomie de la volonté. Nous estimons que John ne peut pas être considéré comme totalement responsable de son action. Nous estimons en effet que John ne remplit pas certaines des préconditions fondamentales qui doivent être satisfaites pour qu'une assignation de valeur morale puisse être appropriée dans son cas. Je suppose que personne ne contestera que la question de la responsabilité touche à la satisfaction ou à la non-satisfaction de telles préconditions fondamentales. Et l'exemple de John illustre parfaitement le rôle central que l'autonomie de la volonté joue quand il s'agit de déterminer le contenu de ces préconditions fondamentales. La tradition philosophique distingue en effet deux grandes classes de facteurs disculpants dans les questions touchant à la responsabilité : facteurs cognitifs d'une part, et facteurs volitifs d'autre part. Ces facteurs ont une très grande plausibilité intuitive et ils exercent une influence déterminante sur nos attributions de responsabilité. En ce qui concerne les facteurs cognitifs, nous estimons en général que la responsabilité d'un agent pour une action particulière est limitée s'il ignore certains éléments décisifs pour l'évaluation de cette action. De façon similaire, dans le domaine volitif, si une déficience importante affecte les capacités de décision de l'agent, nous concluons d'ordinaire que sa responsabilité est limitée. La tradition philosophique assume que l'idée d'un moi capable de résister à ses impulsions et de contrôler ses actions — l'idée d'autonomie — joue un rôle central quand il s'agit de déterminer si les capacités de décision de l'agent fonctionnent de façon adéquate ou non. Les exemples de John et d'autres pathologies similaires confirment le bien-fondé de cette assumption. Comme le reconnaît Arpaly, phobies, compulsions, déficits attentionnels, toutes ces conditions empêchent le moi d'exercer un contrôle normal sur ses désirs et ses actions. En conséquence, nous estimons que la responsabilité des agents souffrant de telles conditions est limitée. C'est dire que l'autonomie de la volonté fait bien partie des préconditions fondamentales dont la satisfaction ou la non-satisfaction détermine si un agent est responsable de ses actions.

Il est bien clair que l'idée d'un moi contrôlant ses actions n'est pas sans problèmes. Personne aujourd'hui ne souhaite faire appel à un moi kantien nouménal, séparé du moi empirique, quand il s'agit d'expliquer l'autonomie de la volonté. Mais il n'est pas aisé de voir comment un moi autonome peut émerger des structures neurologiques et déterministes qui constituent l'esprit humain. Peut-on expliquer le contrôle que nous semblons posséder sur nos actions sans poser la présence d'un « *homunculus* » aux commandes de notre agir ? La grande majorité des philosophes contemporains sont convaincus qu'une telle explication doit être possible, mais nous sommes encore bien loin, malgré les progrès spectaculaires des neurosciences, d'avoir éclairé de façon satisfaisante les mécanismes qui rendent possibles le contrôle que nous exerçons

dans les conditions normales sur nos actions. Même si l'idée d'un moi contrôlant son agir reste donc entourée de mystère, il convient cependant d'insister sur le fait qu'Arpaly ne nous fournit pas de raison de douter de l'existence d'un tel moi et de son importance pour la question de la responsabilité. Arpaly a raison de penser que, lorsqu'il s'agit de déterminer la valeur morale d'une action, la question de la qualité de la volonté de l'agent joue un rôle déterminant. Nous pensons en effet qu'un agent ne mérite d'être blâmé pour une action que s'il a été, dans la formation de ses intentions, insuffisamment sensible aux raisons morales. Mais c'est là une thèse que les défenseurs de l'autonomie peuvent très bien accepter. La présence dans l'agent d'une capacité de résister à ses impulsions et de contrôler son agir n'en demeure pas moins l'une des préconditions essentielles qui doivent être satisfaites pour qu'une assignation de valeur morale soit appropriée. Vouloir nier ce fait, c'est faire violence au sens commun et à nos pratiques morales.

Faut-il conclure que les condamnations sévères qu'Arpaly adresse à l'idée d'autonomie sont injustifiées ? Je crois qu'il convient de tirer de sa contribution une leçon plus positive. Les théories de l'autonomie ont été dominées par le courant intellectualiste kantien, et Arpaly a raison de s'inscrire en faux contre cette domination. Elle semble avoir mal choisi sa cible : ce n'est pas l'idée d'autonomie comme telle qu'il faut rejeter, mais les interprétations intellectualistes que la tradition kantienne a fournies de cette dernière.

En effet, les défenseurs principaux de l'autonomie auxquels Arpaly s'oppose — Harry Frankfurt, Christine Korsgaard, David Velleman, Michael Bratman — acceptent tous le même modèle de base pour expliquer l'autonomie humaine. D'après ce modèle, kantien dans son inspiration, une action est autonome si elle est causée par les jugements évaluatifs réfléchis que le moi porte sur ses désirs et ses actions. Si l'action est déterminée simplement par les désirs immédiats de l'agent, elle n'est pas autonome. Par exemple, je n'agis pas de façon autonome quand je cède à mon désir immédiat de vengeance et agresse mon adversaire. Pour que mon action soit autonome, il faut qu'elle soit causée par un jugement (ou un désir) émanant de mes facultés intellectuelles réflexives, facultés qui me permettent de prendre une distance critique face à mes désirs immédiats. Même si Frankfurt, Korsgaard, Velleman et Bratman diffèrent dans les détails de leurs conceptions de l'autonomie, ils acceptent tous ce même modèle de base intellectualiste⁸.

8. Les travaux de Frankfurt (1988) ont exercé une influence considérable sur la discussion contemporaine de l'autonomie. Frankfurt établit un contraste entre les créatures dont le comportement est dicté par de simples désirs, et les personnes humaines. Ces dernières ont la capacité de réfléchir sur leurs désirs (et de former ainsi des désirs de second ordre au sujet de ces simples désirs). C'est parce que nos capacités réflexives peuvent déterminer nos actions que notre agir est libre et autonome. Pour les modifications plus récentes de la théorie frankfurtienne de l'action (cf. Frankfurt, 1999). Korsgaard, 1996, Velleman, 2000, et Bratman, 2000, reconnaissent tous leur dette envers le modèle de Frankfurt dans leur approche de l'autonomie. Il convient de noter, cependant, que Korsgaard, Velleman et Bratman diffèrent de Frankfurt en ce qu'ils caractérisent en termes de raisons les évaluations réflexives qui jouent le rôle central dans leur approche de l'autonomie. Frankfurt, par contre, ne considère pas ces évaluations réflexives comme des jugements portant sur les raisons pratiques.

C'est quand il s'agit d'illustrer les faiblesses de ce modèle intellectualiste que la critique d'Arpaly est particulièrement utile. En effet, le modèle intellectualiste semble proposer une approche beaucoup trop restrictive de l'autonomie humaine : il n'est pas nécessaire d'agir par devoir, ou d'une manière plus générale de suivre les jugements évaluatifs de son intellect pour agir de façon autonome et être responsable de ses actions.

Considérons un exemple banal. Bien souvent, comme Huck, nous suivons nos inclinations et agissons à l'encontre de nos jugements évaluatifs réfléchis. Je sais que je dois impérativement corriger les copies de mes étudiants, et pourtant je passe ma soirée cloué devant mon téléviseur. Nous estimons que, dans un tel cas, je suis responsable de mon action et mérite d'être blâmé. Cet exemple est donc radicalement différent du cas de John qui, du fait de la profonde pathologie qui l'afflige, ne peut offrir aucune résistance aux impulsions qui s'emparent de lui et ne peut en conséquence être tenu pour responsable de ses actions. Ce qui fait défaut à John, c'est la capacité du moi de résister aux impulsions qui l'assaillent et de contrôler ainsi ses actions — capacité qu'Arpaly nomme autonomie agentielle. Par contre, cette capacité est présente en moi quand je regarde la télévision, et c'est pour cette raison que, contrairement à John, je suis responsable de mes actions.

D'après le modèle intellectualiste de l'autonomie proposé par Kant, Frankfurt et leurs admirateurs, je n'agis pas de façon autonome quand je reste cloué devant mon téléviseur — mon action n'est en effet pas causée par mon jugement évaluatif réfléchi. Ce que les analyses d'Arpaly suggèrent, c'est que l'autonomie, si elle est conçue de manière intellectualiste, ne joue pas un rôle fondamental dans les questions de responsabilité. Le simple fait qu'une action soit acratique suffit, selon le modèle intellectualiste, à établir qu'elle n'est pas autonome. Toutefois, nous sommes de toute évidence moralement responsables pour beaucoup de nos actions acratiques.

Les défenseurs du modèle intellectualiste de l'autonomie — ou en tous cas certains d'entre eux — seront vraisemblablement prêts à accepter la dissociation de l'autonomie et de la responsabilité sur laquelle insiste Arpaly. Ce que le modèle intellectualiste réussit peut-être à articuler, c'est un certain idéal d'autonomie : l'idéal d'une action totalement sous le contrôle de l'agent, ou d'une action à laquelle l'agent « participe » de façon maximale⁹. Une théorie de l'autonomie qui vise un tel idéal ne peut de toute évidence pas prétendre éclairer les conditions nécessaires de la responsabilité morale.

Je ne souhaite pas ici me pencher sur l'avenir du modèle intellectualiste de l'autonomie et déterminer si ce modèle peut vraiment jouer le rôle d'un idéal d'autonomie. Ce qui me paraît par contre capital, c'est de reconnaître la nécessité de distinguer différents *degrés* d'autonomie. Mon action acratique — lorsque je reste cloué devant le téléviseur — est bien éloignée de l'idéal d'autonomie qui intéresse les défenseurs du modèle intellectualiste. Personne ne

9. C'est là par exemple l'opinion de Michael Bratman (communication personnelle).

considérerait une telle action comme un modèle parfait d'autonomie. Pourtant, contrairement à John, je dispose d'un important degré de contrôle (ou d'autonomie agentielle) sur mon action. Le modèle intellectualiste, qui a dominé les discussions contemporaines de l'autonomie, ne procure pas une approche plausible du type de contrôle que j'ai sur de telles actions et qui fonde ma responsabilité dans de tels cas. Ce que la critique d'Arpaly démontre, me semble-t-il, c'est que nous avons besoin d'un nouveau modèle de l'autonomie, plus flexible que le modèle intellectualiste, qui soit capable de différencier les différents degrés d'autonomie qui nous intéressent¹⁰. Comme nous l'avons vu, l'idée d'un moi capable de résister à ses impulsions et de contrôler ses actions joue un rôle important dans les attributions de responsabilité. Cela signifie que l'idée d'un lien essentiel entre l'autonomie et la responsabilité morale doit être réaffirmée. Ce qui est maintenant urgent, c'est de développer une théorie de l'autonomie qui puisse éclairer ce lien essentiel et isoler le type de contrôle qui accompagne les actions dont nous sommes responsables.

4. Devoir moral et autonomie

Pour consolider l'idée d'un tel lien entre l'autonomie et la responsabilité, j'aimerais maintenant me pencher sur certaines difficultés plus fondamentales auxquelles se heurte la tentative d'isoler la responsabilité de l'autonomie. Une telle tentative d'isolement ne fait pas seulement violence aux intuitions que nous avons au sujet de la responsabilité. Elle se heurte également à d'importants problèmes métaéthiques.

Il n'est pas difficile de reconnaître dans l'approche d'Arpaly l'influence du modèle davidsonien, déjà cher à Hobbes et Hume. Selon ce modèle, différents désirs, moraux et immoraux, entrent en conflit, et l'action est simplement causée par le désir le plus fort. C'est bien cette vision fondamentalement atomistique de l'esprit humain, où différentes attitudes entrent en interaction et causent une action sans le contrôle d'un moi autonome, qui est au fondement de la psychologie morale d'Arpaly et de son approche de la valeur morale. Certains individus ont un désir moral¹¹ fort qui, dans les situations appropriées, suffit à causer les actions requises par la moralité. Les actions de tels individus, quand elles sont causées par ce désir moral, sont accomplies pour de bonnes raisons morales et méritent donc l'éloge. D'autres agents ont par contre un désir moral plus faible. Lorsque, dans les situations appropriées, ce désir est dominé par d'autres désirs et qu'une action immorale est ainsi engendrée, elle mérite le blâme.

Il faut reconnaître à Arpaly le mérite d'accepter les conséquences qu'une telle psychologie morale — où l'action est causée par les désirs de l'agent, sans l'intervention ou le contrôle d'un moi autonome — implique pour les questions de la valeur morale et du blâme. « D'après mon approche, nous dit Arpaly,

10. Pour une esquisse d'un tel modèle, cf. Schroeter 2004

11. Arpaly parle de « *moral concern* » qu'elle définit comme un type de désir (84-93).

le blâme est analogue à traiter quelqu'un de mauvais homme d'affaires ou d'artiste médiocre » (173). En effet, du fait de leurs dispositions naturelles et de leurs conditions sociales, certains individus héritent d'un désir moral fort alors que d'autres ont un désir moral plus faible. Nous attribuons éloges aux premiers et blâmes aux seconds. Certes, ce sont ultimement la chance (*moral luck*) et des facteurs sur lesquels nous n'avons aucun contrôle qui déterminent la force du désir moral d'un individu particulier, et donc si les actions de ce dernier auront une valeur morale positive ou négative. Arpaly pense que nous devons nous faire à cette conclusion. La chance joue également un rôle important dans nos évaluations non morales. C'est en effet la loterie génétique qui détermine dans une large mesure les talents économiques et artistiques d'un individu. Pourquoi n'en irait-il pas de même quand il s'agit de la valeur morale de nos actions ?

Les défenseurs de l'autonomie insisteront sur le fait que la position d'Arpaly est ici en sérieux porte-à-faux avec le sens commun. « *Ought implies can* », dit le slogan. Si une raison morale ou un devoir moral s'applique à un agent, il doit être possible à celui-ci d'y conformer son comportement¹². Cette condition fondamentale de la normativité est au cœur de nos pratiques morales et a une très forte plausibilité intuitive. Elle explique pourquoi les évaluations morales sont à nos yeux différentes des évaluations économiques ou artistiques. Ce que le sens commun suggère, c'est que chaque individu normal a la capacité de reconnaître les exigences de la moralité et d'y conformer son comportement. Par contre, bien peu d'êtres humains normaux ont les talents nécessaires pour être de bons artistes ou de bons hommes d'affaires. Ce fait contribue à expliquer pourquoi, contrairement au domaine moral, nous ne considérons pas que tout être humain normal soit soumis à des impératifs normatifs en matière de succès économique et artistique. Nous avons tous le devoir d'agir de façon morale, mais pas le devoir d'être de bons artistes et bons hommes d'affaires.

La présence d'un moi autonome, capable de résister à ses différents désirs et inclinations donne tout son sens au slogan « Devoir implique pouvoir ». C'est cette présence par exemple qui explique pourquoi nous estimons que j'aurais dû ne pas rester cloué devant mon téléviseur et qu'il était en mon pouvoir de corriger les copies de mes étudiants au lieu de perdre mon temps à regarder un programme stupide. Si nous acceptons par contre le modèle davidsonien, d'après

12. Cette condition est appelée « internalisme de l'existence » ; elle est en général contrastée avec l'internalisme du jugement (cf. Darwall, 1983, 54). D'après l'internalisme de l'existence, nécessairement, si une raison s'applique à un agent, cet agent doit être capable d'être motivé à agir en conformité avec cette raison. D'après l'internalisme du jugement, nécessairement, si une personne forme des jugements au sujet des raisons qui s'appliquent à elle, elle doit être motivée à agir en conformité avec ces jugements. Si l'internalisme du jugement est une thèse très controversée, l'internalisme de l'existence — le type d'internalisme dont il est question dans le texte — est très largement accepté, et ce, tant par les néo-humiens (Williams, 1981, chap. 8), que par les néo-kantiens (Korsgaard, 1986) et les néo-aristotéliens (McDowell, 1998, chap. 5).

lequel nos actions sont causées par nos désirs les plus forts sans le contrôle d'un moi autonome, il est beaucoup plus difficile de voir comment, ainsi que l'admet le sens commun, j'aurais pu agir de façon différente dans de telles circonstances.

Proposer une analyse détaillée de la capacité d'agir autrement que semble postuler le sens commun est bien sûr une tâche particulièrement délicate : nous sommes bien loin de disposer d'une approche théorique satisfaisante des mécanismes psychologiques qui sous-tendent cette capacité ou d'avoir clarifié le contexte modal dans lequel s'inscrit cette référence à une telle capacité¹³. Pourtant, il convient de bien mesurer les problèmes métaéthiques auxquels se heurte quiconque prétend, comme Arpaly, expliquer la responsabilité et les devoirs moraux sans l'autonomie. Le problème n'est pas simplement que nos assignations de blâme sont très sensibles à tout ce qui touche à notre capacité d'agir autrement : toute limitation de cette capacité, par exemple dans le cas des pathologies d'addiction, entraîne immédiatement une diminution du blâme et du degré de responsabilité de l'agent. Le problème est que sans l'autonomie il n'est pas facile de voir comment, si devoir implique pouvoir, raisons et devoirs moraux peuvent s'appliquer à tous les êtres humains normaux.

Une comparaison rapide entre la position d'Arpaly et la critique que Bernard Williams adresse à nos pratiques morales peut, je crois, être ici éclairante¹⁴. Comme Arpaly, Williams demeure sceptique face à l'idée d'un moi kantien autonome contrôlant notre agir. Hume et Davidson ont selon lui raison d'insister sur le fait que nos actions sont simplement causées par nos désirs. Mais, contrairement à Arpaly, Williams réalise pleinement combien le modèle humien/davidsonien s'inscrit en opposition à la moralité du sens commun et à la façon dont nous envisageons le blâme. Quand nous blâmons un agent pour son action immorale, nous prétendons, selon Williams, que cet agent avait des raisons d'agir autrement et qu'il était en son pouvoir d'agir autrement. Williams estime que l'éthique kantienne décrit de façon adéquate ce système de présuppositions qui est au cœur de nos pratiques morales. Bien sûr, il rejette l'ensemble de ces présuppositions kantienne, présuppositions qui sont à son avis incompatibles avec la réalité psychologique humaine — alors que Kant décrit correctement les présuppositions de nos pratiques morales, Hume décrit correctement notre réalité psychologique. Mais Williams reconnaît les conséquences très sérieusement révisionnistes de toute tentative de se débarrasser d'un moi kantien aux commandes de son action. Si Hume et Davidson sont dans le vrai, il faut simplement abandonner l'idée que nous sommes tous soumis aux impératifs de la moralité et il faut également renoncer à notre pratique du blâme. En effet, le principe « devoir implique pouvoir » a pour conséquence que les raisons morales ne s'appliquent pas aux individus dépourvus de tout désir moral. Et notre pratique du blâme est à tel

13. Et ce, même si la littérature consacrée à la question du « *could have done otherwise* » est particulièrement volumineuse.

14. Williams, 1985, chap. 10.

point corrompue par le mythe qu'il est toujours en notre pouvoir d'agir de façon morale qu'il convient purement et simplement de s'en défaire.

Tout au long de ses analyses Arpaly assume que les impératifs moraux s'appliquent à tous les êtres humains. Cela signifie-t-il qu'elle entend répudier un principe aussi fortement enraciné dans nos pratiques morales que le principe « devoir implique pouvoir » ? Ou pense-t-elle alors que le modèle humien/davidsonien a les ressources nécessaires pour concilier ce principe avec l'idée que les exigences morales s'appliquent à nous tous, du simple fait que nous avons les capacités intellectuelles requises pour comprendre les impératifs moraux ? Malheureusement, Arpaly n'aborde jamais ces questions métaéthiques pourtant fondamentales.

En conclusion, et contrairement à ce que suggèrent les défenseurs du modèle humien/davidsonien tels qu'Arpaly et Williams, je crois qu'il serait tout à fait prématuré d'abandonner l'idée kantienne d'un moi autonome aux commandes de nos actions. Arpaly ne décrit pas de façon adéquate nos pratiques morales et fait preuve de naïveté métaéthique quand elle prétend que la moralité du sens commun et la manière dont nous envisageons le blâme ne présupposent pas l'existence d'un tel moi autonome. D'une manière plus générale, il convient d'insister, contre les défenseurs du modèle humien/davidsonien, sur le fait que ni une analyse philosophique du sens commun ni les découvertes empiriques récentes ne semblent mettre en doute l'existence d'un moi autonome capable de contrôler ses actions. Mais il faut reconnaître avec Arpaly que les défenseurs de l'autonomie sont, à ce jour, loin d'avoir proposé une caractérisation théorique satisfaisante de l'autonomie et de ses différents degrés.

Bibliographie

- Bratman, M. « Reflection, Planning, and Temporally Extended Agency », *The Philosophical Review*, vol. 109, 2000, pp. 35-61.
- Darwall, S. *Impartial Reason*, Ithaca, Cornell University Press, 1983.
- Davidson, D. *Essays on Actions and Events*, Oxford, Clarendon Press, 1980.
- Dumouchel, P. « Émotion et perception », *Philosophiques*, vol. 29, 2002, pp. 371-77.
- Frankfurt, H. *The Importance of What We Care About*, New York, Cambridge University Press, 1988.
- . *Necessity, Volition, and Love*, New York, Cambridge University Press, 1999.
- Korsgaard, C. « Skepticism about Practical Reason », *Journal of Philosophy*, vol. 83, 1986, pp. 5-25.
- . *The Sources of Normativity*, New York, Cambridge University Press, 1996.
- . « The Normativity of Instrumental Reason » dans G. Cullity et B. Gaut, eds., *Ethics and Practical Reason*, Oxford, Clarendon Press, 1997, pp. 215-54.
- McDowell, J. *Mind, Value, and Reality*, Cambridge, Harvard University Press, 1988.
- Schroeter, F. « Endorsement and Autonomous Agency », *Philosophy and Phenomenological Research*, vol. 69, pp. 633-59.
- Tappolet, C. « Les émotions et leurs conditions d'adéquation. Réponse à Paul Dumouchel », *Philosophiques*, vol. 29, 2002, 378-82.
- Velleman, J. D. « What Happens When Someone Acts ? », *Mind*, vol. 101, 1992, pp. 461-81.

- . *The Possibility of Practical Reason*, New York, Oxford University Press, 2000.
- . « Replies to Discussion on The Possibility of Practical Reason », *Philosophical Studies*, vol. 121, 2004, pp. 277-98.
- Williams, B.A.O. *Moral Luck*, Cambridge, Cambridge University Press, 1981.
- . *Ethics and the Limits of Philosophy*, Cambridge, Harvard University Press, 1985.

Rescher, Nicholas. *Essais sur les fondements de l'ontologie du procès*, traduction et introduction par Michel Weber, Ontos-Verlag, 2006

Frédéric Tremblay

Volume 34, numéro 2, automne 2007

URI : <https://id.erudit.org/iderudit/017429ar>

DOI : <https://doi.org/10.7202/017429ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Société de philosophie du Québec

ISSN

0316-2923 (imprimé)

1492-1391 (numérique)

[Découvrir la revue](#)

Citer ce compte rendu

Tremblay, F. (2007). Compte rendu de [Rescher, Nicholas. *Essais sur les fondements de l'ontologie du procès*, traduction et introduction par Michel Weber, Ontos-Verlag, 2006]. *Philosophiques*, 34 (2), 419–421.
<https://doi.org/10.7202/017429ar>

Comptes rendus

Rescher, Nicholas. *Essais sur les fondements de l'ontologie du procès*, traduction et introduction par Michel Weber, Ontos-Verlag, 2006

Ce livre est la traduction d'une collection de textes tirés de quatre monographies de Rescher : *Process Metaphysics*, *Process Philosophy*, *Nature and Understanding*, et *Inquiry Dynamics*, regroupés sous trois thèmes, soit *Prolégomènes*, *Métaphysiques du procès*, et *Épistémologie du procès*. L'ensemble est précédé d'une brève préface du traducteur qui met en parallèle la philosophie du processus de Rescher avec celle de Whitehead. Dans les *Prolégomènes*, Rescher présente un relevé de l'histoire de la philosophie du processus avant d'en exposer les thèses générales. Il défend ces thèses contre les philosophies anti-processuelles, en particulier celle de P. F. Strawson, et indique quelques difficultés des ontologies substantialistes. Cette première partie forme donc la base des deux autres parties dans lesquelles Rescher applique l'interprétation processuelle à certaines sphères de la science, en commençant avec l'anthropologie philosophique et en proposant une analyse de la personne en termes (partiellement) processuels : une personne n'est pas une substance mais un agrégat de processus, dispositions, tendances, etc. Il l'applique ensuite aux sciences naturelles en s'attardant particulièrement au caractère kaléidoscopique de la complexité de la nature. Rescher s'intéresse en troisième lieu aux processus dans le contexte de la cognition, du développement scientifique et philosophique, et du réalisme métaphysique.

L'application de la perspective processuelle aux différentes sphères de la science n'a, en revanche, une valeur proprement scientifique que dans la mesure où l'ontologie processuelle est elle-même fondée. Attardons-nous donc maintenant sur la question des fondements de l'ontologie du processus qu'annonce le titre. Quels sont ces fondements ? Dans la première partie, une section est consacrée aux « Idées fondamentales » dans laquelle Rescher présente les deux thèses générales de la philosophie du processus : 1) « les choses nécessitent les procès » ; et 2) « les procès sont plus fondamentaux que les choses » (p. 44). La première thèse est causale et veut que les processus soient causalement premiers aux choses, la seconde concerne la constitution ontique des choses et veut que ces dernières soient ultimement constituées de processus. Les deux thèses reposent implicitement sur une troisième suivant laquelle il existe des processus sans propriétaires ou, pour suivre la traduction de Weber, « procès impropres ».

Un processus avec propriétaire (*owned process*) est le processus d'une substance ou, du moins, un processus dont l'existence dépend d'une substance. Un processus sans propriétaire (*unowned process*), au contraire, est un processus sans support substantiel, *free floating* pour employer l'expression de Rescher. Ainsi, le processus sans propriétaire est séparable de toute substance et peut exister indépendamment d'elle. Rescher exprime cette idée de la façon suivante : « Du point de vue du philosophe du procès, l'existence de procès impropres est particulièrement importante, car elle montre que le domaine du procès est complémentaire et séparable du domaine des choses substantielles » (p. 56).

La thèse de l'existence de processus impropres (ou séparabilité du processus) est une condition nécessaire aux deux thèses exposées plus haut. En effet, la thèse suivant laquelle les processus sont causalement premiers aux substances implique qu'il y ait un

processus séparé d'une substance qui soit causalement première à la substance. Et, ensuite, la thèse suivant laquelle les processus sont plus fondamentaux que les substances pré-suppose que les substances sont, au niveau microscopique, ultimement constituées de processus sans propriétaires (je présume que Rescher les conçoit comme étant impropres, car autrement il n'y aurait pas lieu de parler d'une relation de fondation).

Mais cette thèse ne tient pas compte de la corrélation à chaque niveau de granularité entre les mouvements et leurs substances correspondantes : les galaxies sont constituées non pas de rotations impropres mais de systèmes stellaires en rotation, eux-mêmes constitués d'amas de corps célestes en rotation comme notre système solaire. Plus bas dans le spectre des niveaux de granularité, les molécules sont constituées d'atomes, lesquels forment une structure relativement stable moyennant des processus d'attraction et de répulsion, les atomes sont à leurs tours constitués d'électrons en interaction avec le nucleus, dont les parties (protons et neutrons) sont mues par des processus d'attraction réciproque. Ces dernières sont constituées de trois quarks en interaction, etc.

Ainsi, la dichotomie et la correspondance entre substance et processus semblent continues et nécessaires aussi loin qu'il nous est permis de descendre ou de monter dans l'échelle des niveaux de granularité, de sorte que nous n'observons aucun processus réel qui n'est pas corrélé à sa substance, et vice-versa, aucune substance qui n'est pas en pair avec son processus corrélatif propre. En somme, il apparaît invraisemblable de prime abord qu'il existe des processus impropres ou séparés. Et sans une justification de cette thèse invraisemblable, le projet de la philosophie du processus demeure dénué de fondement.

Or, Rescher ne démontre nulle part la validité de cette thèse. Ou peut-être croit-il l'avoir confirmée de façon inductive par les exemples de processus impropres qu'il propose. Ces exemples sont « la baisse de la température, le changement du climat, le flash de l'éclair, la fluctuation d'un champ magnétique » (p. 55). Si ces processus existaient réellement à l'état séparé (c.-à-d., sans propriétaire), alors Rescher pourrait s'appuyer sur des évidences empiriques. Aucun des exemples, cependant, n'est convaincant ; chacun d'eux nous semble réductible à un processus propre (c.-à-d. avec propriétaire). Examinons les exemples un à un.

Réchauffements et refroidissements se produisent toujours dans une portion déterminée d'air ou d'eau (ou autre matière) spatialement localisable, et cette portion déterminée est précisément la substance correspondant au processus. La température est également définie, à un niveau plus raffiné de granularité, comme l'énergie cinétique de particules dans un certain système. Suivant cette définition, le processus de baisse de la température est un processus de diminution de l'énergie cinétique des particules, et le réchauffement une augmentation de celle-ci. L'exemple du changement de climat reçoit la même explication à une plus grande échelle.

Prenons maintenant l'exemple du flash de l'éclair. La traduction de *lightning* par éclair ne rend pas l'effet recherché, car *lightning* est un terme de masse, alors qu'éclair est un terme singulier. Rescher semble inférer qu'aux termes de masses correspondent une masse indéterminée plutôt qu'une substance et qu'en conséquence le processus du flash de l'éclair n'a pas de propriétaire. Mais si certains processus correspondaient à de telles masses, ils ne pourraient pas être dits impropres, car ils auraient des masses comme propriétaires. Il faudrait déjà pour cela que Rescher ait défendu l'existence de masses indéterminées, ce qu'il n'a pas fait (du moins pas dans cette collection de textes).

L'exemple de la fluctuation d'un champ magnétique est plus difficile à discuter, car il pré-suppose une certaine compréhension des équations de Maxwell ou de la

mécanique quantique, mais il est certain qu'il n'y a pas de champ magnétique sans dipôle magnétique, et pas de dipôle magnétique sans un support quelconque (aimant, courant électrique, corps céleste, etc.), de même qu'il n'y a pas de champ magnétique actuel sans électrons qui se meuvent d'un pôle à l'autre. En somme, on n'observe point de champ magnétique sans support.

Plus loin dans le même chapitre, Rescher présente deux nouveaux exemples : la chute de la demande économique et l'érosion de la ligne côtière. Or, de la même façon, la chute de la demande économique est une statistique induite à partir d'un nombre déterminé d'individus, et l'érosion de la ligne côtière est toujours constituée d'un nombre déterminé de processus spécifiques de dissociation, ayant chacun pour support un nombre déterminé de grains de sable (ou autre), spatiotemporellement localisables.

Ainsi, parce que les thèses générales de la philosophie du processus reposent sur la thèse de la séparabilité de certains processus et que Rescher n'a pas démontré que ceux-ci existent effectivement de façon séparée, alors la philosophie (rescherienne) du processus ne repose sur aucun fondement satisfaisant, du moins dans sa version ontologique, dite « forte » (et dans la mesure où la version épistémologique, dite « faible », reposerait sur les mêmes fondements, alors elle serait tout aussi problématique). En effet, sans la thèse de l'existence de processus impropres, celles de la primauté causale et du caractère fondamental des processus se dissolvent, car pour être premier ou fondamental par rapport à autre chose il faut d'abord que l'un puisse se trouver sans l'autre.

La traduction est généralement conforme. Toutefois la reddition de *owned* et *unowned* par « propre » et « impropre » est malheureusement cryptique. De plus, le choix surprenant de « procès » plutôt que « processus » n'est nulle part sérieusement motivé. Weber se contente d'une note dans le lexique où il écrit que *process* est « parfois traduit, pour des raisons euphoniques, [par] "processus" » (p. 276). Or « processus » est préférable pour des raisons plus sérieuses que la simple euphonie. Parmi ces raisons, notons d'abord que le premier sens de « procès » est juridique, de sorte que l'expression « ontologie du procès » pourrait être méprise pour une espèce de l'ontologie du droit et, ensuite, que l'usage déjà établi du terme processus dans le vocabulaire francophone de la philosophie n'a rien de problématique et qu'il est donc superficiel de lui chercher un substitut.

FRÉDÉRIC TREMBLAY

SUNY at Buffalo et IFOMIS (Saarbrücken)

Zeimbekis, John. *Qu'est-ce qu'un jugement esthétique ?* Vrin, coll. Chemins philosophiques, 2006, 128 p.

Cette nouvelle parution relève le défi de présenter une notion incontournable qui a alimenté les querelles et débats en esthétique au dix-huitième siècle. En effet, la notion de jugement esthétique, conçue à l'époque en tant qu'affirmation portant sur la beauté d'un objet (« x est beau », « y est élégant », etc.) a alimenté des querelles épistémologiques sur la nature et le statut de tels jugements. Par exemple : à quelles conditions puis-je dire, si j'affirme que « x est beau », que mon affirmation est *vraie* ? On l'aura deviné : l'enjeu est de déterminer à quelles conditions ce type d'affirmation peut avoir un caractère objectif et par conséquent à quelles conditions on pourrait en faire la pierre d'assise d'un discours sur le beau et sur la notion de plaisir. Le traitement

Zeimbekis, John. *Qu'est-ce qu'un jugement esthétique?* Vrin, coll. Chemins philosophiques, 2006, 128 p.

Mélissa Thériault

Volume 34, numéro 2, automne 2007

URI : <https://id.erudit.org/iderudit/017430ar>

DOI : <https://doi.org/10.7202/017430ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Société de philosophie du Québec

ISSN

0316-2923 (imprimé)

1492-1391 (numérique)

[Découvrir la revue](#)

Citer ce compte rendu

Thériault, M. (2007). Compte rendu de [Zeimbekis, John. *Qu'est-ce qu'un jugement esthétique?* Vrin, coll. Chemins philosophiques, 2006, 128 p.] *Philosophiques*, 34 (2), 421–425. <https://doi.org/10.7202/017430ar>

mécanique quantique, mais il est certain qu'il n'y a pas de champ magnétique sans dipôle magnétique, et pas de dipôle magnétique sans un support quelconque (aimant, courant électrique, corps céleste, etc.), de même qu'il n'y a pas de champ magnétique actuel sans électrons qui se meuvent d'un pôle à l'autre. En somme, on n'observe point de champ magnétique sans support.

Plus loin dans le même chapitre, Rescher présente deux nouveaux exemples : la chute de la demande économique et l'érosion de la ligne côtière. Or, de la même façon, la chute de la demande économique est une statistique induite à partir d'un nombre déterminé d'individus, et l'érosion de la ligne côtière est toujours constituée d'un nombre déterminé de processus spécifiques de dissociation, ayant chacun pour support un nombre déterminé de grains de sable (ou autre), spatiotemporellement localisables.

Ainsi, parce que les thèses générales de la philosophie du processus reposent sur la thèse de la séparabilité de certains processus et que Rescher n'a pas démontré que ceux-ci existent effectivement de façon séparée, alors la philosophie (rescherienne) du processus ne repose sur aucun fondement satisfaisant, du moins dans sa version ontologique, dite « forte » (et dans la mesure où la version épistémologique, dite « faible », reposerait sur les mêmes fondements, alors elle serait tout aussi problématique). En effet, sans la thèse de l'existence de processus impropres, celles de la primauté causale et du caractère fondamental des processus se dissolvent, car pour être premier ou fondamental par rapport à autre chose il faut d'abord que l'un puisse se trouver sans l'autre.

La traduction est généralement conforme. Toutefois la reddition de *owned* et *unowned* par « propre » et « impropre » est malheureusement cryptique. De plus, le choix surprenant de « procès » plutôt que « processus » n'est nulle part sérieusement motivé. Weber se contente d'une note dans le lexique où il écrit que *process* est « parfois traduit, pour des raisons euphoniques, [par] "processus" » (p. 276). Or « processus » est préférable pour des raisons plus sérieuses que la simple euphonie. Parmi ces raisons, notons d'abord que le premier sens de « procès » est juridique, de sorte que l'expression « ontologie du procès » pourrait être méprise pour une espèce de l'ontologie du droit et, ensuite, que l'usage déjà établi du terme processus dans le vocabulaire francophone de la philosophie n'a rien de problématique et qu'il est donc superficiel de lui chercher un substitut.

FRÉDÉRIC TREMBLAY

SUNY at Buffalo et IFOMIS (Saarbrücken)

Zeimbekis, John. *Qu'est-ce qu'un jugement esthétique ?* Vrin, coll. Chemins philosophiques, 2006, 128 p.

Cette nouvelle parution relève le défi de présenter une notion incontournable qui a alimenté les querelles et débats en esthétique au dix-huitième siècle. En effet, la notion de jugement esthétique, conçue à l'époque en tant qu'affirmation portant sur la beauté d'un objet (« x est beau », « y est élégant », etc.) a alimenté des querelles épistémologiques sur la nature et le statut de tels jugements. Par exemple : à quelles conditions puis-je dire, si j'affirme que « x est beau », que mon affirmation est *vraie* ? On l'aura deviné : l'enjeu est de déterminer à quelles conditions ce type d'affirmation peut avoir un caractère objectif et par conséquent à quelles conditions on pourrait en faire la pierre d'assise d'un discours sur le beau et sur la notion de plaisir. Le traitement

philosophique de cette notion a connu des suites multiples. Si Kant a proposé une analyse suffisamment convaincante aux yeux de certains pour qu'ils élèvent sa proposition au statut de dogme (sous la forme du jugement de *goût*), d'autres, en revanche, ont relégué le jugement esthétique au rayon de l'indiscutable en raison de son caractère (en apparence du moins) hautement subjectif. Si « des goûts et des couleurs, il ne faut pas discuter », comment prétendre construire un discours argumentatif fondé sur des jugements portant, justement, sur l'appréciation de ceux-ci ?

L'ouvrage est divisé en deux sections : dans la première, la problématique est présentée et analysée d'un point de vue général. La seconde est une analyse de textes portant sur un extrait de « De la règle du goût » de Hume (la section où il est question de la figure de l'homme de goût est toutefois amputée) et un texte plus récent de Malcolm Budd traduit sous le titre « Jugements esthétiques, principes esthétiques et propriétés esthétiques ». La toute première phrase de l'ouvrage est une mise en garde non superflue : la possibilité même de l'existence du jugement esthétique est remise en question. À cette fin, l'auteur explique comment l'expression s'est imposée dans le langage et précise qu'elle est peut-être, justement, en quelque sorte un abus de langage puisque « *il est possible que les jugements esthétiques ne soient pas des jugements en un sens substantiel et philosophiquement admis du terme* » et que l'on « *ne peut exclure qu'ils soient fondés sur une forme d'appréciation qui relève de la raison pratique* » (p. 7). Autrement dit, bien que cette manière d'exprimer la problématique propre au langage des Lumières soit entrée dans les conventions langagières, la fréquence de son emploi ne met pas la notion à l'abri de la remise en question, et le débat qui l'entoure est loin d'être clos. Toutefois, c'est l'ambition de plusieurs philosophes de montrer « *qu'il existe des jugements proprement esthétiques* » et dans quels cas ils peuvent être bien fondés, notamment en tentant de cerner les conditions de possibilité des jugements esthétiques ainsi que les circonstances dans lesquelles celles-ci pourraient être satisfaites, ne serait-ce qu'en principe (p. 8).

L'ouvrage donne la parole à deux figures historiques incontournables de la discipline (Hume et Kant) ainsi qu'à un interlocuteur plus récent (Frank Sibley) en établissant les parallèles et distinctions qui s'imposent entre leurs solutions respectives. On souligne aussi à juste titre que la problématique des jugements esthétiques est revue dans le contexte actuel à la lumière de spécialisations telles que la psychologie cognitive, l'épistémologie et la philosophie de la perception plutôt que dans le cadre d'une *philosophie de l'art*, qui ne recoupe aujourd'hui que partiellement les préoccupations qui étaient celles de l'esthétique philosophique classique. Comme le rappelle à juste titre Zeimbekis, tous les jugements portant sur les œuvres ne sont pas pour autant des jugements *esthétiques* (p. 13). On s'attarde aussi à tracer les grandes lignes du contexte historique dans lequel s'insérait la notion, et ce, d'une façon claire et détaillée (p. 8-9). Par exemple, si la notion de *beauté* peut nous apparaître comme un concept central dans l'esthétique du dix-huitième siècle, elle était loin d'être dominante (les notions de *sublime* ou de *grâce* occupaient à certains moments des rôles au moins aussi importants (p. 9). Le concept de beauté n'est plus incontournable non plus en esthétique contemporaine et encore moins en philosophie de l'art, surtout depuis que le rapport entre art et beauté s'est effrité avec l'avènement de la modernité artistique. Zeimbekis souligne la variété des concepts esthétiques employés ainsi que les écueils potentiels qui se présentent lorsqu'on se limite aux concepts génériques (beau/laid) possédant une lourde charge évaluative (p. 10). Il précise que l'usage de l'expression « expérience esthétique » n'implique pas la *croissance* en une sorte d'expérience *particulière* qui serait ainsi désignée :

une expérience esthétique est une expérience qui nous porte à dire d'une chose que celle-ci est élégante (p. 11).

On effectue aussi un nécessaire retour sur la distinction entre propriétés *esthétiques* et propriétés *artistiques* : souvent, les affirmations à propos des œuvres d'art portent sur des qualités relevant du contenu représentationnel ou du potentiel symbolique (p. 13) plutôt que des propriétés esthétiques. En ce sens, ce qu'on dit des œuvres d'art prend rarement la forme d'un *jugement esthétique*, même si de telles affirmations sont souvent traitées par l'esthétique en tant que discipline (p. 14). Cependant, puisque l'ouvrage doit se limiter aux jugements esthétiques, l'auteur s'attarde peu à cette nuance pourtant cruciale, pour effectuer plutôt un rappel sur les premières théories autour du jugement esthétique (mises de l'avant notamment par Hutcheson). Celles-ci postulaient l'existence d'un sens interne, le « sens esthétique », c'est-à-dire une « *capacité passive d'avoir des impressions, pour expliquer comment nous jugeons esthétiquement les objets* » (p. 15). Celle-ci agirait comme intermédiaire entre les sensations extérieures reçues par les sens et le sentiment de plaisir/déplaisir éprouvé par le sujet (p. 15) et permettrait en quelque sorte de protéger les jugements esthétiques de la contingence (p. 16). Pourtant, les jugements esthétiques en tant que « *symptômes observables de nos réactions esthétiques* » ne font pas consensus, ce qui met à mal cette hypothèse chère aux empiristes (p. 17). Viennent ensuite les théories physiologiques, dont celle de Burke (qui, au dix-huitième siècle, relie les expériences esthétiques à des instincts) ainsi que sa variante contemporaine neurologique, mise de l'avant par V. Ramachandran, qui s'appuie sur une réflexion autour des mécanismes de la cognition perceptive et cherche à baser l'objectivité des jugements sur une sorte de réalisme biologique (p. 21). Burke formule sa théorie en termes de *sentiments* esthétiques (sentiment de sublime, de beau) et ses limites sont, selon Zeimbekis, surmontées par la variante contemporaine de Ramachandran qui prend assise sur des notions évolutionnistes et tente d'expliquer les réactions esthétiques en « *analysant le fonctionnement de notre physiologie perceptive* », ce qui permettrait de générer des « lois de l'art » que les artistes utiliseraient consciemment ou non pour stimuler les régions visuelles du cerveau (p. 27). Malheureusement, il n'est pas mentionné si la thèse de Ramachandran vise à s'appliquer à autre chose qu'aux arts visuels : on présente cependant les différentes objections qui ont été faites à cette approche, dont sa circularité (p. 29).

L'essentiel de l'analyse de Zeimbekis porte sur la formulation de la problématique par Hume qui a abordé la question sous un angle nouveau : en cherchant à soumettre le concept de jugement esthétique à l'épreuve du scepticisme, il tente de le reformuler sur des assises résistant à ses propres objections (p. 33) afin de répondre à la question : « *comment, et sous quelles conditions, un jugement fondé sur un sentiment peut-il être valide ?* » (p. 37). Certaines distinctions s'imposent alors, par exemple entre jugements objectifs et jugements portant sur des propriétés objectives (p. 38). Après avoir déterminé à quelles conditions un jugement pouvait être objectif, il importe ensuite de voir « *s'il y a bien des faits indépendants qui puissent conférer aux jugements esthétiques le statut de représentations vraies ou fausses* » (p. 39). Une partie de la solution proposée par Hume consiste à attribuer aux désaccords liés au goût aux capacités de *discernement* du sujet : le critique idéal est plus sensible aux nuances et détails parce qu'il a affiné son goût, notamment par la pratique (p. 39). Mais malgré ses tentatives de situer objectivement le jugement de goût, Hume est relégué au camp conventionnaliste dès lors qu'il se résigne à croire que les faits esthétiques ne peuvent être connus que par consensus (p. 43), et on pourrait être tenté, selon Zeimbekis, de lui reprocher de n'avoir comme solution

de rechange au problème du subjectivisme individuel qu'une proposition de substitution par un subjectivisme de groupe (p. 44). C'est alors que sont présentées les solutions de rechange proposées par Kant aux problèmes auxquels Hume n'aurait pas proposé de solution satisfaisante.

Si Hume a tenté de donner au jugement esthétique le statut d'un jugement de faits pour surmonter le défi subjectiviste (position selon laquelle un jugement de sentiments n'est ni vrai ni faux) (p. 44), Kant a procédé selon le chemin inverse en acceptant le caractère subjectif du jugement esthétique mais en lui attribuant une forme de validité particulière (p. 45). C'est le concept même de jugement qui est ainsi transformé, ce qui aura des répercussions au-delà de l'esthétique kantienne déployée dans *Critique de la faculté de juger*. La validité du jugement esthétique dépend de certaines conditions cognitives très précises (p. 45), conditions qui sont explicitées par Kant et distinguées des circonstances où l'on pourrait être porté à croire à tort que notre jugement de goût est bien fondé. La faculté de juger n'est toutefois pas confinée au jugement esthétique et « *est également requise pour accomplir d'autres tâches cognitives* » (p. 45). Quant à l'origine de la satisfaction esthétique, elle est résumée ainsi : « *mon activité cognitive est satisfaisante quand je contemple la forme parce que la forme, étant finalisée, convient aux principes de la faculté de juger, qui gouverne mes tentatives perceptives de cerner des formes finalisées* » (p. 47) : c'est en présence de ce type de satisfaction que l'on qualifie l'objet l'ayant provoquée de « beau ». Kant cherche à expliquer qu'il y a des jugements auxquels on peut attribuer une forme de validité sans en faire pour autant des jugements de faits (p. 47). Le jugement de goût qui, dans le contexte humien était un jugement de sentiment et non factuel, possède selon lui ces caractéristiques : il est non conceptuel, ne peut attribuer de propriété à l'objet et est fondé sur un sentiment de plaisir (p. 49), mais il est valable sur le plan intersubjectif (p. 50). La théorie proposée par Kant comporte aussi ses points de tension, notamment en ce qu'elle oblige à accepter un postulat quant à l'existence d'*idées esthétiques*, principes régulateurs de réactions esthétiques inhérents à la faculté de juger (p. 50). L'auteur relève alors que la prémisse des idées esthétiques « *joue exactement le même rôle que le postulat métaphysique de Hume* » (p. 51) et renvoie les deux théories à dos en leur reprochant un point commun, à savoir, de fonder la normativité des jugements esthétiques « *sur des postulats douteux* », en l'occurrence, selon Zeimbekis, des faits invérifiables dans le cas de Hume et des principes inconnaissables du côté de Kant (p. 56).

L'auteur effectue ensuite une parenthèse visant à faire un état des lieux quant aux différents modèles du jugement esthétique : l'approche la plus courante est de poser le problème en tant que « *relation entre un particulier et un concept évaluatif* », ce qui inclut bien sûr les approches issues de l'empirisme, du réalisme et de la philosophie kantienne pour lesquelles l'auteur avoue son parti pris (p. 57). Il signale aussi l'existence d'un autre type d'approche, représenté par les théories qui « *ne portent pas sur le jugement esthétique mais sur le contenu des œuvres d'art* », parmi lesquelles celles de Hegel et d'Adorno font figure d'illustres exemples (p. 57). À cela s'ajoutent les théories relativistes qui « *critiquent le concept même de jugement esthétique* » : par exemple, Genette fait usage du concept d'*expérience* esthétique mais ne reconnaît pas la validité du jugement qui pourrait être formulé à la suite de cette expérience par le sujet (p. 57). Un autre modèle regroupe les théories « *qui défendent la possibilité qu'il existe des jugements esthétiques bona fide* » (p. 58) et qui se heurtent aux objections classiques évoquées tout au long de l'ouvrage. Parmi celles-ci, la variante proposée par F. Sibley dans les années soixante, qui relève trois types de concepts circonscrits par des frontières

souvent flous : concepts non esthétiques, esthétiques et évaluatifs (p. 59). La théorie de Sibley repose sur une analyse du rapport qu'ils entretiennent les uns aux autres et comporte le mérite de mieux répondre aux critiques de relativisme qu'on pouvait faire aux théories évoquées précédemment (p. 68). Mais elle n'est pas exempte non plus de difficultés, soulignons par exemple le caractère culturellement relatif des concepts esthétiques (p. 71).

La seconde partie de l'ouvrage, composée du commentaire des textes de Hume et de Budd, permet de situer les notions vues précédemment dans une perspective plus générale. Puisque la théorie humienne avait été présentée plus en détail auparavant, Zeimbekis propose un rappel de l'opposition entre conception relativiste (motivée par un constat empirique) et réaliste telle que présentée par Hume (p. 80-81). On insiste aussi sur les préoccupations « méta-esthétiques » de Hume, moins intéressé par la définition des concepts esthétiques que par une discussion autour de leur légitimité (p. 84) ainsi que sur les limites de la forme de réalisme proposée par Hume (p. 90). Le texte de Budd consiste pour sa part en un court extrait qui discute une interprétation contemporaine de Hume, coupé toutefois du contexte de discussion dans lequel il s'insérait. On cherche à y repousser le relativisme des jugements esthétiques en défendant l'existence de « *principes esthétiques logiquement faibles* » (p. 108). Le principal mérite de l'extrait choisi est de servir de prétexte à l'introduction de certaines figures contemporaines dont Alan Goldman et Eddy Zemach.

En conclusion, disons que cette publication relève le défi de rendre compte d'une problématique qui a une longue histoire tout en réussissant à résister à la tentation du parti pris et en sachant mettre en valeur la problématique. La commande implique en soi une grande neutralité de l'auteur, exigence qui a été respectée admirablement, d'autant plus que le format imposé par la collection oblige un traitement très schématique qui ajoute à la difficulté de la tâche. L'auteur a également su mettre en perspective les théories autour du jugement esthétique : bien que populaires chez les historiens qui s'intéressent à l'esthétique, ou dans leurs variantes contemporaines, elles ne sont pas les seules approches possibles, et la pertinence même de traiter de la notion de jugement esthétique a été remise en question par plusieurs. Notons que ce petit volume remplira sa fonction principalement auprès d'un lectorat déjà familier et sympathique aux théories issues de l'empirisme. Mais ce défaut a une contrepartie positive : on va à l'essentiel, ce qui permet de comprendre les enjeux de ce problème historique qui suscite toujours l'intérêt de plusieurs.

MÉLISSA THÉRIAULT

UQAM/Université de Provence

Jean-François Filion, *Dialectique et matière. La conceptualité inconsciente des processus inorganiques dans la Philosophie de la nature (1830) de Hegel*, Québec, Les Presses de l'Université Laval, 2007, 346 p.

Cet ouvrage impose d'abord par son ambition. Reconnaissons d'emblée que la philosophie de la nature n'est ni la partie la plus connue du système hégélien ni encore moins la plus considérée. Or, dans cette étude, qui constitue une version remaniée de sa thèse de doctorat, Jean-François Filion se propose non pas simplement de réhabiliter la philosophie de la nature de Hegel — d'autres auteurs s'y sont déjà

Jean-François Filion, *Dialectique et matière. La conceptualité inconsciente des processus inorganiques dans la Philosophie de la nature (1830) de Hegel*, Québec, Les Presses de l'Université Laval, 2007, 346 p.

Danic Parenteau

Volume 34, numéro 2, automne 2007

URI : <https://id.erudit.org/iderudit/017431ar>

DOI : <https://doi.org/10.7202/017431ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Société de philosophie du Québec

ISSN

0316-2923 (imprimé)

1492-1391 (numérique)

[Découvrir la revue](#)

Citer ce compte rendu

Parenteau, D. (2007). Compte rendu de [Jean-François Filion, *Dialectique et matière. La conceptualité inconsciente des processus inorganiques dans la Philosophie de la nature (1830) de Hegel*, Québec, Les Presses de l'Université Laval, 2007, 346 p.] *Philosophiques*, 34 (2), 425–427.
<https://doi.org/10.7202/017431ar>

souvent flous : concepts non esthétiques, esthétiques et évaluatifs (p. 59). La théorie de Sibley repose sur une analyse du rapport qu'ils entretiennent les uns aux autres et comporte le mérite de mieux répondre aux critiques de relativisme qu'on pouvait faire aux théories évoquées précédemment (p. 68). Mais elle n'est pas exempte non plus de difficultés, soulignons par exemple le caractère culturellement relatif des concepts esthétiques (p. 71).

La seconde partie de l'ouvrage, composée du commentaire des textes de Hume et de Budd, permet de situer les notions vues précédemment dans une perspective plus générale. Puisque la théorie humienne avait été présentée plus en détail auparavant, Zeimbekis propose un rappel de l'opposition entre conception relativiste (motivée par un constat empirique) et réaliste telle que présentée par Hume (p. 80-81). On insiste aussi sur les préoccupations « méta-esthétiques » de Hume, moins intéressé par la définition des concepts esthétiques que par une discussion autour de leur légitimité (p. 84) ainsi que sur les limites de la forme de réalisme proposée par Hume (p. 90). Le texte de Budd consiste pour sa part en un court extrait qui discute une interprétation contemporaine de Hume, coupé toutefois du contexte de discussion dans lequel il s'insérait. On cherche à y repousser le relativisme des jugements esthétiques en défendant l'existence de « *principes esthétiques logiquement faibles* » (p. 108). Le principal mérite de l'extrait choisi est de servir de prétexte à l'introduction de certaines figures contemporaines dont Alan Goldman et Eddy Zemach.

En conclusion, disons que cette publication relève le défi de rendre compte d'une problématique qui a une longue histoire tout en réussissant à résister à la tentation du parti pris et en sachant mettre en valeur la problématique. La commande implique en soi une grande neutralité de l'auteur, exigence qui a été respectée admirablement, d'autant plus que le format imposé par la collection oblige un traitement très schématique qui ajoute à la difficulté de la tâche. L'auteur a également su mettre en perspective les théories autour du jugement esthétique : bien que populaires chez les historiens qui s'intéressent à l'esthétique, ou dans leurs variantes contemporaines, elles ne sont pas les seules approches possibles, et la pertinence même de traiter de la notion de jugement esthétique a été remise en question par plusieurs. Notons que ce petit volume remplira sa fonction principalement auprès d'un lectorat déjà familier et sympathique aux théories issues de l'empirisme. Mais ce défaut a une contrepartie positive : on va à l'essentiel, ce qui permet de comprendre les enjeux de ce problème historique qui suscite toujours l'intérêt de plusieurs.

MÉLISSA THÉRIAULT

UQAM/Université de Provence

Jean-François Filion, *Dialectique et matière. La conceptualité inconsciente des processus inorganiques dans la Philosophie de la nature (1830) de Hegel*, Québec, Les Presses de l'Université Laval, 2007, 346 p.

Cet ouvrage impose d'abord par son ambition. Reconnaissons d'emblée que la philosophie de la nature n'est ni la partie la plus connue du système hégélien ni encore moins la plus considérée. Or, dans cette étude, qui constitue une version remaniée de sa thèse de doctorat, Jean-François Filion se propose non pas simplement de réhabiliter la philosophie de la nature de Hegel — d'autres auteurs s'y sont déjà

attelés de manière convaincante avant lui, notamment Vittorio Hösle dans sa monumentale étude *Hegels System* (1987) —, mais plutôt dans une démarche plus engagée, qui consiste à « prendre la spéculation au sérieux » (p. 3). Malgré les deux siècles qui nous séparent d'elle, la pensée hégélienne, pour peu qu'on prenne toute la mesure de sa « radicalité » (p. 2), peut encore nous aider à penser notre temps.

La thèse que défend J.-F. Filion est que la pensée spéculative hégélienne ouvre une piste de réflexion nouvelle et porteuse permettant de repenser la présente crise écologique. Comme on peut le lire, elle offre « des munitions à la révolution culturelle de l'écologisme » (p. 4). Selon l'auteur, toute l'originalité et la force de cette pensée tient au « renversement ontologique » (p. 317) qu'elle réussit à opérer quant au rapport entre l'être humain et la nature ; rapport dont d'aucuns, parmi les penseurs préoccupés par la question écologique, admettent qu'il s'agit là de la cause la plus profonde de la présente crise. En effet, au cœur de la menace qui guette l'intégrité écologique de l'humanité réside une conception de la nature, fruit d'une tradition métaphysique qui remonte en Occident au début des sciences modernes, qui voit dans celle-ci rien de plus qu'une simple matière brute, un matériau dont l'être humain peut disposer à sa guise pour la poursuite de ses propres fins. La nature n'est toujours considérée qu'en tant qu'altérité radicale, comme l'autre de l'humanité.

Or précisément, toute la philosophie spéculative de Hegel repose sur l'idée qu'il existe dans la nature une rationalité « inconsciente » (p. 23), une raison qui est et qui se donne à voir sous la forme manifeste d'une nature inorganique. Dans l'idéalisme absolu hégélien, la nature n'est point saisie en tant qu'altérité radicale, mais bien toujours en tant que simple moment d'un tout qui l'englobe, ce que Hegel nomme l'« Idée » : « La nature et l'esprit partagent le fait d'être les deux modes de la manifestation de l'Idée » (p. 112). Le réel naturel et le réel spirituel s'offrent dans la philosophie spéculative comme deux moments de la manifestation effective (*wirkliches*) de l'Idée. Autrement dit, « la clef interprétative de la conception hégélienne de la nature inorganique consiste à voir dans celle-ci la forme "aliénée" de la vie cosmique qui se désaliènera ici, sur Terre, avec l'apparition de la vie biologique et d'une espèce animale spéciale qui peut en parler » (p. 18). Aussi, par cette approche spéculative de la nature, la pensée hégélienne permet de surmonter l'aporie à laquelle est confrontée aujourd'hui toute pensée écologique.

À cette fin, *Dialectique et matière* s'articule sous la forme d'une lecture exégétique de la seconde partie de *l'Encyclopédie des sciences philosophiques*, celle qui précisément porte sur la philosophie de la nature, avec un accent mis sur les paragraphes 245 à 336. L'auteur a retenu pour cette lecture la troisième et dernière édition de l'ouvrage, celle de 1830, retravaillée par Hegel avant sa mort, et qui comprend de nombreuses « additions », soit les notes prises par les étudiants pendant ses leçons berlinoises. L'étude de J.-F. Filion s'organise autour de deux parties comportant au total sept chapitres. La première déploie une introduction à la philosophie de la nature chez Hegel avec une analyse de ce qui distingue cette pensée de celle de Friedrich Schelling, penseur à qui sera associée la figure de Hegel, jusqu'à la publication en 1806-1807 de son imposante *Phénoménologie de l'esprit*. La seconde partie s'emploie à expliciter la philosophie de la nature inorganique chez Hegel, soit les paragraphes 253 à 336.

La lecture que nous offre J.-F. Filion est conventionnelle — rien ici d'une lecture inédite — tout en étant néanmoins éclairante. Le travail de pensée ne se situe pas tant dans la lecture de Hegel comme telle que dans ce que cette lecture ouvre comme

possibilité; ce que précisément l'auteur parvient à très bien faire valoir en introduction et en conclusion de l'ouvrage.

Cependant, malgré les grands mérites de cette démarche et la rigueur avec laquelle elle est menée, on ne peut s'empêcher de se demander à quel lectorat un tel ouvrage est destiné. En effet, si cette étude s'adresse avant tout à ceux désireux de trouver dans la pensée spéculative hégélienne une piste d'espoir pour sortir l'humanité de la crise écologique — objectif que l'auteur semble privilégier dans l'introduction — alors, force est d'admettre qu'il manque à cet ouvrage une introduction générale à la pensée de Hegel dans son ensemble. Assurément, un lecteur néophyte risque en de nombreux passages de n'y rien comprendre tant la pensée hégélienne, et peut-être encore plus sa philosophie de la nature, manifeste une impénétrabilité certaine.

Par delà cet élément de critique, ou peut-être même pour y parer, il ne reste maintenant qu'à souhaiter que Jean-François Filion poursuive sur la voie ouverte par cette étude, en mettant à profit l'approche spéculative hégélienne afin de proposer des pistes de solutions pour sortir d'humanité de la présente menace écologique qui pèse sur son habitat. Rendre justice à la profondeur du hégélianisme peut parfois exiger de passer sous silence la pensée du maître.

DANIC PARENTEAU

École d'études politiques, Université d'Ottawa

Alain Beaulieu (dir.), *Gilles Deleuze, héritage philosophique*, Paris, PUF, 2005, 174 pages.

L'année 2005 a marqué le dixième anniversaire de la mort de Gilles Deleuze. À l'occasion d'une publication dans la collection « Débats philosophiques » des Presses Universitaires de France, Alain Beaulieu entend défendre l'idée que Gilles Deleuze fut « *simultanément* historien, créateur et expérimentateur de concepts » (p. 12). En refusant de dissocier catégoriquement et d'étudier séparément — comme on l'avait généralement fait jusque-là — ces différentes facettes de la vie intellectuelle du philosophe, Alain Beaulieu espère donner un second souffle à la réception posthume de ses écrits. À cette fin, il a réuni dans ce volume les contributions de six des principaux commentateurs de l'œuvre deleuzienne, à la fois d'Europe (Manola Antonioli, Arnaud Villani, Olivier Fahle et Stéfan Leclercq) et de ce côté-ci de l'Atlantique (Constantin Boundas et Alain Beaulieu).

1.

Avec son analyse des « Stratégies différentielles dans la pensée deleuzienne », Constantin Boundas cherche à dégager les réquisits et les conséquences d'une authentique philosophie de la différence. L'auteur nous rappelle tout d'abord que, chez Deleuze, « la différence ne se donne pas comme un concept » (p. 16) puisqu'elle est pensée comme un processus, ce qui implique en retour une ontologie des forces et des relations qui s'oppose à l'étude des différents étants dans leur particularité. Certes, les forces se manifestent nécessairement dans des états de choses distincts les uns des autres, mais puisque ceux-ci ne peuvent rendre compte de leur devenir et du processus dont résulte leur différenciation, Deleuze en est conduit à penser, à la suite de Bergson, une différenciation virtuelle

Alain Beaulieu (dir.), *Gilles Deleuze, héritage philosophique*, Paris, PUF, 2005, 174 pages.

Charles Bolduc

Volume 34, numéro 2, automne 2007

URI : <https://id.erudit.org/iderudit/017432ar>

DOI : <https://doi.org/10.7202/017432ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Société de philosophie du Québec

ISSN

0316-2923 (imprimé)

1492-1391 (numérique)

[Découvrir la revue](#)

Citer ce compte rendu

Bolduc, C. (2007). Compte rendu de [Alain Beaulieu (dir.), *Gilles Deleuze, héritage philosophique*, Paris, PUF, 2005, 174 pages.] *Philosophiques*, 34 (2), 427–431. <https://doi.org/10.7202/017432ar>

possibilité; ce que précisément l'auteur parvient à très bien faire valoir en introduction et en conclusion de l'ouvrage.

Cependant, malgré les grands mérites de cette démarche et la rigueur avec laquelle elle est menée, on ne peut s'empêcher de se demander à quel lectorat un tel ouvrage est destiné. En effet, si cette étude s'adresse avant tout à ceux désireux de trouver dans la pensée spéculative hégélienne une piste d'espoir pour sortir l'humanité de la crise écologique — objectif que l'auteur semble privilégier dans l'introduction — alors, force est d'admettre qu'il manque à cet ouvrage une introduction générale à la pensée de Hegel dans son ensemble. Assurément, un lecteur néophyte risque en de nombreux passages de n'y rien comprendre tant la pensée hégélienne, et peut-être encore plus sa philosophie de la nature, manifeste une impénétrabilité certaine.

Par delà cet élément de critique, ou peut-être même pour y parer, il ne reste maintenant qu'à souhaiter que Jean-François Filion poursuive sur la voie ouverte par cette étude, en mettant à profit l'approche spéculative hégélienne afin de proposer des pistes de solutions pour sortir d'humanité de la présente menace écologique qui pèse sur son habitat. Rendre justice à la profondeur du hégélianisme peut parfois exiger de passer sous silence la pensée du maître.

DANIC PARENTEAU

École d'études politiques, Université d'Ottawa

Alain Beaulieu (dir.), *Gilles Deleuze, héritage philosophique*, Paris, PUF, 2005, 174 pages.

L'année 2005 a marqué le dixième anniversaire de la mort de Gilles Deleuze. À l'occasion d'une publication dans la collection « Débats philosophiques » des Presses Universitaires de France, Alain Beaulieu entend défendre l'idée que Gilles Deleuze fut « *simultanément* historien, créateur et expérimentateur de concepts » (p. 12). En refusant de dissocier catégoriquement et d'étudier séparément — comme on l'avait généralement fait jusque-là — ces différentes facettes de la vie intellectuelle du philosophe, Alain Beaulieu espère donner un second souffle à la réception posthume de ses écrits. À cette fin, il a réuni dans ce volume les contributions de six des principaux commentateurs de l'œuvre deleuzienne, à la fois d'Europe (Manola Antonioli, Arnaud Villani, Olivier Fahle et Stéfan Leclercq) et de ce côté-ci de l'Atlantique (Constantin Boundas et Alain Beaulieu).

1.

Avec son analyse des « Stratégies différentielles dans la pensée deleuzienne », Constantin Boundas cherche à dégager les réquisits et les conséquences d'une authentique philosophie de la différence. L'auteur nous rappelle tout d'abord que, chez Deleuze, « la différence ne se donne pas comme un concept » (p. 16) puisqu'elle est pensée comme un processus, ce qui implique en retour une ontologie des forces et des relations qui s'oppose à l'étude des différents étants dans leur particularité. Certes, les forces se manifestent nécessairement dans des états de choses distincts les uns des autres, mais puisque ceux-ci ne peuvent rendre compte de leur devenir et du processus dont résulte leur différenciation, Deleuze en est conduit à penser, à la suite de Bergson, une différenciation virtuelle

des tendances qui ont généré ces distinctions actuelles. Dans la perspective d'une telle ontologie, loin d'être purement informes ou à l'inverse totalement réifiées, les forces sont donc à la fois des grandeurs extensives (différenciées) et intensives (différentiées). Ce départage de l'actuel et du virtuel — tous deux bel et bien réels — amène Constantin Boundas à présenter, dans une seconde partie, la première conséquence philosophique de cette théorie deleuzienne de la différence, soit la remise en question d'une certaine « image de la pensée » qui, en opposant le possible au réel, marque la prédominance de la recognition et de la représentation dans l'exercice de la pensée. Cette critique n'est cependant pas pour nous surprendre, puisqu'une telle conception de la pratique philosophique réduit l'étude des différences réelles à l'élucidation d'une forme commune d'expérience possible, ce dont témoignent par exemple les philosophies transcendantales kantienne et husserlienne. Il s'ensuit conséquemment une redéfinition de la philosophie comme création de concepts et une nouvelle méthode — l'empirisme transcendantal — adaptée au souci de ne pas trahir le concret dans sa singularité : si l'état de choses rencontré dans l'expérience est étudié sous le signe de sa différence constitutive, alors, ce qui est dégagé, ce n'est pas la forme qu'il partage avec d'autres états de choses, mais ses conditions réelles — les tendances virtuelles non représentables — qui en expriment la genèse en tant qu'événement à nul autre pareil. Ces considérations amènent enfin Constantin Boundas à s'interroger sur les conséquences éthiques d'une telle prise en considération de la différence. D'une part, il constate que, pour Deleuze, celle-ci s'accompagne d'une lutte active pour démasquer les prétentions indues de l'actuel à vouloir s'arroger tout le champ du réel. D'autre part, suivant encore une fois Deleuze, ce combat apparaît inséparable d'un renversement des rapports entre l'éthique et la morale. Ainsi, contrairement à ce que l'on pourrait croire de prime abord, ce ne sont jamais des valeurs morales transcendantes qui imposent une conduite particulière : pour qu'une personne accepte de se plier à certaines obligations, celles-ci doivent d'abord avoir été évaluées, c'est-à-dire précédées par une expérience éthique qui en a mesuré la teneur en puissance de vie.

2.

La contribution d'Alain Beaulieu intitulée « Gilles Deleuze et les Stoïciens » questionne l'une des références deleuziennes les plus constantes mais étrangement peu étudiée jusqu'à maintenant. Après avoir rappelé que Deleuze ne pense jamais à partir des distinctions historiques qui divisent le stoïcisme en trois grandes périodes (ancien, moyen et impérial) et après avoir souligné que, par un regard plus attentif aux sources utilisées, Deleuze semble privilégier le stoïcisme athénien (stoïcismes ancien et moyen) plus théorique, au détriment du stoïcisme romain (impérial) plus soucieux des pratiques de soi, l'auteur en vient à la conclusion que, malgré ces distinctions historico-critiques, l'important est que « l'œuvre deleuzienne rappelle l'ancienne démarche stoïcienne qui consiste à développer la philosophie en trois parties complémentaires : la logique, la physique et l'éthique » (p. 50). L'article ouvre alors sur une étude tripartite où, pour chacune de ces rubriques, Alain Beaulieu recense les points de convergence et de divergence entre les principales thèses stoïciennes et celles avancées par Deleuze. D'un côté, tout comme l'ont fait à leur manière les stoïciens, Deleuze substitue une logique propositionnelle à une logique prédicative, pourfend la recherche d'une origine transcendante du monde et enfin promeut un mode de vie qui épouse les forces qui conduisent le monde. D'un autre côté cependant — et c'est ce qui distingue sa pensée des enseignements stoïciens —, puisqu'il ne considère pas la Nature comme étant un système totalement

organisé, la recherche d'une logique universelle pouvant exprimer tous les événements perd tout son sens, et le philosophe ne peut viser un idéal de sérénité. Cela étant dit, bien qu'Alain Beaulieu ait fait un relevé minutieux de ces différents points de rencontre et de désaccord, ce travail mériterait d'être complété par un autre qui départagerait cette fois non pas Deleuze et les stoïciens, mais d'un côté l'héritage que lui ont laissé les principaux commentateurs français du stoïcisme (Émile Bréhier et Victor Goldschmidt) et de l'autre ce qui revient en propre à la lecture deleuzienne de cette tradition.

3.

Avec « La machination politique de Deleuze et Guattari », Manola Antonioli prend le prétexte du politique pour remettre en question les dualismes à partir desquels on a l'habitude d'interpréter leur philosophie. Alors qu'une certaine vulgate la réduit bien souvent à un parti pris pour les notions de minorité, de micropolitique et de déterritorialisation (par opposition à celles de majorité, de macropolitique et de territorialité), l'auteur nous rappelle que, pour Deleuze et Guattari, tout est une question de tendances, et que le crédit accordé à certaines d'entre elles ne relève pas d'une évaluation réalisée dans la perspective d'un idéal à atteindre. C'est ainsi que leur critique du capitalisme se déploie de l'intérieur même du système : la remise en question de tous les codes sociaux étant inhérente au devenir du capitalisme, les mouvements qui s'agitent sous les institutions qui le régulent actuellement à l'échelle mondiale ne peuvent être critiqués unilatéralement au nom de ces mêmes institutions, puisque ces dernières sont justement elles-mêmes le résultat de processus micropolitiques de déterritorialisation. Du coup, si ces derniers revêtent généralement une valeur positive aux yeux de Deleuze et Guattari, ce n'est pas parce qu'ils visent la perte et la destruction de tous les repères, mais uniquement parce qu'ils ne cherchent pas à mesurer la valeur des devenirs sociaux à l'aune d'un code préexistant qui, en occultant sa propre origine et en prenant l'allure d'une norme transcendante, ne fait que réprimer sournoisement et injustement tout désir de changement. Ce n'est donc pas la macropolitique en tant que telle qui est dévaluée (Deleuze et Guattari soulignent souvent les dangers d'une déterritorialisation effrénée et incessante), mais seulement sa tendance à écarter du revers de la main les solutions de rechange qui, d'après elle, ne sont pas viables ou réalistes.

4.

Lecteur fidèle et attentif des écrits de Gilles Deleuze, Arnaud Villani entend pour sa part contribuer aux débats suscités par son œuvre en critiquant les interprétations qu'en donnent Mireille Buydens et Jacques Rancière. Dans « De l'esthétique à l'esthésique : Deleuze et la question de l'art », il remet en cause leurs critères d'évaluation qu'il juge inadéquats pour saisir l'originalité du propos du philosophe. D'un côté, Mireille Buydens soutient qu'il n'y a pas d'esthétique chez Deleuze puisque la forme serait selon lui toujours seconde et contingente par rapport aux forces à exprimer. D'un autre côté, Jacques Rancière en arrive lui aussi à la même conclusion quand il est question d'une potentielle esthétique typiquement deleuzienne puisque, d'après lui, le philosophe se fait le chantre de l'impossibilité de toute œuvre d'art. Selon Arnaud Villani, ces jugements sont irrecevables, car ils ne tiennent pas compte du fait que, chez Deleuze, tout exercice de la pensée met en jeu deux faces (pensée et non-pensée), deux tendances qui s'affrontent continuellement, et que c'est précisément parce que l'esthétique traditionnelle occulte toujours l'une de ces deux faces que Deleuze n'a pas développé une « esthétique » propre, c'est-à-dire une théorie de l'œuvre d'art accomplie et dominée

par la pensée. En d'autres termes, ce ne serait pas l'œuvre d'art qui serait disqualifiée ou jugée impossible par Deleuze, mais seulement l'esthétique traditionnelle, d'où s'explique le dès lors nécessaire passage de l'esthétique (triomphe du pouvoir de la forme dans la représentation) à l'esthésique (« la critique du pouvoir, non de la pensée, mais dans la pensée » [p. 118]). Ainsi, selon Arnaud Villani : « *Il n'y a pas d'allégorie de la fin de l'art chez Deleuze, mais ce qui se passe dans l'art est l'archétype de ce qui constitue le réel* » (p. 118).

5.

Dans « La visibilité du monde. Deleuze, Merleau-Ponty et le cinéma », Olivier Fahle tente un rapprochement entre leurs réflexions respectives sur la question de l'image. D'un côté, d'après Merleau-Ponty, le cinéma nous fait comprendre l'évolution récente des théories psychologiques sur la perception qui, définissant auparavant l'image mentale comme une synthèse intellectuelle d'impressions sensorielles, la considèrent maintenant comme un processus spontané toujours déjà structuré. D'un autre côté, chez Deleuze, le passage de l'image-mouvement à l'image-temps est l'expression d'un décentrement du primat de l'action représentée vers celui d'un visible qui n'est plus subordonné à la narration chronologique de cette action. Selon Oliver Fahle, ces deux analyses trouvent leur point de convergence dans le fait que, pour les deux philosophes, le visible en vient progressivement à se définir en fonction de l'invisible, ce qui sonne ainsi le glas de la distinction traditionnelle du sujet percevant et de l'objet perçu. Une lecture attentive des derniers écrits de Merleau-Ponty offrirait donc un démenti à l'interprétation qu'en donne Deleuze lui-même, qui confondait peut-être un peu négligemment la méditation de Merleau-Ponty avec une phénoménologie de type husserlienne se fondant en dernière instance sur un sujet transcendantal. Malgré ses qualités indéniables, il me semble qu'un bémol doit cependant être apporté à cette démonstration. Au tout début de son article, l'auteur distingue les conceptions merleau-pontienne et deleuzienne de l'exercice de la pensée : pour la première, « c'est la philosophie qui élabore des notions « applicables » dans les autres domaines » alors que « ce sont pour Deleuze les médias qui représentent les modèles sensoriels d'organisation à partir desquels une explication des notions philosophiques devient possible » (pp. 125-126). Alors qu'on pourrait s'attendre à ce que l'auteur mesure l'importance de cette différence d'approche, il la relativise au contraire sous le prétexte qu'elle n'a aucune importance puisque leur conception du visible est sensiblement la même. Mais n'est-ce pas occulter le fait primordial que, pour Deleuze, la philosophie est essentiellement une création perpétuelle de nouveaux concepts, tandis que pour Merleau-Ponty elle demeure essentiellement l'interrogation et l'explicitation d'une vérité primordiale, soit l'incarnation de l'être humain dans le monde ?

6.

La dernière contribution du collectif porte sur « La réception posthume de l'œuvre de Gilles Deleuze ». Stéfán Leclercq y distingue d'une part la perpétuation de l'œuvre elle-même et d'autre part l'influence qu'elle a eue sur des artistes contemporains. Les divers moyens de diffusion de la philosophie deleuzienne étant largement connus, c'est surtout la deuxième partie de l'article qui retient l'attention. On lit souvent que la pensée de Deleuze a davantage influencé les artistes que les philosophes et, à cet égard, plusieurs publications érudites établissent des liens entre ses concepts et divers mouvements artistiques. Par contre, on ne relève presque jamais ce que les artistes ont vraiment soutiré

de leur lecture des écrits de Deleuze. En laissant la parole à un peintre (Ange-Henri Pieraggi), un poète (Jean-Philippe Cazier) et un dramaturge (Alain Jugnon), l'auteur cherche à combler cette lacune. On y apprend alors quelles sont les idées maîtresses qui les ont le plus influencés, soit respectivement le Figural (par opposition à la figuration), le corps sans organes ouvert aux forces du dehors (par contraste avec l'organisme exclusivement structuré de l'intérieur) et la machination (dirigée contre l'interprétation). Ce qui est par contre un peu regrettable, c'est qu'on a davantage affaire ici à des témoignages de reconnaissance qu'à des confrontations serrées avec les problèmes soulevés par Deleuze, ce qui s'explique il est vrai dans une large mesure par certaines contraintes matérielles, comme l'espace restreint accordé à chacun. On ne peut donc qu'espérer que ce travail sera l'amorce d'un autre plus approfondi.

Tout compte fait, conformément au projet initial défini par Alain Beaulieu, cet ouvrage offre bel et bien un aperçu des multiples facettes de la pensée d'un philosophe qui fut « *simultanément* historien, créateur et expérimentateur de concepts » (p. 12). Ce qui me semble la plus importante contribution de cette publication, ce sont les mises au point faites par les différents auteurs : en reformulant clairement les problèmes soulevés par l'œuvre, ils écartent d'emblée du débat certaines interprétations polémiques mais peu profitables à ceux qui désirent approfondir sincèrement leur compréhension de la philosophie deleuzienne.

CHARLES BOLDUC
Université de Sherbrooke

Livres reçus (automne 2007)

Volume 34, numéro 2, automne 2007

URI : <https://id.erudit.org/iderudit/017433ar>

DOI : <https://doi.org/10.7202/017433ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Société de philosophie du Québec

ISSN

0316-2923 (imprimé)

1492-1391 (numérique)

[Découvrir la revue](#)

Citer ce document

(2007). Livres reçus (automne 2007). *Philosophiques*, 34 (2), 433–434.
<https://doi.org/10.7202/017433ar>

Livres reçus (automne 2007)

- Angelelli, Ignacio, *Études sur Frege et la philosophie traditionnelle*, trad. par J.-F. Courtine,
- A. de Libera, J.-B. Rauzy et J. Schmutz, Paris, Vrin (coll. Problèmes et controverses), 2007, 365 p.
- Angenot, Marc, Eddi, Maï-Linh, Paule-Monique Vernes, *La tolérance est-elle une vertu politique ?*, Québec, Presses de l'Université Laval (coll. Verbatim), 2006, 70 p.
- Billouet, P., J. Gaubert, N. Robinet, et A. Stanguennec, (dir.), *L'homme et la réflexion. Actes du XXX^e congrès de l'Association des sociétés de philosophie de langue française*, Paris, Vrin, 2006, 512 p.
- Clavien, Christine et Catherine El-Bez, (éd.), *Morale et évolution biologique : entre déterminisme et liberté*, Presses polytechniques et universitaires romandes (coll. anthropolos), 2007, 352 p.
- Courtine, Jean-François, *La cause de la phénoménologie*, Paris, Presses Universitaires de France (coll. Épiméthée), 2007, 276 p.
- De Libera, Alain, *Archéologie du sujet*. Vol. I, *Naissance du sujet*, Paris, Vrin (coll. Histoire de la philosophie), 2007, 448 p.
- Engel, Pascal, *Va savoir ! De la connaissance en général*, Paris, Hermann, 2007, 256 p.
- Hintikka, Jaakko, *Les principes des mathématiques revisités*, Paris, Vrin (coll. Mathesis), 2007, 319 p.
- Huglo, Pierre-André, *Sartre : questions de méthode*, Paris, L'Harmattan (coll. Épistémologie et philosophie des sciences), 2005, 105 p.
- Ogien, Ruwen, *L'éthique aujourd'hui. Maximalistes et minimalistes*, Paris, Gallimard (coll. Folio essais), 2007, 253 p.
- Příonský, Franz, *Bolzano contre Kant. Le nouvel anti-Kant*, Paris, Vrin (coll. Textes philosophiques), Paris, Vrin, 2006, 185 p.
- Seron, Denis, *Théorie de la connaissance du point de vue phénoménologique*, Liège, Bibliothèque de la Faculté de philosophie et lettres de l'Université de Liège, 2006, 307 p.

Outre les ouvrages qui apparaissent dans cette liste, la rédaction de *Philosophiques* peut obtenir, pour fin de compte-rendu ou d'étude critique, la plupart des livres parus récemment.

Les personnes désireuses de faire un compte-rendu
ou une étude critique sont priées de s'adresser à :

Jimmy Plourde
Philosophiques
Département de philosophie
Université du Québec à Montréal
c.p. 8888, succ. Centre-ville, Montréal, Qc
H3C 3P8
plourde7@yahoo.com

Index du volume 34 (2007)

Volume 34, numéro 2, automne 2007

URI : <https://id.erudit.org/iderudit/017434ar>

DOI : <https://doi.org/10.7202/017434ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Société de philosophie du Québec

ISSN

0316-2923 (imprimé)

1492-1391 (numérique)

[Découvrir la revue](#)

Citer ce document

(2007). Index du volume 34 (2007). *Philosophiques*, 34 (2), 435–436.

<https://doi.org/10.7202/017434ar>

Index du volume 34 (2007)

Articles

BRISART, ROBERT, <i>Les premières articulations du fonctionnement intentionnel : le projet d'un Raumbuch chez Husserl entre 1892 et 1894</i>	259
CHUNG, RYOA, <i>Domination, vulnérabilité et inégalité d'accès aux soins de santé</i>	133
COUTURE, JOCELYNE et NIELSEN, KAI, <i>Introduction : Cosmopolitisme et particularisme</i>	3
GAGNÉ, LEARRY, <i>La modélisation des comportements non conséquentialistes en théorie du choix rationnel</i>	329
GALIBERT, PABLO, <i>La justice globale, le multiculturalisme et les revendications des immigrants</i>	41
LIMA, MARIA HERRERA, <i>Prendre les autres au sérieux : sur le débat cosmopolitique-particularisme</i>	17
MACLURE, JOCELYN, <i>La reconnaissance engage-t-elle à l'essentialisme ?</i>	77
MOELLENDORF, DARREL, <i>La justice et les associations</i>	61
SMITH, JUSTIN E. H., <i>La génération spontanée et le problème de la reproduction des espèces avant et après Descartes</i>	273
TAN, KOK-CHOR, <i>Nationalisme libéral et internationalisme égalitaire</i>	113
VANDERVEKEN, DANIEL, <i>Principes de pragmatique formelle du discours</i>	229
VILMER, JEAN-BAPTISTE, <i>La prudence de Descartes face à la question de l'infini en mathématiques</i>	295
WENAR, LEIF, <i>L'individu, l'État et les droits de base</i>	97
WOLFE, CHARLES T., <i>Le rêve matérialiste, ou « Faire par la pensée ce que la matière fait parfois »</i>	317

Disputationes

BISHOP, MICHAEL, Notice critique : <i>Reconstructing Reason and Representation</i> — Murray Clarke	367
CLARKE, MURRAY, Précis de <i>Reconstructing Reason and Representation</i>	353
CLARKE, MURRAY, Réponses à mes critiques	385
FAUCHER, LUC, L'intentionnalité, naturellement !	165
FAUCHER, LUC, <i>Faraway, so close ?</i>	375
HUDSON, ROBERT, Améliorer les théories de la fiabilité	363
JACOB, PIERRE, Précis de <i>L'Intentionnalité, problèmes de philosophie de l'esprit</i>	153
JACOB, PIERRE, Réponses à mes critiques	173
VALLÉE, RICHARD, Intentionnalité et référence directe	159

Études critiques

- HACKER, P.M.R. AND BENNETT, M. R. *Philosophical Foundations of Neuroscience* (Pierre Poirier et Nicolas Payette) 183
- ARPALY, N., *Unprincipled Virtue. An Inquiry into Moral Agency* (François Schroeter) 403

Comptes rendus

- BEAULIEU, ALAIN, (dir.), *Gilles Deleuze, héritage philosophique* (Charles Bolduc) 427
- FILION, JEAN-FRANÇOIS, *Dialectique et matière. La conceptualité inconsciente des processus inorganiques dans la Philosophie de la nature (1830) de Hegel* (Danic Parenteau) 425
- GNASSOUNOU, BRUNO, et KISTLER, MAX, *Causes, pouvoirs, dispositions en philosophie. Le retour des vertus dormitives* (Aurélien Robert) 215
- KHUSHF, GEORGE, (dir.), *Handbook of bioethics : Taking stock of the field from a philosophical perspective* (Soumaya Mestiri) 212
- RESCHER, NICOLAS, *Essais sur les fondements de l'ontologie du procès* (Frédéric Tremblay) 419
- RESCHER, NICOLAS, *Metaphysics : The Key Issues From A Realistic Perspective* (Frédéric Tremblay) 217
- SEIDENGART, JEAN, *Dieu, l'univers et la sphère infinie. Penser l'infinité cosmique à l'aube de la science classique* (Yvon Gauthier) 210
- SEYMOUR, MICHEL, *L'institution du langage* (Pascal Engel) 201
- STELZNER, WERNER und KREISER, LOTHAR, *Traditionelle und nichtklassische Logik* (Guillaume Fréchette) 206
- ZARKA, YVES CHARLES, *Un détail nazi dans la pensée de Carl Schmitt* (Luc Vigneault) 220
- ZEIMBEKIS, JOHN. *Qu'est-ce qu'un jugement esthétique ?* (Mélessa Thériault) 421