

Université de Montréal

Nature, Fiabilité et Rôle des Intuitions Utilisées en Philosophie Morale Contemporaine

par Jessy Giroux

**Département de Philosophie
Faculté des Arts et des Sciences**

Mémoire présenté à la Faculté des Arts et des Sciences en vue de l'obtention du grade M.A. en philosophie

Juillet 2012

I. RÉSUMÉ

Ce mémoire porte sur les intuitions utilisées en philosophie morale contemporaine. Les intuitions sont interrogées sur trois aspects : leur nature, leur fiabilité et leur rôle. Concernant la nature des intuitions, celles-ci sont présentées comme des jugements spontanés ayant une valeur épistémique « à première vue ». Concernant la fiabilité des intuitions, trois grands types de raisons de douter de leur fiabilité sont considérés, à savoir les résultats de certaines études scientifiques montrant de diverses manières leur instabilité, les études réalisées par le mouvement de philosophie expérimentale afin d'illustrer le phénomène du désaccord intuitif, et enfin, le problème de l'origine douteuse des intuitions. Aucune de ces raisons n'est jugée efficace afin de discréditer l'usage des intuitions en philosophie. Concernant le rôle des intuitions, une distinction est proposée entre deux grands types d'usage philosophique contemporain. D'une part, ce qu'on peut appeler la méthodologie standard en philosophie morale contemporaine consiste en une combinaison d'expériences de pensée et de la méthode de l'équilibre réfléchi. Différentes manières d'aborder la méthode de l'équilibre réfléchi sont présentées, et le modèle de l'équilibre réfléchi large développé par Norman Daniels est jugé le plus efficace. D'autre part, les intuitions sont parfois utilisées en philosophie morale contemporaine dans le cadre de la théorie de l'intuitionnisme éthique. La théorie est défendue contre diverses objections, et elle est montrée compatible avec la méthode de l'équilibre réfléchi.

MOTS-CLÉS : Intuitions, Nature des intuitions, Rôle des intuitions, Fiabilité des intuitions, Intuitionnisme éthique, Expériences de pensée, Équilibre réfléchi, Méthodologie morale, Épistémologie morale, Justification morale.

ABSTRACT

The object of this Master's thesis is the intuitions used in contemporary moral philosophy. Intuitions are investigated on three aspects: their nature, their reliability, and their role. Regarding the nature of intuitions, they are presented as spontaneous judgments with "prima facie" epistemic value. Regarding the reliability of intuitions, three general types of reasons to question their reliability are considered, namely, the results of certain scientific studies showing in various ways the unreliability of intuitions, studies conducted in the field of experimental philosophy which emphasize the phenomenon of intuitive disagreement, and finally, the problem of the origin of intuitions. None of these reasons is judged good enough to discredit the use of intuitions in philosophy. Regarding the role of intuitions, a distinction is drawn between two main types of usage in contemporary philosophy. On the one hand, what can be called the "standard methodology" in contemporary moral philosophy consists in the combination of thought experiments and the method of reflective equilibrium. Different ways of construing the method of reflective equilibrium are considered, and the "Wide Reflective Equilibrium" model developed by Norman Daniels is judged the most successful. On the other hand, intuitions are sometimes used in contemporary moral philosophy in the context of the theory of ethical intuitionism. The theory is defended against various objections, and it is shown to be compatible with the method of reflective equilibrium.

KEYWORDS : Intuitions, Nature of intuitions, Role of intuitions, Reliability of intuitions, Ethical Intuitionism, Thought Experiments, Reflective Equilibrium, Moral Methodology, Moral Epistemology, Moral Justification, Moral Knowledge.

II. TABLE DES MATIÈRES

I. RÉSUMÉS ET MOTS-CLÉS.....	ii
II. TABLE DES MATIÈRES.....	iv
III. REMERCIEMENTS.....	v
1. INTRODUCTION.....	1
2. QUE SONT LES INTUITIONS?	5
2.1. Usage populaire	6
2.2. Étymologie.....	7
2.3. Faculté « étrange »?	9
2.4. Intuitions, émotions et sciences cognitives	12
2.5. Conclusion : domaine spécifique et crise existentielle	17
3. LES INTUITIONS SONT-ELLES FIABLES?	20
3.1. Recherches scientifiques	20
3.1.1. Économie comportementale : la faiblesse évaluative	21
3.1.2. Sciences cognitives : catégorisation et nature des concepts	25
3.2. Désaccord intuitif et philosophie expérimentale	29
3.3. Origines problématiques	36
3.3.1. La socialisation comme source de nos intuitions	38
3.3.2. Les théories tacites comme sources de nos intuitions	40
3.3.3. Les théories explicites comme sources de nos intuitions	45
3.4. Conclusion	49
4. MÉTHODOLOGIE STANDARD : EXPÉRIENCES DE PENSÉE ET ÉQUILIBRE RÉFLÉCHI	52
4.1. Expériences de pensée	52
4.1.1. Les grands classiques	53
4.1.2. Nature et valeur épistémique	55
4.2. Équilibre réfléchi	59
4.2.1. La méthode rawlsienne	61
4.2.2. L'équilibre réfléchi large	66
4.2.3. Autres modèles	73
4.3. Conclusion	77
5. INTUITIONNISME ÉTHIQUE	80
5.1. Intuitionnisme méta-éthique	82
5.2. Intuitionnisme normatif	89
5.3. Intuitionnisme et équilibre réfléchi	94
5.4. Conclusion	97
6. CONCLUSION	99
7. BIBLIOGRAPHIE	105

III. REMERCIEMENTS

La rédaction d'un mémoire apporte inévitablement un certain lot d'inquiétudes et d'anxiété. Il s'agit après tout du plus important travail que l'étudiant aura eu à accomplir jusqu'à présent dans son parcours académique. Ce vertige initial est cependant largement compensé par l'expérience acquise tout au long du processus de recherche et de rédaction, et bien sûr par le sentiment de devoir accompli une fois le mémoire complété. J'ai personnellement traversé toute cette gamme d'émotions, mais le principal sentiment qui m'a habité du début à la fin fut un sentiment de plaisir! Cette année dédiée à mon mémoire en fut une de pur bonheur! Après tout, j'ai eu le privilège de consacrer tout mon temps à lire et à écrire dans divers cafés et bibliothèques sur un sujet qui me passionnait! Que demander de plus? Cette expérience a dissipé tout doute que je pouvais avoir concernant la poursuite de mes études supérieures en philosophie.

Je ne saurais suffisamment remercier ma directrice de recherche, Christine Tappolet, pour tout l'aide et le soutien qu'elle m'a offerts tout au long de mon parcours de maîtrise. C'est un atout précieux pour un étudiant de savoir qu'il a à sa disposition toute l'érudition et l'expérience d'un chercheur accompli. Christine a toujours été disponible pour m'aider lorsque j'en avais besoin, et ses conseils et suggestions m'ont été d'une grande utilité. Elle fut à bien des égards la superviseure idéale! Je tiens également à la remercier pour m'avoir fait découvrir différentes dimensions de la vie académique, incluant le travail d'assistant-chercheur, la participation à différents colloques et la rédaction d'articles scientifiques. C'est un bagage d'expériences précieux que Christine m'a permis d'acquérir, et je lui en serai toujours reconnaissant.

Je tiens également à remercier Christian Nadeau, qui a toujours été pour moi une inspiration et un modèle, tant sur le plan professionnel que personnel. Son mélange parfait de bonté, d'intelligence et d'engagement m'auront marqué plus que tout au cours de mes six années à l'Université de Montréal. Ce serait également un crime d'oublier de remercier Benoît Dubreuil, qui a été pour moi un véritable mentor à partir de ma dernière année de baccalauréat. Quel étudiant sérieux ne rêverait pas de bénéficier d'un « coach » personnel, aux compétences exceptionnelles, pour le guider à travers l'univers académique? Merci Benoît pour toute ta générosité!

Enfin, merci à mes parents de m'avoir hébergé pendant toute une année de rédaction sans jamais se plaindre (!). Merci, surtout, pour votre amour inconditionnel.

Nature, Fiabilité et Rôle des Intuitions Utilisées en Philosophie Morale Contemporaine

1. INTRODUCTION

La scène se déroule dans un cours de baccalauréat de troisième année portant sur l'éthique et la méta-éthique. La deuxième partie de chaque cours est consacrée à l'échange et au débat entre étudiants sur le thème ayant été abordé en première partie. Cette période donne généralement lieu à des échanges passionnés entre les étudiants, échanges dont moi et des camarades faisons généralement partis. La semaine précédente, après avoir abordé les théories intuitionnistes de G.E. Moore et W. D. Ross, le débat avait porté sur le rôle des intuitions en philosophie morale. Cette semaine, le débat porte sur les « désirs des morts » et les « devoirs envers les morts ». Initialement, les étudiants sont unanimes pour dire que ce sont là des idées tout simplement absurdes : les morts n'ayant aucun désir, nous n'avons à leur endroit aucun devoir d'une quelconque sorte.

Cependant, à mesure que le débat avance et que les exemples se succèdent, certains étudiants (dont je fais partie) commencent à se montrer plus sympathiques face à ces idées qu'ils jugeaient à prime abord absurdes. Après une heure de discussions, seul un petit groupe de conséquentialistes convaincus demeurent totalement inflexibles sur la question. Le débat finit par se recentrer sur un problème plus précis : a-t-on une obligation (*prima facie*) de tenir une promesse faite à une personne un peu avant sa mort? Les conséquentialistes répondent par la négative. Une telle promesse a au plus une valeur *instrumentale*, disent-ils : si le fait de tenir cette promesse permet de donner bonne conscience à l'entourage du décédé et d'apporter un peu de bonheur en cette période difficile, alors oui, tenir cette promesse peut devenir moralement souhaitable. Cependant, le désir formulé par le défunt n'a aucune valeur *intrinsèque*. S'il n'y avait absolument aucune retombée du fait de tenir la promesse, il n'y aurait alors aucune obligation à la tenir. Et si les retombées étaient globalement *négatives*, tenir la promesse serait alors *condamnable*.

Sur cette question précise, je finis par « croiser le fer » avec un camarade qui, la semaine précédente, s'était montré farouchement opposé au recours aux intuitions en philosophie. Ce camarade avec qui j'ai une amicale rivalité philosophique – appelons-le David – défend l'idée que l'obligation qui nous lie à un individu suite à une promesse disparaît dès lors que cet individu s'éteint (ou dès lors que l'individu devient fou). Il va également plus loin en affirmant qu'il serait irrationnel d'accorder une véritable importance à une promesse de ce type, et qu'une personne qui sentirait une importante contrainte morale à tenir une telle promesse serait irrationnelle. Je soupçonne pour ma part que les convictions profondes de David sont contraires à la position qu'il

défend, et je décide alors, dans l'esprit de notre amicale rivalité, de rendre le débat un peu plus « personnel ». Je lui demande :

« Tu dis donc, David, que si *moi* je *te* demandais, un peu avant ma mort et sans que personne d'autre que toi et moi ne soit au courant, d'aller enterrer sous un arbre précis dans une région éloignée un objet qui m'était précieux, que tu n'aurais aucune obligation, même minimale, d'aller enterrer l'objet en question après mon décès même après m'avoir donné ta parole? ».

David répond par l'affirmative sans la moindre hésitation. Si tenir sa promesse n'implique aucun sacrifice réel, alors il peut être bien d'aller enterrer l'objet si on le souhaite. Cependant, tenir un tel engagement demeure entièrement surrogatoire, pour la simple et bonne raison que je serais mort et que son action n'aurait aucune conséquence sur moi ou sur quiconque d'autre.

Je décide alors de personnaliser encore davantage le débat :

« Laisse-moi te poser une question un peu différente alors. Si je te fais la même demande et que tu me donnes ta promesse, et que je meurs, je ne veux pas savoir si tu crois que tu aurais *le devoir* d'aller enterrer l'objet sous un arbre, je veux simplement savoir ceci : *est-ce que tu le ferais?* Si je te le demande avant de mourir, est-ce que tu irais enterrer mon objet sous l'arbre? Si oui, pourquoi? »

La réponse que donne alors David s'apparente davantage à une réflexion à voix haute, parsemée de silences : « Est-ce que j'irais enterrer ton objet? ... Je ne sais pas, peut-être que oui... Mais si c'était compliqué, peut-être que je... »

Un autre étudiant intervient alors: « Tu es sérieux? Tu ne ferais pas l'effort d'aller enterrer l'objet, même si tu sais que c'était important pour ton ami décédé? »

David semble alors « réaliser » qu'il ne souhaite pas réellement endosser une telle position, et sur un ton blagueur, mais dévoilant une subtile insécurité, il répond : « Bon Dieu... mais je suis un monstre! »

Cette réaction, ou « réalisation », de David illustre à merveille le type d'expériences intuitives qui constitue le centre d'intérêt de mon mémoire. En dépit de son opposition farouche à l'utilisation d'intuitions en philosophie, c'est bel et bien une intuition qui est venue ébranler sa certitude théorique quant à l'insignifiance morale du désir des morts. Pendant un bref instant, David a fait l'expérience de ce qu'il avait discrédité la semaine précédente, à savoir la contrainte qu'exercent les intuitions sur nos théories.

Les intuitions qui apparaissent spontanément en notre esprit nous *révèlent* en quelque sorte à nous-mêmes en dévoilant des convictions que nous ignorons souvent avoir. Le fait qu'elles se

présentent ainsi comme des *découvertes* explique sans doute pourquoi elles jouissent d'une force contraignante sur nos constructions théoriques – force que nombre de philosophes contemporains s'assurent d'utiliser à leur avantage. D'autres, cependant, critiquent farouchement la pertinence du renvoi aux intuitions en philosophie. Se fondant entre autres sur diverses études empiriques, ces « sceptiques de l'intuition » soutiennent que les intuitions ne sont pas fiables et que leur utilisation en philosophie doit être limitée autant que possible, voire complètement éliminée. À cela, plusieurs répondent qu'un discrédit général des intuitions mènerait ni plus ni moins à l'arrêt de mort de la philosophie en tant que discipline académique, puisque l'intuition constitue véritablement l'objet d'étude spécifique de cette discipline. La réhabilitation des intuitions n'est donc rien de moins qu'un enjeu existentiel pour certains philosophes.

Le problème de la fiabilité des intuitions est l'un des trois grands enjeux que j'aborderai dans ce mémoire. Le premier enjeu constitue le point de départ obligé de toute enquête philosophique et concerne *la nature* du phénomène central à mon analyse. Le chapitre 2 portera ainsi entièrement sur la question : que sont les intuitions? Une fois la nature des intuitions clarifiée, je serai en mesure de me tourner vers la deuxième question de mon mémoire : les intuitions sont-elles *fiables*? Je présenterai trois principaux types de raisons pouvant nous faire douter de la fiabilité des intuitions : des recherches scientifiques venant de différentes manières jeter un doute quant à la fiabilité des intuitions (3.1); le problème du désaccord intuitif tel qu'étudié par le mouvement de philosophie expérimentale (3.2); le problème de *l'origine* des intuitions (3.3). Je défendrai l'idée que ces trois types d'arguments ne viennent pas réellement discréditer l'utilisation des intuitions en philosophie mais démontrent plutôt le besoin d'utiliser les intuitions *avec prudence*.

Après avoir discuté des intuitions elles-mêmes aux chapitres 2 et 3, j'aborderai dans les deux chapitres suivants la question du *rôle* des intuitions en philosophie morale contemporaine. Je distinguerai trois grandes façons d'utiliser les intuitions en philosophie morale contemporaine, soit à travers les expériences de pensée, la méthode de l'équilibre réfléchi et la théorie de l'intuitionnisme éthique. Puisque la « méthodologie standard » en philosophie contemporaine consiste en une combinaison d'expériences de pensée et d'équilibre réfléchi, j'aborderai conjointement ces deux thèmes au chapitre 4. Je commencerai par donner quelques exemples classiques d'expériences de pensée, puis j'interrogerai la nature et la valeur épistémique de ce type d'exercice intellectuel (4.1). J'aborderai ensuite la méthode de l'équilibre réfléchi, en distinguant différentes versions de la méthode et en soulignant les principales critiques soulevées envers chacune d'elles (4.2). Le chapitre 5 sera ensuite entièrement consacré à la théorie de l'intuitionnisme éthique. Je distinguerai

deux grandes formes d'intuitionnisme, à savoir l'intuitionnisme *méta-éthique* (5.1) et l'intuitionnisme *normatif* (5.2), et je chercherai à voir dans quelle mesure on peut réconcilier l'une ou l'autre forme d'intuitionnisme avec la méthode de l'équilibre réfléchi (5.3).

Mon objectif dans ce mémoire est modeste; bien que je prenne position sur certaines questions centrales, par exemple sur la question de la fiabilité des intuitions ou sur la possibilité de réconcilier l'intuitionnisme éthique et la méthode de l'équilibre réfléchi, mon objectif est davantage d'effectuer un « état de la question » des différents débats contemporains portant sur les intuitions.

Je débute sans plus tarder avec le point de départ obligé de la présente enquête, à savoir la question : « Que sont les intuitions? ».

2. QUE SONT LES INTUITIONS?

Le concept d'intuition est l'un de ces concepts difficiles à cerner. Même en repoussant la tentation de rechercher des conditions nécessaires et suffisantes, en acceptant l'existence de frontières floues et en approchant le concept sous l'angle de la typicité plutôt que de l'essentialité des constituants, on se retrouve néanmoins confrontés à des désaccords majeurs quant à la nature du concept. Pour illustrer ce constat, notons que même un regroupement d'experts des intuitions, réunissant divers psychologues et philosophes, fut incapable de s'entendre sur une définition technique du terme pour l'ouvrage *Rethinking Intuition : The Psychology of Intuition and its Role in Philosophical Inquiry* (DePaul et Ramsey, 1998).

Il n'est donc guère surprenant que le concept donne lieu à un large éventail de caractérisations, dont voici un aperçu rapide; variant selon les auteurs, les intuitions peuvent être considérés comme des « apparences intellectuelles » (Huemer 2005), des « appréhensions directes de propositions évidentes en elles-mêmes » (Audi 2005, Gaut 2002), des « attitudes propositionnelles » (Bealer 1998), des « croyances non-inférentielles » (Sosa 1998, Hooker 2002), des « flashes de sentiments évaluatifs » (Haidt 2008), des « inclinations à croire » (Earlenbaugh et Molyneux, 2009), des « jugements spontanés » (Goldman et Pust 1998), des « sentiments de véridicalité ou de convenance » (Bedke 2008) et des « simili-croyances immédiates et non-réflexives » (Foley 1998).

Certaines de ces caractérisations portent plus spécifiquement sur les intuitions qualifiées de philosophiques (e.g. Audi 2005, Gaut 2002 et Bedke 2008) mais la plupart cherchent à englober toutes les formes d'intuitions. On peut également constater un certain désaccord quant à savoir si les intuitions sont principalement un phénomène rationnel (e.g. Huemer 2005) ou émotionnel (e.g. Haidt 2008). Je tenterai d'apporter quelques clarifications à ce sujet à la section 2.4.

Je ne cherche aucunement dans ce chapitre à trancher de façon définitive entre ces différentes caractérisations pour identifier la meilleure d'entre elles. Sans tomber dans une forme de relativisme définitionnel, il me paraît juste de soutenir que chacune de ces caractérisations comporte des forces et des faiblesses, et que chacune propose un découpage du phénomène qui peut s'avérer plus utile dans certains contextes que d'autres. Ce serait en mon sens une erreur d'évaluer ces différentes caractérisations en se référant à une quelconque « essence » du concept de laquelle on parviendrait plus ou moins à se rapprocher.¹ Ma démarche ne doit d'ailleurs pas être comprise

1 Il n'est bien sûr pas impossible que les intuitions constituent véritablement un "natural kind", en quel cas une

comme une « analyse conceptuelle » au sens philosophique traditionnel; même lorsque je tente de décortiquer le concept dans son usage populaire (section 2.1), mon but en est simplement un de clarification et est en cela beaucoup moins ambitieux que ce que cherchaient à démontrer certains philosophes par ce type d'analyse au siècle dernier.

Après avoir rapidement abordé l'usage populaire du concept à la section 2.1, j'effectuerai une analyse étymologique à la section 2.2. Je tenterai ensuite, à la section 2.3, de démystifier les intuitions en montrant que celles-ci ne sont aucunement le produit d'une « faculté étrange » comme l'ont soutenu certains philosophes comme John Mackie (Mackie 1977), et j'appuierai ce travail de démystification par quelques études portant sur l'intuition effectuées en sciences cognitives qui montrent que l'intuition n'a rien d'un phénomène mystérieux (2.4). Je profiterai également de ce détour vers les sciences cognitives pour apporter quelques distinctions importantes entre l'intuition, l'émotion et le raisonnement. Enfin, après avoir clarifié ce que sont et ne sont pas les intuitions, je conclurai le chapitre en me demandant si les intuitions peuvent ou non être considérées comme les objets d'études spécifiques à la philosophie (2.5).

2.1. Usage populaire

Dans la mentalité populaire, l'intuition est plus souvent qu'autrement assimilée à une faculté spéciale qui mènerait l'individu à faire l'expérience de différentes formes de prémonitions, ou qui le connecterait à une certaine « sagesse universelle transcendante ». L'intuition, ainsi comprise comme forme de sixième sens, séduit particulièrement les adeptes de l'ésotérisme, qui exhorteront souvent leur entourage à « suivre leur intuition » ou qui affirmeront à l'occasion « avoir l'intuition que quelque chose va se produire ». Une recherche rapide sur Amazon.com, le site d'achats en ligne, nous donne un aperçu de cet univers ésotérique généralement associé à l'intuition : *Navigating by Intuition : How to Follow the Signs* (Artemis, 2011); *The Intuitive Way : The Definitive Guide to Increasing Your Awareness* (Peirce, 2009); *Discover Your Psychic Type : Developing and Using Your Natural Intuition* (Dillard, 2008) sont parmi les premiers résultats de livres proposés suite à une recherche du mot-clé « intuition ». Or, ce que l'on appelle intuition en philosophie est bien sûr fort différent du phénomène auquel ces livres se réfèrent.

Il existe cependant un autre usage du concept populaire d'intuition qui s'éloigne de l'univers ésotérique et se rapproche davantage du concept philosophique. Cet autre usage populaire est

approche essentialiste paraîtra plus légitime. Je préfère cependant éviter le débat sur les "natural kinds", et laisserai ainsi la question en suspens.

présupposé dans des expressions courantes telles que « contre-intuitif » ou « intuitivement ». Dire qu'une idée ou proposition est « contre-intuitive » dans le langage populaire revient à dire qu'elle ne semble pas faire beaucoup de sens ou qu'elle va à l'encontre de notre façon courante de raisonner ou de penser. L'adverbe « intuitivement » soulève pour sa part une autre dimension qui se retrouve également dans le concept philosophique d'intuition; lorsqu'un individu répond à une question en affirmant : « intuitivement, je te dirais que je suis d'accord », « intuitivement » est synonyme de « à première vue » ou « ce qui paraît être le cas » ou « spontanément, avant une réflexion approfondie ».

Bref, s'il est vrai que le concept populaire d'intuition prend parfois une dimension ésotérique ou transcendante, il existe néanmoins un dénominateur commun entre les usages populaires et spécialisés : dans un cas comme dans l'autre, les intuitions sont un type d'épisode mental spontané, à première vue, non-réflexif et relevant de notre façon courante de penser. Ces quelques attributs permettent d'offrir une *caractérisation minimale* de l'intuition. C'est cette caractérisation minimale que j'aurai en tête lorsque je parlerai des intuitions dans mon mémoire. Des définitions plus pointues du concept peuvent bien sûr être données, mais seule cette caractérisation minimale me paraît englober les usages populaires et spécialisés de manière non-controversée.

Cette similarité entre les deux types d'usage soulève par ailleurs le problème de l'oeuf et de la poule : le concept d'intuition est-il une construction philosophique qui a fait son chemin dans le langage populaire, ou est-il plutôt un concept populaire repris par la communauté philosophique ? Une analyse étymologique du concept semble être la meilleure façon de répondre à cette question.

2.2. Étymologie

Le terme « intuition » provient du latin *intuitionem*², nom commun construit à partir du verbe *intueri*. *Intueri* est généralement traduit simplement par « voir », « observer » ou « contempler » (source : www.etymonline.com).

Bien qu'on ne retrouve aucune trace du terme « intuition » dans le Dictionnaire du Moyen Français (1330-1500) – www.atilf.fr/dmf –, le terme était vraisemblablement déjà utilisé à cette époque. Plusieurs éléments pointent en effet dans cette direction. D'abord, le site-web spécialisé en étymologie et généalogie de mots anglais, www.etymonline.com, soutient que le terme apparaît au XV^e siècle. Or, considérant que l'anglais a ses sources dans le « latin français », il serait surprenant

2 Le diminutif *intuitio* signifie "de l'intuition", i.e. quelque chose qui est un produit ou un objet de l'intuition.

que le mot « intuition » se retrouverait dans le vocabulaire anglais de l'époque mais non dans le vocabulaire français. Le même site-web soutient également que le verbe anglais « *to intuit* » apparaît en 1776, alors que la signification plus précise du terme comme le fait de « percevoir directement sans raisonnement » apparaît en 1840. Le site-web n'apporte cependant aucune information supplémentaire concernant cette première apparition en 1840, mais il est fort probable qu'il s'agisse de la première apparition d'une telle définition *dans un dictionnaire*. Une chose est certaine : ce n'est pas en 1840 que le terme est *compris* ou *utilisé* avec une telle signification pour la première fois, puisque plusieurs philosophes anglais modernes – e.g., Balguy, 1728; Clarke, 1728; Cudworth, 1731; Price, 1758 – comprenaient déjà le terme comme étant une forme de compréhension immédiate, non-réflexive.

Une enquête plus approfondie permet également de voir que le terme « intuition » était utilisé bien avant le XVI^e siècle. Les locutions latines *intuitio*, *intueri* et *intuitif* sont en effet fréquemment utilisées au Moyen-Âge par les plus importants penseurs, notamment Duns Scot (1265-1308) et Guillaume D'Occam (1285-1347).³ Le terme intuition fait alors référence à l'idée d'une « appréhension intellectuelle directe » (ou « cognition directe ») de « substances singulières » par opposition aux substances universelles. Comme le souligne Joel Biard dans *The Encyclopedia of Medieval Philosophy* (2011) :

« The theory of intuitive and abstractive cognition [...] is linked to the emergence of the idea that a singular can be grasped directly by the intellect itself. It was in the Franciscan, anti-Thomist, tradition that this theory was developed. In the *Correctorium fratris Thome* of William de la Mare, which was adopted in 1282 as an official position of the Franciscan order, Aquinas was criticized for having maintained that the intellect does not cognize singulars. Without a doubt, this is a simplification. But what is at stake is the possibility of a direct cognition of the singular. And it is from such a direct cognition that the idea of intuition or intuitive cognition is formed. In this sense, the theory is not truly formed until the last quarter of the thirteenth century, and finds its canonical expression in Duns Scotus before becoming a central component of theories of cognition from the fourteenth until the sixteenth century. » (Biard, 2011)

On retrouve donc dès le XIII^e siècle l'idée d'une intuition comme d'une « saisie intellectuelle directe », soit deux siècles avant la première apparition du terme dans le dictionnaire. Biard poursuit cependant son analyse et montre que le terme a des racines encore plus lointaines :

3 Merci à Christine Tappolet de m'avoir souligné la possibilité que le terme "intuition" ait ses origines dans la philosophie médiévale, et merci à David Piché de m'avoir confirmé cette information et de m'avoir référé à l'article de Joel Biard sur le sujet.

« The idea of intuition refers to vision. Since Augustine, the idea of vision has been applied not only to sensible vision, and not only to the inner vision of an image, but also to thought. Augustine generally used the term *visio* in this sense, but he introduces the verb *intueri* to describe the act of grasping the eternal reasons, and calls *intuitus* this immediate act of the mind (*De trinitate*, IX, vi, 9–11). The verb *intueri* and the adverb *intuitive* were first used in conjunction with singular cognition by Matthew of Aquasparta. We find them again used by Vital du Four, John Duns Scotus, and William of Ockham. The Augustinian origin is reinforced by the theory of vision, as it was received from the optics of Alhazen, where certain and direct vision by a straight ray, is named *intuitio*. » (*ibid.*).

C'est donc à St-Augustin (354-430) que l'on semble devoir l'idée d'une intuition comme d'une saisie intellectuelle immédiate, analogue à la perception sensorielle. Le terme ne remonte donc pas à la philosophie médiévale, mais bien à la philosophie antique tardive. Qui sait par ailleurs si une enquête plus approfondie ne montrerait pas les sources platoniciennes et/ou aristotéliennes du concept d'intuition utilisé par St-Augustin.

On peut également noter que l'usage augustinien du concept d'intuition s'inscrivait dans une conception théologique particulière à celui-ci. « Intuitionner » (*intueri*) renvoyait chez Augustin non pas à toutes formes d'intuitions intellectuelles, mais bien à l'intuition de « raisons éternelles », concept étroitement associé au monde divin. Les sources théologiques du concept expliqueraient-elles pourquoi certains philosophes assimilent encore aujourd'hui l'intuition à une « faculté étrange »? J'aborde cette idée dans la section suivante.

2.3. Faculté « étrange »?

L'intuition fait parfois l'objet de méfiance en philosophie contemporaine. La critique adressée aux défenseurs de l'intuitionnisme éthique, qui s'apparente davantage selon certains de ces défenseurs à une caricature, est que la faculté d'intuition serait une sorte de sixième sens qui permettrait de percevoir des propriétés spéciales ou étranges. Il suffit d'un pas pour dire que ces propriétés étranges ont une dimension *surnaturelle*. Or, en se basant sur les conclusions de la section précédente, on pourrait facilement conclure que ce soupçon de surnaturalisme s'explique par les sources théologiques du concept d'intuition.

Évidemment, le fait que le concept d'intuition ait à l'origine été associé à une vision religieuse du monde n'est pas suffisant pour expliquer la méfiance *contemporaine* face à ce concept. Il suffirait en effet que le concept ait été sécularisé à un moment ou un autre de l'histoire pour que la dimension théologique cesse de lui être ainsi associé, et on pourrait facilement penser que c'est ce

qui s'est produit avec les philosophes intuitionnistes suivant la Renaissance. Or, les premiers philosophes intuitionnistes modernes semblent en fait en grande partie avoir préservé la connexion entre la connaissance morale et le divin. Ralph Cudworth soutenait par exemple dans *A Treatise on Immutable and Eternal Morality* que les « idées » morales (en un sens platonicien), qui sont connues par intuition, sont éternelles et immuables et résident dans l'esprit divin (Cudworth, 1731). Des thèses similaires se retrouvent chez John Balguy (Balguy, 1728) et, dans une moindre mesure, chez Samuel Clarke (Clarke, 1728). Bref, si une sécularisation du concept a eu lieu, elle est arrivée plutôt tardivement dans la pensée philosophique, ce qui pourrait peut-être expliquer pourquoi le concept est encore perçu avec une certaine méfiance à notre époque.

Une autre source possible de cette méfiance philosophique pourrait se trouver dans l'assimilation populaire de l'intuition à l'univers ésotérique. Tel que décrit à la section 2.1, le concept d'intuition est souvent utilisé afin de décrire un mode de connaissance ésotérique qui n'est ni rationnel, ni empirique. Cette façon particulière de concevoir l'intuition colore la compréhension populaire du concept et pourrait ainsi influencer la perception philosophique.

Cependant, bien qu'il soit tout à fait possible que la source théologique du concept et/ou l'ésotérisme qui lui est associé dans la pensée populaire expliquent – ou renforcent – une certaine méfiance philosophique à son endroit, la raison première de cette méfiance est toutefois d'une autre nature. En effet, les sceptiques de l'intuition en philosophie contemporaine s'opposent non pas au possible caractère *surnaturel* de l'intuition, mais plutôt à son (apparent) caractère *non-naturel*. Dit autrement, même si l'on rejette tout référent divin ou transcendant ou ésotérique associé à l'intuition, l'intuition semble tout de même inévitablement supposer une forme de « non-naturalisme ». La critique en ce sens la plus célèbre vient de John Mackie, qui, dans *Ethics : Inventing Right and Wrong* (1977), soutient que toute théorie objectiviste en éthique est ultimement condamnée à postuler l'existence d'une « faculté étrange », qu'il nomme l'intuition :

« If we were aware [of objective values], it would have to be by some special faculty of moral perception or intuition, utterly different from our ways of knowing everything else. These points were recognised by Moore when he spoke of non-natural qualities, and by the intuitionists in their talk about a faculty of moral intuition. Intuitionism has long been out of favour, and it is indeed easy to point out its implausibilities. What is not so often stressed, but is more important, is that the central thesis of intuitionism is one to which any objectivist view of values is in the end committed: intuitionism merely makes unpalatably plain what other forms of objectivism wrap up » (Mackie, 1977, 38).

Bien que plusieurs aient reproché à Mackie de transformer l'intuitionniste en homme de

paille, sa critique semble tout de même refléter la pensée de plusieurs intuitionnistes. En effet, l'idée que l'objet de l'intuition est une propriété non-naturelle est défendue par un nombre important de philosophes intuitionnistes.

La raison principale pour laquelle des propriétés non-naturelles sont postulées par ces philosophes nous renvoie au problème soulevé par David Hume dans *A Treatise of Human Nature* (1739), que Hume nomme le « ought-is problem » (Hume, 1739, 335) et qu'Allan Gibbard décrit ainsi: « From nonethical premises there is no reasonable inference to any ethical conclusion; in other words, no *ought* can be deduced from an *is* » (Gibbard, 2002, 193).⁴ Dit autrement, on ne peut déduire une proposition morale entièrement à partir de propositions descriptives/non-morales. Pour parvenir à une conclusion morale, il faut inclure parmi les prémisses au moins une proposition morale non-réductible à une proposition non-morale. C'est ce type de propositions morales non-réductibles que les intuitionnistes considèrent être l'objet propre de l'intuition. Par exemple, la proposition morale selon laquelle « réparer un tort causé à autrui est (prima facie) *bien* » ne repose sur aucune autre proposition et ne peut au final qu'être acceptée *intuitivement*, c'est à dire qu'elle doit être conçue comme obtenant notre approbation spontanée. Or, pour un philosophe naturaliste, l'idée que certaines propositions soient non-réductibles à des propriétés naturelles est tout simplement inconcevable.

Un mémoire de maîtrise entier ne serait pas suffisant pour aborder convenablement tous les enjeux soulevés par le « ought-is problem », et je me contenterai donc des quelques remarques suivantes. D'abord, même en supposant que l'intuitionnisme soit effectivement incompatible avec le naturalisme – compris comme la thèse que toute propriété ne peut ultimement qu'être naturelle ou réductible à une propriété naturelle – on ne doit pas pour autant conclure à l'arrêt de mort de l'intuitionnisme. Plusieurs défenses sophistiquées du non-naturalisme ont été offertes au cours des dernières années, et il faudrait démontrer que ces différentes approches échouent toutes dans leur tentative de défendre le non-naturalisme pour en conclure qu'il s'agit là d'une thèse insoutenable qui, par ricochet, disqualifie la théorie intuitionniste.⁵

Ensuite, plusieurs philosophes soutiennent qu'intuitionnisme et naturalisme peuvent facilement être réconciliés. Roger Crisp affirme par exemple:

4 On peut également aborder le problème sous la forme proposée par G.E. Moore et connue sous le nom "open-ended question" (Moore, 1903).

5 Pour un exemple récent de défense originale du non-naturalisme, voir Huemer 2005. Huemer offre dans la première partie de son livre une défense "négative" ou "par défaut" du non-naturalisme, c'est à dire qu'il montre que toutes les autres possibilités théoriques en méta-éthique sont problématiques, et que seul le non-naturalisme permet d'éviter les problèmes auxquels les autres théories sont confrontées.

« [A] philosophical naturalist may accept intuition if she allows that we have the capacity to form non-inferential beliefs about what is ultimately right or reasonable to do. Whether the objects of those beliefs can be understood naturalistically or not is a metaphysical question, to be kept apart from epistemology or psychological questions concerning moral phenomenology » (Crisp, 2002, 59)

Allan Gibbard propose quant à lui une certaine réinterprétation de la pensée de G.E. Moore qui permette une réconciliation entre intuitionnisme et naturalisme:

« The *property* of being good is a natural property, though perhaps a complex one, and the like. The *concept* of being good, though, is special, and what's non-naturalistic is this concept. This is not how Moore would have put the matter, but it does have affinities with Moore's claims as he put them » (Gibbard, 2002, 212).

Enfin, les intuitionnistes paraissent unanimes pour décrier l'idée que l'intuition serait une faculté « étrange ». George Bealer affirme par exemple : « [By intuition,] we do not mean a magical power or inner voice or special glow or any other mysterious quality. When you have an intuition that A, it [simply] *seems* to you that A » (Bealer, 1998, 207) Allan Gibbard abonde dans le même sens: « Intuition figures in thinking what to do, and *oughts* in conclusions of what to do. Neither requires a mysterious metaphysic or mysterious psychic powers » (Gibbard, 2002, 228). La même idée est résumée par Philip Stratton-Lake de la façon suivante : « [Intuitionists] claim that certain moral propositions can be known by intuition, not because they think we have such a sixth sense, but because they think these propositions are self-evident. Intuition is not for them a way of perceiving a property, but is a way of grasping the truth of certain (*a priori*) propositions » (Stratton-Lake, 2002, 22).⁶

Afin de montrer en quoi l'intuition n'a pas à être considérée comme une faculté « étrange », on peut se tourner vers les sciences cognitives, qui peuvent aisément démystifier le phénomène en nous aidant à le « localiser » dans le cerveau humain.

2.4. Intuitions, émotions et sciences cognitives

6 Cette dernière citation de Stratton-Lake mérite une certaine nuance: s'il est vrai que la majorité des intuitionnistes ne conçoivent pas leur théorie comme appliquant une perception de propriétés, certains intuitionnistes ont cependant une telle idée. C'est pourquoi Henry Sidgwick distinguait entre trois formes d'intuitionnisme dont l'une était un intuitionniste "perceptuel" (Sidgwick 1874).

La philosophie n'est pas la seule discipline qui s'intéresse aux intuitions. La psychologie cognitive et les neurosciences s'intéressent également à ce phénomène, qu'elles abordent bien sûr avec des outils fort différents de ceux déployés en philosophie. Ces sciences cherchent surtout à préciser la nature des intuitions en clarifiant leur rôle dans la cognition humaine et en identifiant les mécanismes neuronaux et les zones du cerveau qui leur sont associés.

La théorie la plus répandue pour expliquer le phénomène intuitif dans ces disciplines est appelée « théorie à processus duaux » (*dual-process theory*). On retrouve différentes versions de cette théorie, et elle peut s'appliquer à différentes formes de cognitions humaines⁷, mais on y retrouve toujours la même idée centrale, à savoir qu'il existe deux types de processus cognitifs distincts que l'on nomme généralement Système 1 et Système 2. Ces deux systèmes fonctionnent en parallèle chez l'être humain et peuvent à l'occasion s'influencer.⁸ Le Système 1 est généralement décrit comme étant le fruit de processus inconscients, rapides et automatiques, alors que le Système 2 est au contraire associé à des processus conscients, lents et délibératifs. Le Système 1 est souvent nommé le système *intuitif* alors que le Système 2 est nommé le système *réflexif* ou *analytique*.

Plusieurs chercheurs considèrent que ces deux systèmes ont des fonctions de même que des origines évolutives bien distinctes, et certains croient pouvoir identifier une base neuronale spécifique à chacun d'eux. Matthew Lieberman (Lieberman, 2003) a par exemple identifié deux systèmes neuronaux précis qui sous-tendraient les Système 1 et 2. Le premier système neuronal, qu'il nomme le Système-X (X pour *refleXive*), serait composé de l'amygdale, des noyaux gris centraux et du cortex latéral temporel, alors que le second système neuronal, qu'il nomme le Système-C (C pour *refleCtive*), serait composé du cortex cingulaire antérieur, du cortex préfrontal et du lobe temporal médian.

Une idée souvent défendue par les partisans de la théorie à processus duaux est que le système intuitif serait beaucoup plus ancien que le système réflexif.

« The idea that System 1 cognition evolved earlier than System 2 is a recurring theme in dual-process theories (Evans & Over 1996, Epstein & Pacini 1999, Reber 1993, Stanovich 1999). System 2 is thought to be associated with language, reflective consciousness, and higher-order control and with the capacity to think hypothetically about future and counterfactual possibilities. Authors have often asserted that such

7 Jonathan Evans (Evans, 2008), par exemple, distingue entre des théories à processus duaux qui s'appliquent aux jugements et prises de décision (e.g., Kahneman et Frederik, 2002), au raisonnement (e.g., Stanovich, 2004) et aux cognitions sociales (e.g., Chen et Chaiken, 1999)

8 Certains soutiennent que le Système 1 jouit d'une autonomie importante et n'est pas facilement influencable par le Système 2. Ronald de Sousa affirme par exemple: « The efficiency of our First Track functions would be at risk if they were too easily influenced by external information » (DeSousa, 2009, 113).

characteristics of thought are distinctively or uniquely human, while at the same time arguing that System 1 cognition, shared with other animals, continue to control much of our behavior » (Evans, 2008, 259).

Bien qu'il soit beaucoup plus ancien ou *primitif*, le système intuitif serait, selon plusieurs chercheurs, beaucoup plus fréquemment utilisé par les individus que le système réflexif, ce qui amène Ronald de Sousa à souligner : « Notice [...] that while we, like Descartes, continue to regard explicit thinking as essentially human, most of our life skills are governed by the intuitive track » (De Sousa, 2009, 101). Cette idée est renforcée par les travaux de Jonathan Haidt, qui montrent que la plupart des individus, lorsque confrontés à des problèmes moraux, tendent à prendre position sur la base de jugements spontanés (i.e., des intuitions) plutôt que sur la base de raisonnements. C'est seulement après avoir *tranché intuitivement* qu'ils se mettent à la recherche de raisons justifiant leur position, et dans plusieurs cas, les individus s'accrochent à leur intuition même lorsque confrontés à des raisons leur indiquant qu'ils sont vraisemblablement dans l'erreur – ce que Haidt appelle le phénomène du « moral dumbfounding ». ⁹

Certains chercheurs mettent cependant en doute l'idée que le Système 1 serait un système unique et distinct, et soutiennent plutôt qu'il existerait un grand nombre de processus mentaux rapides, automatiques et inconscients qui ne constitueraient pas à proprement parler un « système ». Jonathan Evans conclut d'ailleurs son survol de la littérature sur les théories à processus duaux sur une mise en garde :

« In short, my conclusion is that although dual-process theories enjoy good empirical support in a number of fields in psychology, the superficially attractive notion that they are all related to the same underlying two systems of cognition is probably mistaken, at least in the way that Systems 1 and 2 are being defined in the current literatures. For example, it is almost certainly wrong to think of System 1 as one system, all of which is old and shared with other animals. Equally, it is probably a mistake to think of System 2 as the conscious mind, all of whose processes are slow and sequential. » (Evans, 2008, 271).

Un exemple parfait de processus cognitif qui possède les caractéristiques du Système 1 (rapide, automatique et inconscient) sans pour autant être une intuition est celui des *émotions*. Evans souligne qu'étonnamment, la plupart des théoriciens des processus duaux ont ignoré les émotions comme exemple de phénomènes cognitifs typiques du Système 1 qui ne sont pas à proprement parler des intuitions (*ibid.*, 256-7). Il faut plutôt se tourner vers la philosophie des

⁹ Voir sur ce point les recherches de Haidt qui montrent les réactions d'individus lorsque questionnés sur un cas fictif d'inceste consensuel entre un frère et une soeur (Haidt 2001).

émotions pour que le rapprochement entre émotions et Système 1 soit établi clairement. Ronald de Sousa affirme par exemple : « Many of the features we associate with the modus operandi of emotions [...] can be classified as belonging to the First or Intuitive Track in the two-track model » (De Sousa, 2009, 101).

Il faut dire que la distinction entre intuitions et émotions n'est pas toujours évidente, ce qui peut mener à plusieurs confusions lorsque vient le temps d'évaluer le rôle et la pertinence des intuitions en philosophie. Par exemple, Peter Singer, dans sa critique du renvoi aux intuitions en philosophie morale, utilise souvent les termes intuitions et émotions de manière interchangeable, comme le suggère le passage suivant :

« Haidt's behavioral research and Greene's brain imaging studies suggest the possibility of distinguishing between our immediate emotionally based responses, and our more reasoned conclusions. In everyday life, as Haidt points out, our reasoning is likely to be nothing more than a rationalization for our intuitive responses – as Haidt puts it, the emotional dog is wagging the rational tail. But Greene's research suggests that in some people, reasoning can overcome an initial intuitive response. » (Singer, 2005, 350).

Dans ce passage, Singer utilise les expressions « immediate emotionally based responses » et « initial intuitive responses » comme des synonymes. Une telle indifférenciation s'explique sans doute en grande partie par le fait que les intuitions et les émotions partagent toutes deux les caractéristiques du Système 1, et que l'intuition est un phénomène qui, comme l'émotion, comporte une part importante de « ressenti ». D'ailleurs, si l'on retourne aux nombreuses définitions du terme intuition que j'ai mentionnées en introduction de chapitre, on remarque que certaines définitions – par exemple celle de Haidt (« flashes de sentiments évaluatifs ») et celle de Bedke (« sentiments de véridicalité ou de convenance ») – mettent l'accent sur cette dimension ressentie.

Le problème cependant est que les phénomènes intuitifs et émotifs, bien que souvent associés, sont clairement distincts et ne doivent pas être confondus. Une intuition peut être dépourvue de toute émotion et une émotion peut être dépourvue de toute intuition – bien que l'émotion semble toujours avoir un « contenu » ou « objet »¹⁰. Une analyse phénoménologique semble suffire à démontrer cette distinction. Pour prendre des exemples qui illustrent de manière flagrante la différence entre les deux types de phénomènes, on peut comparer l'émotion de peur avec l'intuition mathématique ou logique. Lorsqu'on finit par comprendre une longue formule

10 Certains soutiennent cependant que les émotions contiennent toujours une composante intuitive – voir par exemple le passage de Jonathan Haidt au paragraphe suivant. J'interprète ici Haidt comme voulant dire que les émotions ont toujours un objet ou un « contenu propositionnel ».

mathématique ou qu'on parvient à saisir la validité d'une inférence logique, le « dé clic » qui caractérise notre compréhension est véritablement une intuition. C'est également ce type d'intuitions qui semble suscité dans la plupart des expériences de pensée développées en philosophie : « It feels appropriate to deny that XYZ is water [just like] the formula $\neg(p \wedge q)$ iff $(\neg p) \vee (\neg q)$ feels right after some reflection » (Bedke, 2008, 260). Ce type d'expérience intuitive semble n'avoir pratiquement rien en commun avec l'expérience ressentie lorsqu'un ours surgit devant notre tente lors d'une activité de camping. On a réellement affaire ici à deux phénomènes différents.

Dans le même ordre d'idées, Jonathan Haidt, dont la définition de l'intuition semble servir davantage à entretenir qu'à défaire la confusion entre intuition et émotion, s'efforce parfois de distinguer les deux phénomènes. Par exemple, après qu'on lui eut adressé une critique (Jacobson, 2008) selon laquelle sa théorie réduirait le raisonnement moral au fait de ressentir des émotions, Haidt répondit que sa théorie (SIM)¹¹ portait sur les intuitions plutôt que sur les émotions :

« The SIM is about intuition more than emotion. [...] Intuitions are often accompanied by affect, and all emotions include an intuitive component (the automatic appraisal which launches the emotion; see Ortony, Clore & Collins, 1988). [...] Thus, when Jacobson says that 'to put it most crudely, social intuitionism holds that we arrive at moral judgments by feeling rather than thinking', we protest that this portrait is indeed too crude. Social intuitionism holds that we arrive at moral judgments by *intuition*, which is a kind of thinking that is not reasoning, and in which emotion often plays a role » (Haidt et Bjorklund, 2008, 250-1).

Cette citation de Haidt et Bjorklund est particulièrement importante puisqu'elle permet de clarifier la relation complexe entre intuition, émotion et raisonnement. Comme le souligne Haidt, l'intuition n'est pas un phénomène émotionnel mais plutôt un phénomène rationnel, ce à quoi deux nuances importantes doivent être ajoutées : (1) l'intuition est un phénomène rationnel qui, contrairement au raisonnement, est spontané et apparaît tel un « flash évaluatif » en notre conscience, et (2) bien qu'étant un phénomène distinct de l'émotion, l'intuition est très souvent accompagnée d'émotion, ce qui explique en grande partie pourquoi les deux phénomènes sont si souvent indifférenciés. Même les intuitions qui semblent les plus « froides », par exemple les intuitions mathématiques ou logiques, peuvent très bien susciter de fortes émotions dans certains contextes. Par exemple, le mathématicien ou le logicien qui, après avoir passé des jours à analyser une formule, arrive finalement à la solution qu'il recherchait ressentira certainement une grande

11 SIM = Social Intuitionist Model.

dose de soulagement, d'excitation voir d'euphorie lorsque la solution tant recherchée émergera finalement en sa conscience sous la forme d'une intuition. À l'inverse, si ce même individu finit par découvrir, après des heures ou des jours d'analyse, que la piste qu'il suivait menait ultimement à un cul-de-sac, l'intuition dont il fera l'expérience sera fort probablement accompagnée d'une bonne dose d'affect négatif.

Bref, aucune intuition ne semble *a priori* imperméable au ressenti émotif. On pourrait aller plus loin et soutenir, comme le fait Aaron Ben-Ze'ev, que tout état mental humain comporte à la fois une dose d'affect et une dose de contenu propositionnel – ou « intentionnel » :

« Typical mental states in human beings consist of both intentional and feeling dimensions. Thus, seeing something often evokes pleasant or unpleasant feelings relating to the content of what we see. The relationships between the two dimensions vary in type and degree for different mental states. Whereas in emotions both dimensions are central, in most mental states only one of these is dominant. » (Ben-Ze'ev, 2009, 49).

Pour résumer, on peut dire que les intuitions sont un type de jugements qui émergent spontanément en la conscience et qui sont souvent accompagnées d'émotions – bien qu'étant un phénomène distinct de l'émotion. L'intuition est également distincte du raisonnement, compris comme un dialogue interne consistant en l'analyse de divers arguments ou raisons pouvant être soulevés pour ou contre une position. Mais bien que distincte, l'intuition joue un rôle essentiel dans le processus de raisonnement. En effet, c'est très souvent une intuition qui *initie* un raisonnement, et diverses intuitions surgissent tout au long du processus, venant parfois rediriger le dialogue interne dans une direction ou une autre. Darcia Narvaez décrit la relation dynamique entre raisonnement et intuitions lors de délibérations morales comme étant une sorte de « danse mentale » ou encore un « jeu de chaise musicale moral » (Narvaez, 2008, 235). Sur le plan neuronal, on peut concevoir cette danse mentale comme étant le reflet de l'interaction entre le Système 1 et le Système 2.

2.5. Conclusion : domaine spécifique et crise existentielle

J'ai tenté de montrer dans ce chapitre ce que sont et surtout ce que ne sont pas les intuitions : les intuitions utilisées en philosophie n'ont pas un caractère ésotérique ou transcendant et ne relèvent pas d'une faculté étrange. Elles sont tout simplement des jugements spontanés produits par ce que l'on nomme en sciences cognitives le « Système 1 ». Elles sont distinctes à la fois des

émotions et du raisonnement, bien qu'elles leurs soient étroitement reliées.

Certains philosophes soutiennent que les intuitions ainsi comprises constituent le domaine spécifique de la philosophie. Par là, ces philosophes entendent non pas que *seule* la philosophie s'intéresse aux intuitions, mais plutôt que les intuitions sont ultimement le seul véritable objet d'analyse de la philosophie. Garry Gutting défend cette idée lorsqu'il affirme : « In contemporary analytic discussions [...] "intuition" has become the name for whatever it is that might provide philosophy with a distinctive method and hence preserve it as a separate (in principle) intellectual domain » (Gutting, 1998, 7). Michael DePaul et William Ramsey rappellent quant à eux que l'intuition a toujours été l'objet propre de la philosophie : « [F]rom its beginning in Greek philosophy right through to the present, intuitions have always played an extremely important role in Western philosophy. [...] And [philosophers] continue to use them in the absence of any well articulated, generally accepted account of intuitive judgment » (DePaul et Ramsey, 1998, viii).

Cette dernière phrase de DePaul et Ramsey laisse deviner l'existence d'un certain malaise concernant l'utilisation des intuitions en philosophie. Les deux auteurs clarifient ce malaise quelques pages plus loin en soulignant que les intuitions sont devenues l'objet de suspicion de la part de bon nombre de philosophes :

« There is a growing consensus among philosophers that there is a serious and fundamental problem here that needs to be addressed. In fact, we do not think it is an overstatement to say that Western analytic philosophy is, in many respects, undergoing a crisis where there is considerable urgency and anxiety regarding the status of intuitive analysis ». (*ibid.*, x).

Le malaise décrit par les deux auteurs cache donc une véritable crise que traverse la philosophie – ou plutôt, que *traversait* la philosophie lors de la parution de ces lignes il y a 14 ans. Certains allaient même jusqu'à suggérer que la philosophie devait entièrement se débarrasser des intuitions si elle souhaitait préserver une quelconque pertinence : « Philosophical intuition [...] is epistemologically useless. [...] I do not know if something [...] will replace intuition in every branch of philosophy, but *something* had better replace it. It cannot support any conclusion worth drawing. » (Cummins, 1998, 118). L'idée essentielle de ces adversaires de l'intuition était que seule la méthode expérimentale avait une légitimité intrinsèque comme démarche menant à la connaissance : « It is illegitimate to put an analysis of our intuitive concepts to work as a block to the movement of science in one direction or another, as though it revealed some independent truth that science was powerless to overthrow » (Gopnick et Schwitzgebel, 1998, 86). Certains étaient

cependant prêts à reconnaître une pertinence minimale à la démarche intuitionniste, mais en ajoutant que les intuitions ne jouaient un rôle qu'au début de l'enquête scientifique : « Intuition must be taken seriously in the absence of substantial theoretical understanding, but once such theoretical understanding begins to take shape, a priori intuitive judgments carry little weight unless they have been endorsed by the progress of theory » (Kornblith, 1998, 135).

Le problème, évidemment, est que si l'utilisation des intuitions à des fins de construction théorique est effectivement illégitime ou inutile, c'est la philosophie toute entière qui s'en retrouve menacée, pour la simple raison que ce type de démarche semble constituer l'objet spécifique de la philosophie, un constat qui amène Robert Cummins à dire : « To many, this will appear as an unemployment opportunity » (Cummins, 1998, 118). Avant de se réorienter vers la pâtisserie ou la mécanique cependant, il est bon de s'intéresser davantage aux raisons qui ont mené certains philosophes à prôner un virage radical de leur discipline. Quel est le problème avec les intuitions? Pourquoi devrait-on les délaissier même au péril de la philosophie? Tel est le type d'interrogation qui guidera le prochain chapitre, qui porte sur le problème de la fiabilité des intuitions.

3. LES INTUITIONS SONT-ELLES FIABLES?

Un argument récurrent utilisé par les adversaires des intuitions est que celles-ci ne sont pas fiables. En effet, que la critique porte sur les expériences de pensée, la méthode de l'équilibre réfléchi ou encore l'intuitionnisme éthique, on souligne toujours cet argument selon lequel le manque de fiabilité des intuitions devrait suffire pour nous décourager de les utiliser en philosophie. Mais d'où provient cette idée que les intuitions ne sont pas fiables? Dans ce chapitre, je souhaite aborder ce qui me semble être les trois principales sources de méfiance à l'endroit des intuitions. D'abord, de nombreuses recherches scientifiques sont à l'origine de ce discrédit. J'aborderai certaines de ces recherches à la section 3.1. Ensuite, l'existence de désaccords intuitifs importants sur pratiquement tous les sujets entre individus est une deuxième source importante de méfiance envers les intuitions. J'aborderai ce phénomène à la section 3.2. Enfin, le problème de l'*origine* des intuitions semble être un autre objet d'inquiétude important; certains soulignent que nos intuitions ne sont que le produit d'une intériorisation de normes sociales arbitraires, alors que d'autres soutiennent qu'elles sont des vestiges de notre passé évolutif qui n'ont plus aucune pertinence aujourd'hui. Je discuterai de ce problème originaire à la section 3.3.

Il existe certainement de nombreuses autres critiques pouvant être soulevées contre la fiabilité des intuitions, mais les trois types de critiques que j'aborderai ici me semblent clairement être les plus importantes. J'espère démontrer par mon analyse que les critiques soulevées à l'endroit des intuitions, bien que montrant l'importance d'une certaine *prudence* dans notre utilisation des intuitions en philosophie, ne portent aucunement un coup fatal à cette pratique.

3.1. Recherches scientifiques

Les recherches scientifiques portant directement ou indirectement sur les intuitions sont légion. Je me concentrerai ici sur trois grands types de recherches qui ont contribué à soulever des doutes importants quant au recours aux intuitions en philosophie : les recherches en économie comportementale portant sur les évaluations irrationnelles des agents (3.1.1), les recherches en psychologie cognitive portant sur les catégorisations et la nature des concepts (3.1.2) et les recherches en philosophie expérimentale portant sur la non-universalité des intuitions philosophiques. Ce dernier type de recherches sera abordé à la section 3.2, en lien avec le problème du désaccord. La section 3.3 portera en grande partie sur les intuitions morales, ce qui m'amènera à

aborder d'autres domaines de recherche, tels la psychologie développementale et la psychologie morale. Dans tous les cas, j'offrirai un survol bien modeste n'ayant aucune prétention d'exhaustivité.

3.1.1. Économie comportementale: la faiblesse évaluative

Les recherches en économie comportementale sont sans doute les recherches les plus souvent citées par les adversaires des intuitions. L'économie comportementale est une discipline à mi-chemin entre la psychologie et la science économique dont les nombreuses recherches ont principalement servi à discréditer le modèle de l'« agent économique rationnel ». Ces recherches ne concernent donc pas directement l'usage des intuitions en philosophie, mais elles sont tout de même importantes dans la mesure où elles montrent comment les évaluations humaines, aussi bien délibératives (raisonnement, Système 2) que spontanées (intuitions, Système 1), sont souvent irrationnelles et peu fiables.

La théorie la plus connue dans cette discipline est la « théorie des perspectives » (*Prospect theory*), développée initialement en 1979 par Daniel Kahneman et Amos Tversky (Kahneman et Tversky, 1979).¹² La théorie des perspectives démontre essentiellement que les comportements et évaluations économiques des individus diffèrent grandement de ceux prédits par la théorie classique, i.e., la théorie de l'utilité espérée. Les recherches de Kahneman et Tversky montrent que les évaluations des agents économiques reflètent d'importants biais cognitifs, notamment une aversion aux risques et l'utilisation d'un point de référence dans la comparaison de différentes options. De tels biais cognitifs amènent parfois les agents à effectuer des choix irrationnels.¹³

S'inspirant de ces travaux, de nombreux chercheurs ont voulu approfondir le phénomène d'irrationalité des agents économiques, cherchant à révéler comment, outre les biais cognitifs décrits par Kahneman et Tversky, de nombreux facteurs d'apparence anodine influencent grandement les

12 Kahneman a d'ailleurs remporté le prix Nobel d'économie pour ses travaux sur ce sujet.

13 Il existe bien sûr de nombreuses conceptions différentes de l'irrationalité (et de la rationalité). En acceptant une conception particulière de l'irrationalité, on pourrait conclure que les comportements et évaluations décrits par Kahneman, Tversky et autres chercheurs ne sont pas véritablement irrationnels. Je préfère éviter ici de comparer différentes conceptions de l'irrationalité, et je me contenterai de souligner que le concept d'irrationalité utilisé par les chercheurs en économie comportementale comporte généralement, en plus du critère d'incohérence des préférences, les deux (ou l'un ou l'autre des deux) éléments suivants : d'une part, l'irrationalité implique le fait d'agir à l'encontre de son meilleur intérêt (compris en un sens économique classique), et d'autre part, elle implique le fait d'accorder un poids important dans le processus décisionnel à des éléments pourtant anodins ou futiles. Chacun de ces deux éléments soulève des questions importantes, notamment en ce qui a trait à la possibilité de déterminer objectivement le « meilleur intérêt » d'un individu ou en ce qui a trait aux critères permettant de déterminer ce qui rend certaines considérations « anodines » ou « futiles ». Je ne discuterai pas davantage de ces questions, me contentant de souligner qu'il s'agit là de critiques possibles à l'endroit des études de Kahneman et autres chercheurs en économie comportementale.

évaluations des individus. Par exemple, de légères modifications dans le contexte de l'agent ou dans la manière de présenter les différentes options peuvent mener à de véritables « renversements des préférences » (« preference reversals »). Un exemple classique d'études démontrant un renversement des préférences ressemble à ceci :

« Subjects are first asked to choose between two gambles with nearly the same expected value. One gamble, called the *H* bet (for high chance of winning) has a high chance of winning a relatively small prize (say 8/9 chance to win \$4) while the other gamble, the *L* bet, offers a lower chance to win a larger prize (say, a 1/9 chance to win \$40). Most subjects choose the *H* bet. Subjects are then asked to price each of the gambles. Specifically, they are asked to state the lowest price at which they would be willing to sell each gamble if they owned it. Surprisingly, most subjects put a higher price on the *L* bet. [...] This pattern is called a *preference reversal*. Sarah Lichtenstein and Paul Slovic (1971, 1973) first demonstrated such reversals in a series of studies, one of which was conducted for real money with gamblers on the floor of the Four Queens Casino in Las Vegas » (Tversky and Thaler, 1990, 202-3)

Ce type de renversements de préférences des agents a été observé et reproduit dans de nombreuses études (voir par exemple Tversky, Slovic, et Kahneman, 1990; Lichtenstein et Slovic, 1973; Slovic et Lichtenstein, 1983; Kachelmeier et Shehata, 1992). Un sceptique pourrait répondre que ce type de recherches démontre simplement que les individus ne sont pas doués pour évaluer différentes *probabilités*, et que cette prétendue irrationalité ne doit pas être généralisée à tous les domaines. Cependant, force est de constater que de telles évaluations irrationnelles dépassent largement le domaine des probabilités. Pour prendre un exemple parmi bien d'autres, le phénomène d'irrationalité s'observe également dans les évaluations de compensation devant être offertes à des victimes. Par exemple, un homme qui se fait tirer par balle par un cambrioleur en se rendant à un dépanneur où il n'a pas l'habitude d'aller est jugé devoir recevoir en moyenne 100,000\$ de plus en compensation qu'un homme qui se fait tirer par balle en se rendant au dépanneur où il a l'habitude d'aller (Miller et McFarland, 1986). On observe des évaluations de compensation similaires pour un soldat qui se fait tuer une journée avant la fin de la guerre plutôt que six mois avant la fin de la guerre (Shafir, 1998, 64), ou encore pour un individu qui meurt dans un accident d'avion en ayant changé d'avion à la dernière minute plutôt qu'en ayant pris l'avion qu'il devait prendre (*ibid.*).

Une autre réponse que l'on pourrait être tenté d'offrir est que ces études montrent simplement que les individus ne sont pas doués pour assigner des *valeurs monétaires*, ce qui semble être le point commun des études soulevées jusqu'à présent. Mais le phénomène d'irrationalité s'étend, encore une fois, bien au-delà de ce domaine. Plusieurs études révèlent que même en

l'absence de valeur monétaire, les individus effectuent souvent des choix incohérents ou se laissent influencés par des éléments banals. Par exemple, des individus à qui l'on demande de déterminer qui, entre deux parents aux personnalités et circonstances différentes, mérite la garde d'enfants donnent des réponses complètement différentes si on change un simple mot dans la présentation du problème : si on leur demande à qui on doit *accorder* la garde, les individus donnent une réponse différente de celle donnée si on demande à qui on doit *refuser* la garde (Shafir, 1993). Des exemples de ce type se retrouvent dans plusieurs études qui montrent l'impact important que peut avoir un simple changement de « cadrage » (framing) sur les évaluations des individus. De nombreuses études ont par exemple montré qu'une simple modification dans l'*ordre de présentation* des options pouvait parfois mener à des évaluations différentes de la part des individus.

Une réplique légitime que l'on peut cependant offrir à toutes ces études est que celles-ci ne concernent pas spécifiquement les intuitions, mais s'adressent plus largement aux capacités évaluatives des individus. C'est un constat que soulève entre autre Richard Foley : « Many of the objections against intuitions in philosophy and other kinds of theoretical reasoning are not objections against intuitions *per se* but rather objections that, when valid, apply equally to other sorts of intellectual judgments » (Foley, 1998, 245). Comment doit-on alors interpréter la portée de ces études? Dans la mesure où celles-ci montrent les failles non seulement de nos évaluations spontanées (intuitions) mais également de nos évaluations réfléchies (raisonnements), doit-on conclure que l'on doit se méfier de toute forme d'évaluations? C'est là une conclusion qui paraît beaucoup trop forte. Après tout, les raisonnements et intuitions s'avèrent très efficaces dans plusieurs domaines, et conclure à partir de ces recherches qu'on ne saurait se fier à aucune forme de raisonnement humain revient à jeter le bébé et toute la pouponnière avec l'eau du bain!

Une conclusion plus raisonnable à tirer de ces études est que celles-ci doivent être comprises comme lançant un appel à la *prudence* dans l'utilisation des intuitions, plutôt qu'un discrédit total de notre faculté d'intuition. Or, comme le souligne Ernest Sosa, un tel appel à la prudence n'a rien de surprenant en soi :

« Actually, we have long known of the fallibility of apparent rational intuition, even in the best conditions of alertness, normality, and reflection time. [...] We humans are indeed sometimes systematically in error when we apparently intuit while normal, alert, and amply reflective. Our reasoning *competence* seems flawed; not just our performance. What follows from this? Does it seriously undermine our continued use of intuition? [...] Any such conclusion would be premature. » (Sosa, 1998, 261).

La nécessité d'aborder les intuitions avec prudence est en réalité un impératif qui a toujours été défendu par les défenseurs des intuitions eux-mêmes. Les intuitions nous jouent parfois des tours, et il faut prendre garde de ne pas systématiquement endosser leur contenu – mais cela ne doit pas mener à un rejet systématique des intuitions. C'est là une conclusion à laquelle arrive Sosa dans un autre article où il effectue un survol des études en philosophie expérimentale :

« The upshot is that we have to be *careful* in how we use intuition, not that intuition is useless. It is of course helpful to be shown how intuition can go astray in unfavorable conditions, just as perception can go similarly astray. But the important question is untouched: Can intuition enjoy relative to philosophy an evidential status analogous to that enjoyed by perception relative to empirical science? » (Sosa, 2007, 105)

Sosa soulève ici une analogie entre l'intuition et la faculté de perception, analogie qui est fréquemment utilisée afin de défendre le recours aux intuitions. De la même manière que la perception – ou l'introspection, la mémoire ou autre faculté – nous joue à l'occasion des tours sans que cela ne nous conduisent à endosser un scepticisme radical à la Descartes, il devrait également être possible de se fier à nos intuitions même si celles-ci se révèlent peu fiables à l'occasion. C'est là un argument repris par de nombreux philosophes, notamment David Sosa (Sosa, 1998, 638) et S. Matthew Liao (Liao, 2008, 255). Le point qui divise cependant les défenseurs de l'intuition est celui de la *fréquence* à laquelle l'intuition nous joue des tours. Alors que David Sosa affirme que l'intuition n'est fautive que dans quelques circonstances particulières bien connues, Liao et d'autres soutiennent que l'erreur intuitive est plus fréquente. Or, déterminer la fréquence à laquelle nos intuitions sont erronées peut s'avérer important, dans la mesure où si cette fréquence est très élevée, l'analogie avec nos autres facultés (perception, mémoire, etc.) serait alors moins adéquate.

Il est tout de même possible de légitimer l'usage des intuitions même en étant incapable de déterminer la fréquence de l'erreur intuitive. Cela est en effet possible si l'on possède une méthode permettant de séparer le bon grain de l'ivraie dans nos intuitions. Ainsi, même si nos réactions intuitives immédiates étaient largement fautives, on pourrait légitimement continuer à utiliser nos intuitions à condition que l'on soit en mesure de les *filtrer* afin de ne préserver que les plus fiables d'entre elles. Or, une telle méthode existe effectivement; il s'agit de la méthode de l'équilibre réfléchi, que j'aborderai au chapitre 4.

Avant d'aborder cette méthode, il me reste encore à considérer d'autres sources possibles du manque de fiabilité de nos intuitions. Je poursuis à la section suivante avec un autre type de

recherches scientifiques qui tente de démontrer par un chemin complètement différent le manque de fiabilité des intuitions.

3.1.2. Sciences cognitives : catégorisations et nature des concepts

Les études en sciences cognitives portant sur la catégorisation et la nature des concepts tendent presque toutes à la conclusion suivante : rechercher des conditions nécessaires et suffisantes derrière nos concepts normatifs est une mission vaine, vouée à l'échec dès le départ. On aurait sans doute plus de chance de trouver le pot d'or au bout de l'arc-en-ciel. Ces études, initiées principalement par Eleanor Rosch dans les années 1970 (Rosch et Mervis 1975, Rosch 1978) renforcent l'idée qu'il n'existe aucune « essence » derrière nos concepts, aucun ensemble de conditions nécessaires et suffisantes qui nous permettraient de déterminer véritablement ce qui constitue *la justice* ou *la connaissance* ou *le beau* ou *le bien*. Certes, pour plusieurs propriétés naturelles, il est possible de trouver un petit nombre d'éléments qui sont véritablement essentiels ou constitutifs de cette propriété : l'eau est essentiellement un composé de H₂O, etc. Mais c'est une erreur de croire que l'on puisse identifier pour nos concepts normatifs des propriétés essentielles analogues aux éléments chimiques, pour la simple raison que nos concepts ne se construisent pas sur la base de caractéristiques essentielles.

Le modèle théorique qui a principalement servi à discréditer le modèle classique (ou essentialiste) est le modèle *prototypique* des concepts. Selon ce modèle, nos concepts ont des frontières floues et regroupent sous une même étiquette des objets qui ont des « ressemblances de famille » plutôt qu'une essence conjointe – idée qui remonte bien sûr à Wittgenstein¹⁴. Un exemple classique de l'idée que nos concepts se construisent par prototypes plutôt que par conditions nécessaires et suffisantes est celui du concept « oiseau ». Des études (Rosch 1978, Smith et Medin 1981) ont montré que les individus ont beaucoup plus de facilité à reconnaître que les merles sont des oiseaux qu'à reconnaître que les autruches sont des oiseaux. Un tel constat n'a en soi rien pour causer une révolution, mais ce sont des études de ce genre qui ont permis de renforcer la thèse principale de la théorie des prototypes, qui soutient essentiellement que ce qui permet ou non de situer un objet sous un concept est sa proximité avec un prototype. Cette idée de proximité avec un prototype est étroitement associée à la thèse d'une « structure probabiliste » des concepts, que Laurence et Margolis (2011) décrivent ainsi : « Concepts have a probabilistic structure in that

14 Voir en particulier l'investigation philosophique 66 (Wittgenstein 1953).

something falls under C just in case it satisfies a sufficient number of properties encoded by C's constituents » (Laurence et Margolis, 2011, §2.2). Plus un objet est similaire à un prototype, plus il est probable qu'il soit compris sous un concept rattaché à ce prototype. Une telle structure probabiliste est contraire à l'idée d'une structure définitionnelle associée au modèle classique des concepts.¹⁵

À la lecture de ces études, le défenseur des intuitions aura sans doute le désir de répondre par un haussement d'épaule. Il est vrai qu'à première vue, de telles études ne semblent pas avoir de portée sur l'usage des intuitions en philosophie. Cependant, en s'y attardant davantage, on finit par voir les conséquences potentielles de telles études, puisque le renvoi aux intuitions en philosophie semble effectivement étroitement associée à l'idée qu'il existerait des critères nécessaires et suffisants derrière nos concepts et que nos intuitions sur des cas particuliers permettraient d'en délimiter les contours. Les expériences de pensée développées par Gettier¹⁶ visent par exemple à montrer que la définition classique de la connaissance – croyance vraie justifiée – n'offre pas des conditions *suffisantes* au concept de connaissance. Les expériences de pensée développées en éthique impliquant des trains hors de contrôle (*trolley problems*)¹⁷ visent à montrer que le modèle conséquentialiste n'offre pas une conception satisfaisante du bien ou du devoir moral. Bref, il semblerait que le renvoi aux intuitions en philosophie, au moins dans plusieurs cas, soit associé à une recherche (implicite) de conditions nécessaires et suffisantes définissant nos concepts. Le problème, évidemment, est que si l'on se fie aux études, une telle quête est tout simplement vouée à l'échec. William Ramsey résume bien cette idée :

« If being an intuitive instance of X is simply a matter of having a cluster of properties that is sufficiently similar to some prototype representation, and if there are a number of different ways this can be done, some of which may vary over different contexts, then any crisp definition comprised of some subset of these properties and treating them as necessary and sufficient is never going to pass the test of intuition. It will *always* admit of intuitive counterexamples because the range of diversity sanctioned by our conceptual representation scheme will be much greater than that allowed by any tidy, straightforward definition. » (Ramsey, 1998, 171).

15 La théorie des prototypes et la thèse d'une structure probabiliste des concepts ne font bien sûr pas l'unanimité, et des théories alternatives ont été développées, par exemple la « théorie théorie » des concepts (*theory theory of concepts*) ou encore la théorie de l'atomisme conceptuel (*conceptual atomism*), sans parler de la théorie plus radicale de l'éliminativisme conceptuel. Voir Laurence & Margolis (1999) pour une présentation de ces alternatives. Toutes ces théories ont cependant en commun le rejet de la théorie classique des concepts, dont la nature inadéquate semble être l'objet d'un consensus.

16 Je présente ces expériences de pensée à la section 4.1.1.

17 Voir note précédente.

Que doit-on conclure de ces observations? La question qu'il semble important d'aborder est la suivante : la méthode philosophique de renvoi aux intuitions peut-elle préserver sa pertinence même s'il est vrai que nos concepts n'ont pas de conditions nécessaires et suffisantes? On peut au minimum répondre que rien ne semble *a priori* empêcher une telle possibilité. Il ne semble pas en effet y avoir un lien de nécessité entre la méthode philosophique et ce qu'on appelle la conception classique des concepts. Même en l'absence de conditions nécessaires et suffisantes, un appel aux intuitions peut s'avérer utile dans la mesure où il permet d'éclairer nos concepts. Comme le soulignent George Graham et Terry Horgan : « The exercise of generating counterexamples to putative analyses is often ideologically illuminating anyways. We understand the concept of *knowledge* better, for instance, upon appreciating the kinds of scenarios in which justified true belief that *p* does not constitute knowledge that *p*. » (Graham et Horgan 1998, 274).

Si l'on juge cette réponse trop timide, on pourrait être tenté d'offrir une réponse beaucoup plus forte, qui consiste à dire que les études sur la nature des concepts n'ont pas ou n'ont que très peu d'impact sur la méthode philosophique, pour la simple raison que nos intuitions ne se réduisent pas à des concepts mais renvoient d'abord et avant tout à des *phénomènes*. Un tel argument est notamment développé par Ernest Sosa :

« The use of intuitions in philosophy should not be tied exclusively to conceptual analysis. Consider some main subjects of prominent debate: utilitarian versus deontological theories in ethics, for example, or Rawls's theory of justice in social and political philosophy, or the externalism/internalism debate in epistemology; and many others could be cited to similar effect. They are not controversies about the conceptual analysis of some concept. They seem moreover to be disputes about something more objective than just a description or analysis of our individual or shared concepts of the relevant phenomena. Yet, they have been properly conducted in terms of hypothetical examples, and intuitions about these examples. The questions involved are about rightness, or justice, or epistemic justification. Some such questions concern an ethical or epistemic subject matter, and not just our corresponding concepts. » (Sosa, 2007, 101).

Un tel argument a le mérite de soulever plusieurs questions importantes : quelle est la nature de la relation entre intuitions et concepts? Une intuition renvoie-t-elle toujours nécessairement à un concept, ou est-il possible d'avoir une intuition sur un sujet sans que celle-ci ne reflète un quelconque concept que l'on posséderait sur ce sujet? La position défendue par Sosa semble cependant soulever plus de problèmes qu'elle n'apporte de solutions. Sans vouloir élaborer en

longueur sur cette question, la thèse selon laquelle toute intuition porte sur un concept paraît plausible. Après tout, il semble que l'on ne puisse avoir une intuition que sur des sujets que l'on *comprend*, et il semble également que l'on ne puisse comprendre un sujet que si l'on possède un concept correspondant à ce sujet. Par exemple, pour que j'aie l'intuition que les personnages décrits par Gettier ne sont pas véritablement en possession de connaissance, il me faut posséder le concept de connaissance, sans quoi je ne pourrais tout simplement pas comprendre le scénario décrit par Gettier. Idem pour les scénarios de trains à la dérive, qui requièrent la possession de plusieurs concepts moraux. Évidemment, c'est une chose de dire que pour toute intuition, nous possédons toujours un concept qui est rattaché au contenu de cette intuition, et une autre chose de dire qu'une intuition *porte toujours sur* un concept. Néanmoins, entre admettre la première proposition et admettre la seconde, il semble n'y avoir qu'un petit pas à franchir.

Quoiqu'il en soit, il est clair que l'argument défendu par Sosa ouvre une boîte de Pandore, et il s'agit d'une boîte qu'il ne semble pas nécessaire d'ouvrir, pour la simple raison que l'on peut défendre la méthode philosophique de renvoi aux intuitions même s'il s'avérait exact que toute intuition porte ultimement sur un concept. Car même s'il est vrai que nos concepts ne comportent aucun critère de nécessité et de suffisance, certaines études tendent à montrer qu'il y a des ressemblances considérables dans la manière dont les individus conçoivent la typicité des concepts : « [T]here appears to be considerable overlap between intuitive judgments of different subjects with regard to highly typical instances[...] » (Ramsey, 1998, 175)¹⁸. On peut donc espérer parvenir à un certain accord intuitif entre individus même en l'absence de frontières rigides délimitant nos concepts.

Comme ce fut le cas pour les études en économie comportementale discutées précédemment, on doit tout de même tirer quelques leçons de prudence de ces études sur la nature des concepts. La leçon principale que l'on doit tirer est qu'il est peu probable que l'on puisse jamais parvenir à offrir des définitions *courtes et simples* de concepts normatifs qui obtiendraient une approbation intuitive universelle. Cela semble être une conséquence de la structure probabiliste des concepts : plus on vise la simplicité dans nos définitions, plus on risque de rencontrer de contre-intuitions à ces définitions; inversement, plus on cherche à intégrer un grand nombre d'intuitions associées à un concept, plus on risque de se retrouver avec des définitions longues et complexes : « [I]f philosophers are going to insist on air-tight analyses that admit no intuitive counterexamples, then they must abandon the hope for tidy definitions and develop a tolerance for highly disjunctive

¹⁸ Ramsey poursuit cependant en soulignant qu'il y a moins de similarités entre individus pour les cas « modérément typiques ».

and heavily qualified definitions » (*ibid.*).

Or, une telle inévitabilité semble avoir été assimilée par plusieurs philosophes contemporains qui utilisent les intuitions à des fins de construction théorique. Un exemple récent parmi d'autres est celui de George Sher qui, dans le livre *Who Knew? Responsibility Without Awareness* (2009), développe une série d'expériences de pensée visant à offrir une définition satisfaisante de la responsabilité morale. Or, la définition finale offerte par Sher est hautement disjonctive, ce qui semble illustrer la reconnaissance par Sher de la nécessité de sacrifier un peu de simplicité afin d'atteindre une définition potentiellement universelle.

Ramsey répondrait sans doute qu'aucune définition ne peut aspirer à être véritablement universelle, puisque l'application d'un concept donnera toujours lieu à des contre-exemples intuitifs (« It will *always* admit of intuitive counterexamples » (Ramsey, 1998, 171)). Une telle déclaration semble cependant relever davantage d'un acte de foi qu'être véritablement une conséquence de la structure probabiliste des concepts. Quoiqu'il en soit, ce genre de considérations sur l'universalité des intuitions nous renvoient à un autre débat, étroitement lié au problème de la fiabilité des intuitions, qui concerne le phénomène du désaccord intuitif entre individus. C'est vers ce débat que je me tourne dans la prochaine section.

3.2. Désaccord intuitif et philosophie expérimentale

Le phénomène du désaccord entre individus rationnels est souvent souligné en philosophie pour défendre une forme de relativisme ou pour s'opposer à une forme d'objectivisme. Ce qu'on appelle l'argument du désaccord (ou argument de la relativité) trouve également écho dans le domaine des intuitions, où plusieurs adversaires de l'utilisation des intuitions en philosophie renvoient au phénomène du désaccord intuitif afin de critiquer cette pratique. Ces sceptiques de l'intuition soupçonnent que les intuitions suscitées par les philosophes lors d'expériences de pensées sont loin d'être universelles, et qu'en sortant de leur tour d'ivoire académique, ils risqueraient de constater un important désaccord intuitif entre individus. Si ce constat s'avérait exact, la philosophie spéculative traditionnelle, ou « armchair philosophy », pourrait voir sa pertinence menacée, dans la mesure où une forte relativité des intuitions limiterait grandement la portée des théories philosophiques construites sur la base de ces intuitions.

Un mouvement de philosophes s'est développé au tournant du millénaire ayant pour but de vérifier empiriquement les intuitions du commun des mortels. Ces « philosophes expérimentaux »

ont mené de nombreuses enquêtes afin de tester les expériences de pensée développées par les philosophes sur un public plus large. La mission du mouvement de philosophie expérimentale est décrite explicitement sur le site-web officiel du mouvement :

« Experimental philosophy, called x-phi for short, is a new philosophical movement that supplements the traditional tools of analytic philosophy with the scientific methods of cognitive science. So experimental philosophers actually go out and run systematic experiments aimed at understanding how people ordinarily think about the issues at the foundation of the philosophical discussion. »
(<http://pantheon.yale.edu/~jk762/ExperimentalPhilosophy.html>)

Plus d'une centaine de philosophes s'identifient à ce mouvement, dont plusieurs ont publié au cours des dix dernières années diverses enquêtes empiriques testant les intuitions de la population sur des enjeux reliés à l'épistémologie, la philosophie du langage, la philosophie de l'esprit, la métaphysique, la philosophie de l'action ou la philosophie morale.¹⁹ La plupart de ces études ont été menées à partir de questionnaires distribués à un échantillon d'individus – parfois des classes d'étudiants de baccalauréat, parfois des échantillons plus aléatoires représentant mieux l'ensemble de la population, parfois encore des échantillons provenant de différentes cultures. Dans la plupart des cas, les questionnaires reproduisaient mots pour mots des expériences de pensée classiques développées par des philosophes et demandaient aux sujets de choisir parmi différentes options celle qui représentait le mieux leur intuition ou jugement par rapport à cette expérience de pensée. Voici deux exemples d'expériences de pensée testées lors de ces études :

« Bob has a friend, Jill, who has driven a Buick for many years. Bob therefore thinks that Jill drives an American car. He is not aware, however, that her Buick has recently been stolen, and he is also not aware that Jill has replaced it with a Pontiac, which is a different kind of American car. Does Bob really know that Jill drives an American car, or does he only believe it? (a) Really knows; (b) Only believes »
(Weinberg et al., 2001).

« One day Charles was knocked out by a falling rock; as a result his brain was “rewired” so that he is always right whenever he estimates the temperature where he is. Charles is unaware that his brain has been altered in this way. A few weeks later, this brain rewiring leads him to believe that it is 71 degrees in his room. Apart from his estimation, he has no other reasons to think that it is 71 degrees. In fact, it is 71

19 On peut trouver sur le site-web du mouvement une liste des principaux articles et livres, classés par thème, publiés par des philosophes expérimentaux. Un site-soeur donne également une liste quasi-exhaustive des philosophes qui s'identifient au mouvement. Voir la section « Contributors » dans le menu de gauche sur le site <http://experimentalphilosophy.typepad.com/é>

degrees. Please indicate to what extent you agree or disagree with the following claim: "Charles knows that it is 71 degrees in his room". (a) Strongly agree; (b) Agree; (c) Neutral; (d) Disagree; (e) Strongly disagree. » (Swain et al., 2008).

La première expérience de pensée est un exemple classique de type Gettier (1963) visant à susciter l'intuition qu'une croyance vraie justifiée n'est pas toujours une connaissance, tandis que la seconde expérience de pensée provient de Lehrer (1990) et vise à susciter des intuitions favorables à la théorie « internaliste ». Or, tant pour l'étude de Weinberg et al. que pour celle de Swain et al., les sujets se sont montrés beaucoup moins unanimes dans leurs réponses intuitives que ce qu'auraient sans doute espéré Gettier ou Lehrer. Dans l'étude sur la voiture américaine, 26% des sujets (N=66) ont conclu que Bob avait une véritable connaissance plutôt qu'une simple croyance. L'étude sur le thermomètre humain (généralement nommée « Truetemp Case ») était quant à elle particulière en ce qu'elle visait à tester si les réponses des sujets étaient affectées par le fait de présenter une autre expérience de pensée, dont la « bonne réponse » est plus évidente, avant celle de Truetemp. Les auteurs résument leur conclusion ainsi :

« Our results show that compared to subjects who receive the Truetemp Case first, subjects first presented with a clear case of knowledge are less willing to attribute knowledge in the Truetemp Case, and subjects first presented with a clear case of non-knowledge are more willing to attribute knowledge in the Truetemp Case » (Swain et al., 2008, 138).

Les auteurs interprètent ces résultats comme constituant une menace sérieuse à la philosophie spéculative traditionnelle : « We contend that this instability undermines the supposed evidential status of these intuitions, such that philosophers who deal in intuitions can no longer rest comfortably in their armchairs » (*ibid.*). Weinberg et al. concluent également que leur étude devrait soulever d'importants doutes quant à la pertinence de ce qu'ils nomment le « romantisme des intuitions » (« intuition driven romanticism »), qu'ils décrivent comme l'acceptation de la thèse suivante : « Knowledge of the correct norms [...] is implanted within us in some way, and with the proper process of self-exploration we can discover them » (*ibid.*, 194). Bref, à en croire ces auteurs, le mouvement de philosophie expérimentale sonne ni plus ni moins l'arrêt de mort du renvoi aux intuitions en philosophie.

Que penser de ce type de nécrologies? À mon sens, les philosophes expérimentaux sont encore bien loin de pouvoir légitimement crier victoire. Les critiques sont en effet nombreuses de la part de ceux que Weinberg et al. qualifient de « romantiques ». Certains, dont Ernest Sosa, jugent

que ce type d'enquêtes empiriques s'apparentent à des coups d'épée dans l'eau, pour la simple raison qu'on ne devrait aucunement s'attendre à ce que les intuitions de la population se comparent à celles des philosophes : « When we rely on intuitions in philosophy [...] we manifest a competence that enables us to get it right on a certain subject matter, by raising our beliefs on the sheer understanding of their contents. How might survey results create a problem for us? » (Sosa, 2007, 102). Michael Huemer défend une thèse similaire lorsqu'il soutient que l'argument du désaccord intuitif s'apparente à une forme de « veto de l'idiot » :

« The fact that some people fail to perceive a fact simply does not destroy the possibility that other people do perceive it. This would be to invoke what I call "the idiot's veto". [...] The thesis [of the idiot's veto is] that any individual has the power to block a fact from the realm of objectivity or knowledge, merely by persistently refusing to agree with it, and resisting all efforts to educate him » (Huemer, <http://home.sprynet.com/~owl1/rand.htm#5.4.1>)

Un philosophe expérimental pourrait cependant répliquer que les intuitions de non-philosophes ne peuvent aussi facilement être jugées inintéressantes ou inutiles pour la simple raison que les intuitions philosophiques ont une prétention à la validité qui s'étend bien au-delà du cercle restreint de la philosophie académique. La méthode de renvoi aux intuitions et les expériences de pensée semblent en effet présupposer la croyance que ces intuitions peuvent être reconnues par toute personne minimalement compétente, et il semble exagéré de soutenir que la population en général n'est pas minimalement compétente. C'est là un point que certains défenseurs de la philosophie spéculative traditionnelle sont prêts à concéder aux philosophes expérimentaux. Simon Cullen (2010) affirme par exemple :

« It is true that when a philosopher says 'We find X intuitive' ('we find it natural to say X'), she does not mean 'I find X intuitive', and she does not, in general, mean 'We philosophers find X intuitive'. Rather, she means that competent, thoughtful speakers would, under certain circumstances, find X intuitive. Experimental philosophers are right to emphasize that appeals to intuition typically engender empirical commitments. » (Cullen, 2010, 279)

D'autres seront cependant moins prêts à concéder ce point, et répliqueront que la plupart des gens se laissent trop facilement influencer par des facteurs anodins et n'ont pas suffisamment cultivé de bonnes habitudes intellectuelles pour que leurs intuitions puissent être jugées aussi valables que celles de philosophes. C'est ce que Sosa soutient lorsqu'il affirme que les philosophes manifestent

une compétence qui leur permet d'avoir raison sur certains sujets (Sosa, 2007, 102). De par leur formation, les philosophes (entre autres) sont plus à même d'éviter certains pièges cognitifs dans lesquels la plupart des gens tombent facilement.²⁰ Ces cas d'irrationalité peuvent nous amener à conclure que même s'il est vrai que les intuitions philosophiques devraient en principe être reconnues par tout individu compétent, le fait est que la compétence intellectuelle est une denrée malheureusement plus rare qu'on pourrait croire. Pour cette raison, le constat d'un désaccord intuitif important dans la population ne devrait pas empêcher les adeptes de philosophie spéculative traditionnelle de bien dormir.

Si l'on juge cette ligne argumentative insatisfaisante, par exemple parce qu'elle suppose un niveau trop généralisé d'incompétence intellectuelle, plusieurs autres stratégies demeurent disponibles afin de défendre la philosophie spéculative traditionnelle. Une première stratégie consiste à dire que le désaccord intuitif demeure l'exception plutôt que la règle, et que les études réalisées par les philosophes expérimentaux portent sur des cas ambigus où il est normal de constater des intuitions populaires partagées. Par exemple, dans « A Defense of Intuitions » (2008), S. Matthew Liao analyse l'étude de Swain et al. (2008) mentionnée plus haut de même qu'une autre étude menée par Weinberg et al. (2001) et conclut que dans les deux cas, les auteurs ne portent pas assez attention au fait que les intuitions des sujets sont quasi-unanimes lorsqu'on leur présente des expériences de pensée moins ambiguës (Liao, 2008, 253). Weinberg et al. admettent que ces cas consensuels supportent l'idée qu'il existerait un « noyau universel d'épistémologie populaire »,²¹ mais Liao juge qu'ils ne réalisent pas réellement l'importance d'un tel constat pour la théorie qu'ils cherchent à défendre : “[T]hey fail to recognize that this is a large concession [...] to the extent that it undermines their claim that there is not a fixed set of intuitions about a particular thought-experiment to which we can appeal” (Liao, 2008, 254).

Une autre stratégie disponible consiste à attaquer plus directement la méthodologie utilisée par les philosophes expérimentaux, plus précisément leur recours à des questionnaires. Par exemple, Simon Cullen (2010), après avoir effectué un survol d'un grand nombre d'études en philosophie expérimentale, souligne que ces études ont un commun leur manque d'attention portée à la méthodologie :

« In my review of the growing body of experimental philosophy literature I have found just one reference to a serious discussion of survey methodology. The conclusion to be drawn is that, despite their

20 Voir section 3.1.1 pour des exemples de tels pièges cognitifs.

21 Les termes exacts utilisés par les auteurs sont « a universal core of folk epistemology » (Weinberg et al., 201, 240).

pretensions, many experimental philosophers have given no serious thought to methodology. This not only undermines their claim to be doing science, as we shall see, it often leaves the philosophical significance of their findings unclear » (Cullen, 2010, 279).

Cullen se livre ensuite à une analyse impressionnante des nombreux problèmes pouvant résulter de l'utilisation de questionnaires, le plus important étant que la réponse donnée à un questionnaire risque souvent de ne pas refléter une véritable « intuition » du sujet. Ce constat amène Cullen à reprendre à son compte l'expression développée par Weinberg et al. (2006) et à parler de la philosophie expérimentale comme étant une forme de « romantisme du questionnaire » (« survey driven romanticism »), qu'il définit, en parallèle avec Weinberg et al., comme l'acceptation de la thèse suivante : « People's philosophical intuitions are implanted within them in some way, and by administering simple surveys we can discover them » (*ibid.* 278).

Cullen est allé jusqu'à effectuer sa propre étude par questionnaire afin de montrer les dangers potentiels de ce type d'enquête. Bien qu'il n'ait pas lui-même utilisé une telle expression, l'enquête de Cullen tend à soutenir la croyance populaire selon laquelle « on peut faire dire ce qu'on veut à des sondages ». Cullen a repris les deux expériences de pensée reproduites plus haut – la voiture américaine et l'homme thermomètre – et a simplement modifié les choix de réponses offerts aux sujets. Pour le cas de la voiture américaine, Cullen a comparé deux types de choix de réponses : dans un premier cas, il a repris la formule de Weinberg et al. en offrant les choix « Really knows » et « Only believes », et dans un deuxième cas, il a simplement posé la question « Does Bob know that Jill has an American car? » en offrant les options Oui et Non. Dans le premier cas, les réponses obtenues sont similaires à celles de Weinberg et al. : 71% des sujets (N=233) ont répondu « Only believes », comparativement à 74% (N=66) pour Weinberg et al. Dans le deuxième cas cependant, le changement de formulation de la question a mené à des résultats assez différents : seulement 58% des répondants ont répondu que Bob n'étaient pas en possession d'une connaissance. Les résultats sont encore plus spectaculaires pour l'expérience de pensée du thermomètre humain :

« The results were even more striking. Only 28% of subjects in the non-dichotomous group answered “really knows” compared with 57% of subjects in the dichotomous group who answered “knows” (p < .0001) – that is, subjects were *twice as likely* to respond “knows” than “really knows”. » (*ibid.* 290).

Ces résultats amènent Cullen à conclure que les philosophes expérimentaux ont encore bien des croûtes méthodologiques à manger avant de pouvoir annoncer de façon crédible la mort de la

philosophie spéculative traditionnelle. De façon plus constructive, il décrit ce qui lui semble être la principale faille de la démarche des expérimentaux : « Whenever experimental philosophers force subjects to respond to conceptually complex scenarios by choosing one of two mutually exclusive answers, they risk misrepresenting the structure and subtleties of folk concepts, and their subjects' understanding of the scenarios » (*ibid.*, 292).

Cette idée de « complexité conceptuelle » est étroitement associée à une troisième stratégie argumentative utilisée par certains défenseurs de la méthode traditionnelle, qui consiste à soutenir que le désaccord observé ne serait pas véritablement un désaccord *substantiel* mais refléterait simplement une différente compréhension de certains concepts. C'est là une critique qui a été soulevée par plusieurs philosophes, mais qui est principalement associée à Ernest Sosa :

« The bearing of these surveys on traditional philosophical issues is questionable [...] because the experimental results really concern in the first instance only people's responses to certain words. But verbal disagreement *need* not reveal any substantive, real disagreement, if ambiguity and context might account for the verbal divergence. [...] The experimentalists have not yet done enough to show that they have crossed the gaps created by such potential differences in meaning and context, so as to show that supposedly commonsense intuitive belief is really not as widely shared as philosophers have assumed it to be » (Sosa, 2007, 102-3)

De nombreuses autres réponses peuvent être et ont été offertes aux philosophes expérimentaux, mais les quelques exemples que j'ai présentés ici me semblent suffisants pour montrer qu'il y a encore loin de la coupe aux lèvres dans le camp expérimental. Le débat entre les deux camps se poursuit bien sûr, et il est tout à fait possible que le mouvement X-Phi parvienne à surmonter ses défis méthodologiques et à montrer qu'il existe réellement des cas de désaccords intuitifs *substantiels* entre individus *compétents*. Un tel constat aurait-il réellement des conséquences dramatiques sur la méthode philosophique traditionnelle? Il faudrait d'abord savoir si de tels désaccords constituent l'exception ou bien la règle. Comme le souligne Liao, certaines études de philosophes expérimentaux tendent en fait à montrer qu'il existe véritablement un noyau universel d'épistémologie populaire (Liao, 2008, 254), et un tel noyau pourrait être encore plus important dans d'autres domaines, notamment dans le domaine moral. Or, il y a un important travail philosophique à faire à partir d'un tel noyau. Par exemple, comme je le montrerai au chapitre 5, l'intuitionnisme éthique cherche précisément à construire un système moral à partir d'un petit noyau de principes moraux élémentaires. La méthode de l'équilibre réfléchi, présentée au chapitre 4, offre

également une manière de concevoir le travail philosophique malgré le phénomène du désaccord.

Bref, les travaux en philosophie expérimentale ne semblent pas sérieusement menacer la philosophie traditionnelle. Au contraire, la philosophie traditionnelle peut grandement bénéficier de telles études, comme le souligne Liao en conclusion de son article :

« Through quantitative and qualitative research [...] experimentalism can help us identify areas in which we have conflicts of intuition, and may be able to help us distinguish among Apparent Conflicts, Conflicts out of Biases and Genuine Conflicts. When Genuine Conflicts have been uncovered, [traditional philosophical methodology] may then be able to help us resolve these conflicts by providing positive theories for why certain intuitions should be excluded » (*ibid.* 260).

Je me tourne à présent vers un dernier grand type d'argument souvent employé pour souligner le manque de fiabilité des intuitions, à savoir l'argument de *l'origine problématique* des intuitions.

3.3. Origines problématiques

D'où proviennent nos intuitions? Cette question ne doit pas être confondue avec celle soulevée à la section 2.4, alors que l'on cherchait à localiser la *faculté* intuitive dans le cerveau humain. La question que je pose ici concerne non pas l'origine ou l'emplacement de notre faculté, mais plutôt l'origine du *contenu* de nos intuitions. Cela revient à se demander pourquoi nous avons les intuitions que nous avons plutôt que des intuitions contraires. Pourquoi, par exemple, la plupart des gens jugent-ils spontanément qu'il est moralement juste ou obligatoire de tirer le levier afin de faire dévier le train vers le rail ne contenant qu'un seul individu attaché, alors qu'ils jugent par ailleurs qu'il est mal de pousser l'homme obèse sur les rails afin d'éviter la mort de quatre individus?²²

Plusieurs philosophes affirment qu'une enquête plus approfondie de l'origine de nos intuitions permet de révéler que celles-ci ne sont pas fiables. C'est cet argument que je souhaite explorer dans cette section. Mon analyse portera entièrement sur les intuitions morales, principalement parce que l'origine de nos intuitions épistémiques, métaphysiques et autres me semble être un sujet beaucoup plus ardu. Il est en effet beaucoup moins clair que l'on puisse offrir

²² Voir section 4.1.1 pour une présentation de cette expérience de pensée classique en philosophie morale.

une généalogie de nos intuitions épistémiques (par exemple) et d'en retracer les sources évolutives que cela semble possible dans le cas des intuitions morales, ce qui explique sans doute pourquoi l'argument originaire concernant les intuitions a principalement été soulevé et débattu dans un contexte de philosophie morale.

À la question « d'où proviennent nos intuitions? », Robert Cummins (1998) offre l'une des analyses philosophiques les plus éclairantes qui soient. Cummins distingue entre cinq sources possibles de nos intuitions, à savoir (1) nos croyances ordinaires, qui sont le fruit de notre éducation/socialisation; (2) nos concepts; (3) notre langue; (4) des théories tacites; (5) des théories explicites. La distinction entre (2) et (3) n'est pas très claire, et Cummins n'analyse que rapidement ces deux cas, soulignant simplement qu'ils se réduisent ultimement aux options (1), (4) et (5), dans la mesure où notre langue est acquise à travers le processus de socialisation et que nos concepts fonctionnent à bien des égards comme des théories. Il concentre donc son analyse sur les options (1), (4) et (5), jugeant qu'il s'agit des trois principales sources de nos intuitions. Or, Cummins conclut que, dans la mesure où elles proviennent de notre éducation, nos intuitions ne sont tout simplement pas fiables, et dans la mesure où elles proviennent de théories (tacites ou explicites), elles sont tout simplement inutiles. Voilà l'essentiel de l'argument originaire qui l'amène à discréditer le renvoi aux intuitions en philosophie. Il me paraît bon de prendre le temps d'analyser chacune de ces trois sources afin de voir si effectivement elles sont toutes problématiques; la section 3.3.1 abordera donc la socialisation comme source de nos intuitions, la section 3.3.2 abordera les théories tacites et la section 3.3.3, les théories explicites.

Avant de débiter l'analyse de ces trois sources, notons qu'une explication alternative de l'origine des intuitions n'est pas soulignée par Cummins, à savoir la possibilité que nos intuitions soient des perceptions de propriétés morales *sui generis*. Ce type d'explication s'inscrirait dans le cadre d'une théorie réaliste des valeurs, théorie défendue par plusieurs philosophes contemporains.²³ Il est surprenant que Cummins ne considère aucunement l'option réaliste, dans la mesure où il s'agit d'une théorie qui semble parvenir à éviter les pièges qu'il prête aux autres explications. Pour ma part, je préfère ne pas discuter en longueur de l'option réaliste, pour la simple raison qu'on ne peut aborder adéquatement cette théorie sans considérer sérieusement les enjeux métaphysiques et ontologiques qu'elle soulève, ce que je n'ai pas le luxe de faire dans le cadre de mon mémoire. Je soulignerai néanmoins quelques points importants concernant la théorie réaliste en fin de chapitre.

23 Voir entre autres Tappolet, 2000.

3.3.1. La socialisation comme source de nos intuitions

La première source considérée par Cummins est sans doute la plus évidente : nos intuitions morales pourraient simplement provenir de notre éducation morale, ou plus largement, du processus de socialisation qui nous permet à tous d'assimiler durant l'enfance les valeurs et les normes de notre société. Des conditionnements sociaux de plusieurs sortes forment notre caractère et notre conception de ce qui est bien ou mal, obligatoire ou surrogatoire.

De nombreux philosophes considèrent l'éducation morale comme la source principale de nos jugements et intuitions morales. Jesse Prinz est probablement le plus connu, sinon certainement le plus convaincu des portes-étendards de l'explication éducationnelle de nos intuitions morales. Dans « The Emotional Construction of Morals » (2007) et ailleurs (e.g., Prinz, 2008), Prinz décrit certains des nombreux mécanismes de renforcements émotionnels qui amènent les jeunes enfants à intérioriser très vite les valeurs et normes de leur société :

« When children violate harm norms, three responses are most typical (Hoffman, 2000). First, there's power assertion: if Lulu hits her baby brother, her parents are likely to get very angry, and they may even take hold of Lulu in a threatening way. Second, there's love withdrawal: if little Lulu smacks Daddy's face while playing, he may say, "That's not nice; I'm not going to play with you any more." Most importantly, there's induction of empathetic distress: when Lulu hits her baby brother, and he begins to cry, they will say, "Look what you've done; you made your brother cry." Each of these methods can be quite effective » (Prinz, 2008, 431).

Prinz se dit par ailleurs confiant qu'il n'existe pratiquement aucune limite aux valeurs qu'on peut en principe faire assimiler à des individus grâce à de tels mécanismes de renforcements émotifs, ce qui le rapproche à bien des égards de la théorie behavioriste classique :

« I tend to think, somewhat cynically, that the range of moral rules is relatively unconstrained. [...] I adamantly believe that we could teach people to value recreational torture of babies. [...] I'm sure a search of the anthropological record would uncover groups that tortured babies for fun – especially if the babies belonged to enemy groups that were defeated in battle » (*ibid.*, 429).

Si l'explication éducationnelle de l'origine de nos intuitions morales s'avérait exacte, on se retrouverait avec un problème potentiel quant à la légitimité du renvoi aux intuitions en philosophie morale. En effet, si nos intuitions ne sont que le produit d'une intériorisation de normes sociales, et que ces normes varient considérablement d'une société à l'autre, on voit mal comment on peut

renvoyer à de telles intuitions afin de justifier nos principes ou théories éthiques. Il semblerait y avoir là une forme de *chauvinisme*, comme le décrit Robert Cummins :

« I am inclined to think a lot of our so called intuitions about fairness are beliefs of exactly this kind. In the current jargon, they are a part of the “values” we picked up in our families, schools, and play groups, our reading and TV watching. It would take a real mad kind of chauvinist, however, to suppose that these have the sort of special epistemological status required to ground a philosophical theory of distributive justice. If I can show that your reflective equilibrium depends on intuitions with this source, I can dismiss it out of hand. » (Cummins, 1998, 119).

Il est facile de voir en quoi l'explication éducationnelle pourrait miner la valeur épistémique de nos intuitions : si la seule différence entre un individu A, qui a l'intuition que X est mal, et un individu B, qui a l'intuition que X est bien, est que A provient d'une société où X est condamné alors que B provient d'une société où X est approuvé, un renvoi aux intuitions dans un cas comme dans l'autre reviendrait simplement à un endossement de la norme sociale en vigueur. S'il s'avérait vrai que nos intuitions se réduisent entièrement à une intériorisation de normes sociales, le renvoi aux intuitions paraîtrait aboutir ni plus ni moins à un relativisme culturel.

On répondra à ce type d'arguments que le fait qu'une intuition résulte de l'intériorisation d'une norme sociale ne signifie pas pour autant qu'elle perde sa valeur. En effet, toutes les normes sociales transmises aux individus ne sont pas arbitraires : certaines connaissances scientifiques peuvent par exemple être transmises culturellement, et le fait qu'elles soient transmises de cette façon ne vient aucunement compromettre leur validité.²⁴ On pourrait donc tenter de montrer que plusieurs de nos normes morales n'ont rien d'arbitraire malgré le fait qu'elles soient des produits culturels transmis à travers le processus de socialisation. Dans cette optique, on pourrait utiliser la théorie des jeux afin de montrer que certaines règles morales résultent d'une forme de nécessité pratique. On pourrait également souligner que certaines normes morales, bien que transmises culturellement, n'ont rien d'arbitraire puisqu'elles sont profondément enracinées en l'être humain. Par exemple, le principe de justice selon lequel il est bien de remettre à chacun ce qui lui est dû est certes transmis culturellement, mais plusieurs études montrent qu'il s'agit d'un principe qu'endossent ou développent naturellement non seulement tout être humain normalement constitué, mais également plusieurs autres grands primates.²⁵ Ce dernier type d'explication renvoie à l'existence de « théories tacites », que j'aborde dans la section suivante.

24 Merci à Christine Tappolet d'avoir offert cet exemple.

25 Voir par exemple les études résumées par Franz De Waal (De Waal, 2009).

3.3.2. Les théories tacites comme sources de nos intuitions

Après avoir conclu que l'explication éducationnelle de nos intuitions morales nous condamnerait à une forme de chauvinisme, Cummins se tourne vers la possibilité que nos intuitions soient plutôt le produit de théories tacites ou explicites. C'est là, selon lui, la seule explication qui permette au défenseur des intuitions de préserver un certain espoir quant à la fiabilité des intuitions : « Every remotely plausible story about the origin of philosophical intuitions, then, leads us to the view that they are generated by explicit or tacit theories of the target properties » (*ibid.*, 121).

Cummins ne prend pas la peine de définir ce qu'est une « théorie tacite », mais on parvient facilement à déduire ce qu'il entend grâce aux exemples qu'il en donne. Une théorie tacite est essentiellement un ensemble de croyances que l'individu possède naturellement, qui lui sont en quelque sorte *constitutives*. L'individu peut constater qu'il/elle possède ce type de croyances lorsque confronté/e à une situation qui fait surgir ces croyances en sa conscience. Il s'agit de croyances qu'il/elle accepte sans qu'on ait à lui enseigner explicitement. Le principe de justice décrit à la fin de la section précédente est un bon exemple de ce type de croyances. Or, si certains soutiennent que ces théories tacites sont entièrement *acquises* durant l'enfance, plusieurs philosophes et scientifiques contemporains affirment au contraire qu'elles sont innées.

Différentes théories contemporaines postulent l'existence de croyances morales innées ou autres formes de prédispositions morales. La théorie de la grammaire morale universelle, défendue principalement par Marc Hauser (2006, 2008), Susan Dwyer (2008), and John Mikhail (2008, 2011), affirme par exemple que les individus naissent avec une faculté morale qui s'apparente à une « grammaire morale », c'est à dire un ensemble de principes et de paramètres innés qui guident le développement moral des individus et leur permettent de *connaître intuitivement* certains faits moraux. Cette théorie s'inspire largement de la théorie développée par Noam Chomsky (1965, 1971) en linguistique. Une autre théorie innéiste contemporaine importante est celle de l'« intuitionnisme social », développée principalement par Jonathan Haidt (2001, 2004, 2008), qui, lui, postule l'existence de cinq modules moraux spécialisés – harm/care, fairness/reciprocity, in-group/loyalty, authority/respect, and purity/sanctity – qui prédisposeraient les individus à ressentir des sentiments d'approbation ou de désapprobation dans certains contextes pertinents : « [Children are born equipped with] an innate preparedness to feel flashes of approval or disapproval toward certain patterns of events involving other human beings » (Haidt et Joseph, 2004, 56).

Au-delà de leurs différences, les théories de la grammaire morale universelle et de l'intuitionnisme social ont en commun l'idée que la moralité n'est pas une simple fabrication humaine et que nos intuitions et jugements moraux ne se réduisent pas à la simple intériorisation de normes sociales : « The basic idea is that morality[...] is better described as *emerging from* the child (externalized) on a particular developmental schedule rather than being *placed into* the child from outside (internalized) on society's schedule » (Haidt et Bjorklund, 2008, 206). Cette hypothèse est renforcée par de nombreuses recherches en psychologie développementale qui tendent à montrer que les enfants ont une compréhension de l'univers moral dès un très jeune âge, un phénomène que l'on nomme la « précocité morale » et dont John Mikhail donne un aperçu général :

« Three- and four-year-old children use intent or purpose to distinguish two acts with same result (Baird, 2001). They also distinguish 'genuine' moral violations (e.g. theft, battery) from violations of social conventions (e.g. wearing pajamas to school; Smetana, 1983, Turiel, 1983). Four- and five-year-olds use a proportionality principle to determine the appropriate level of punishment for principals and accessories (Finkel, Liss & Moran, 1997). Five-year-olds display a nuanced understanding of negligence and restitution (Shultz, Wright & Schleifer, 1986). One man shoots and kills his victim on the mistaken belief that he is aiming at a tree stump. A second man shoots and kills his victim on the mistaken belief that killing is not wrong. Five- and six-year-olds distinguish cases like these in conformity with the distinction between mistake of law and mistake of fact, recognizing that false actual beliefs may exculpate, but false moral beliefs do not (Chandler, Sokol, & Wainryb, 2000). Five- and six-year-olds also calibrate the level of punishment they assign to harmful acts on the basis of mitigating factors, such as provocation, necessity, and public duty (Darley, Klosson, & Zanna, 1978). Six- and seven-year-olds exhibit a keen sense of procedural fairness, reacting negatively when punishment is inflicted without affording the parties notice and the right to be heard (Gold, Darley, Hilton & Zanna, 1984) ». (Mikhail, 2008, 354).

Tous ne sont bien sûr pas d'accord avec le modèle innéiste. Jesse Prinz est l'un des plus farouches opposants à ce type de théories, comme on peut le deviner en prenant connaissance de son explication essentiellement éducationnelle de l'origine des intuitions et jugements moraux. Mon but ici n'est cependant pas d'offrir un compte-rendu des arguments et contre-arguments entourant le débat sur l'innéisme moral, mais plutôt de considérer les conséquences possibles de telles prédispositions morales sur la fiabilité des intuitions qui en découleraient.²⁶

Cummins se montre lui-même sympathique à l'idée que nous aurions des « théories morales tacites » qui se présenteraient sous la forme de modules moraux spécialisés : « No doubt there is some innate theory [...] It looks like there might be an innate module responsible for at least some

26 J'aborde plus en détails le débat sur l'innéisme dans Giroux, 2011.

moral reasoning (Cummins 1996a, 1996b). I do not find it implausible to suppose that we are endowed with some innate philosophy, especially moral and social philosophy. » (Cummins, 1998, 123). Or, malgré son acceptation du modèle innéiste, ou à tout le moins son ouverture à ce modèle, Cummins juge que les intuitions qui seraient le produit de telles théories tacites ne seraient aucunement plus fiables ou pertinentes que les intuitions qui seraient le produit de l'éducation.

« But when you come to think of what these theories are for, there is no more reason to think that innate philosophy is a good basis for philosophy than that innate physics is a good basis for physics. We have cultural institutions like science and philosophy largely to *overcome* the limitations of our innate endowment. There is, then, no good reason to think that innate tacit theories that might underlie philosophical intuition are true, and some compelling general reasons for thinking that they are likely to be pretty inaccurate » (*ibid.*)

Le fait qu'une théorie ou croyance soit innée n'est donc aucunement un gage de vérité, pour la simple raison que le processus ayant mené au développement de ces théories ou croyances innées, à savoir le processus de sélection naturelle, n'est aucunement un processus qui garantit une quelconque *vérité*. Ces théories ou croyances innées ont été sélectionnées en raison de leurs qualités adaptatives, et aucunement en raison d'une quelconque véracité intrinsèque à ces croyances ou théories. Comme le souligne Cummins de façon fort amusante : « You would have to be either extremely cynical about morals, or extremely naive about social interactions among primates and other social species, to suppose that adaptive moral concepts are accurate » (*ibid.*, 122).

Peter Singer défend une position similaire concernant les intuitions morales résultant (vraisemblablement) du processus de sélection naturelle. Dans « Ethics and Intuitions » (2005), il s'intéresse aux intuitions suscitées lors des expériences de pensées impliquant des trains (« trolley thought experiments »), et plus spécialement à la différence de réactions intuitives selon qu'il s'agisse du scénario où il faut tirer le levier ou du scénario où il faut pousser un homme obèse sur les rails – qu'on distingue comme étant le « switch dilemma » et le « footbridge dilemma ». Se fondant sur les travaux de Joshua Greene (2001, 2007), il explique cette différence intuitive entre les deux scénarios comme étant le résultat d'une charge émotionnelle négative beaucoup plus forte associée au « footbridge dilemma ». Il interprète ensuite cette charge émotionnelle comme étant un produit évolutif, arguant qu'une telle source devrait nous amener à discréditer la pertinence de notre réaction face au scénario :

« So the salient feature that explains our different intuitive judgments concerning the two cases is that the

footbridge case is the kind of situation that was likely to arise during the eons of time over which we were evolving; whereas the standard trolley case describes a way of bringing about someone's death that has only been possible in the past century or two, a time far too short to have any impact on our inherited patterns of emotional response. But what is the moral salience of the fact that I have killed someone in a way that was possible a million years ago, rather than in a way that became possible only two hundred years ago? I would answer: none. » (Singer, 2005, 348).

La solution proposée par Singer est de séparer nos évaluations véritablement « rationnelles » de nos réactions émotives immédiates. Il souligne que certains des sujets qui participaient à l'étude de Greene parvenaient à surpasser leur réaction émotive immédiate et adoptaient la réponse que Singer juge être la bonne, à savoir qu'il est tout aussi moralement obligatoire de pousser le gros homme sur les rails qu'il est obligatoire de tirer le levier, pour la simple raison que les conséquences sont les mêmes dans les deux scénarios. Singer mentionne également l'étude sur l'inceste de Jonathan Haidt (Haidt 2001), où certains sujets parvenaient aussi à dépasser leur révolusion immédiate et à reconnaître qu'il n'y avait rien de mal à ce que le frère et la sœur ait une relation sexuelle consentante. Bref, selon Singer, la solution consiste simplement à séparer le bon grain de l'ivraie dans les éléments qui constituent nos évaluations.

Le problème avec ce type de solution, cependant, est qu'elle ne fait en vérité que repousser légèrement le problème de la fiabilité des intuitions. Car une fois que l'on reconnaît que l'on doit accepter seulement nos « jugements rationnels » comme étant fiables, on se retrouve confrontés au besoin d'expliquer pourquoi ces jugements doivent être considérés fiables. Après tout, tant notre « jugement rationnel » que notre réaction émotive immédiate reposent ultimement sur une intuition, et la seule différence entre les deux cas est l'intensité de l'affect rattaché à l'intuition. Or, qu'est-ce qui confère plus de fiabilité à nos intuitions moins chargées émotivement? Singer répondrait sans doute que l'émotion vient obscurcir le fait évident que la seule considération pertinente est que la mort d'une personne est moins grave que la mort de cinq personnes. Mais, sans vouloir nier qu'il s'agisse là d'une évidence, on peut se demander comment on parvient au juste à connaître une telle évidence? La réponse à cette question, comme le reconnaît Singer lui-même, est qu'on connaît la nature évidente de cette proposition *par intuition* :

« It might be said that the response that I have called “more reasoned” is still based on an intuition, for example the intuition that five deaths are worse than one, or more fundamentally, the intuition that it is a bad thing if a person is killed. But if this is an intuition, it is different from the intuitions to which Haidt and Greene refer. It does not seem to be one that is the outcome of our evolutionary past [...] Thus the

“intuition” that tells us that the death of one person is a lesser tragedy than the death of five is not like the intuitions that tell us we may throw the switch, but not push the stranger off the footbridge. It may be closer to the truth to say that it is a rational intuition, something like the three “ethical axioms” or “intuitive propositions of real clearness and certainty” to which Henry Sidgwick appeals in his defense of utilitarianism in *The Method of Ethics*. » (*ibid.*, 350).

Singer reconnaît donc que des jugements tels que « cinq morts est pire qu'un mort » ou « la mort d'une personne est quelque chose de mauvais » sont ultimement des intuitions, mais il spécifie qu'il s'agit cependant d'intuitions différentes, car elles ne sont pas des produits évolutifs. Mais si elles ne sont pas des produits évolutifs, d'où proviennent-elles? S'agit-il de quelque chose que nous avons simplement appris durant notre enfance? Dans un tel cas, il faut montrer comment on évite le piège du chauvinisme décrit à la section précédente. Une réponse plus probable est qu'il s'agit là d'une évidence qui peut être reconnue par tout individu rationnel qui prend la peine de réfléchir sur de telles questions. C'est là une réponse qui me semble plus satisfaisante, mais c'est néanmoins une erreur de prétendre qu'il s'agisse d'une réponse qui satisfait au problème de l'origine de nos intuitions. Car le type de conclusions auxquelles des individus rationnels qui réfléchissent aboutissent dépend largement du type d'individus qu'ils sont, ce qui renvoie bien sûr à leur histoire évolutive. Qu'est-ce qui nous dit, par exemple, qu'une espèce extra-terrestre intelligente n'aurait pas des intuitions fort différentes des nôtres sur de tels sujets?

Mon point ici n'est pas qu'une proposition du type « cinq morts est pire qu'un mort » n'est pas une vérité objective parce qu'elle dépend de notre constitution particulière d'être humain. Mon point est plutôt qu'indépendamment de la valeur de vérité de telles propositions, on ne peut pas détacher ce qui nous apparaît intuitivement évident de l'histoire évolutive de notre faculté intuitive. Par exemple, si on affirme que la proposition « cinq morts est pire qu'un mort » est une vérité *objective* plutôt qu'une simple vérité *anthropocentrique* (pour reprendre l'expression de David Wiggins (1987)), on doit alors soutenir que des extra-terrestres qui auraient une intuition différente démontreraient simplement les failles de leur propre faculté intuitive, de la même manière qu'ils démontreraient leurs failles s'ils avaient l'intuition que $2+2=6$. Or, avoir (ou prétendre avoir) une faculté qui fonctionne bien est quelque chose qui dépend grandement de notre histoire évolutive.

Bref, lorsque Singer affirme, concernant l'intuition « cinq morts est pire qu'un mort » que, contrairement à l'intuition « il est mal de pousser le gros homme sur le rail pour sauver la vie des cinq individus sur les rails », il s'agit d'une intuition qui n'est pas le résultat de notre passé évolutif, ce qu'il veut réellement dire est qu'il s'agit d'une intuition qui, bien qu'étant un produit évolutif, est

encore appropriée et pertinente aujourd'hui. Ce qu'il veut dire est que si nous avons uniquement des intuitions appropriées, nous n'aurions pas l'intuition qu'il est mal de pousser le gros homme sur les rails. La question importante devient alors : comment sait-on quelle intuition est appropriée et laquelle ne l'est pas? Répondre à une telle question n'est pas évident lorsqu'il s'agit d'intuitions morales, puisqu'on ne peut pas espérer trouver une réponse qui transcenderait nos intuitions, en raison du « is-ought problem » : « From nonethical premises there is no reasonable inference to any ethical conclusion; in other words, no *ought* can be deduced from an *is* » (Gibbard, 2002, 193). Ce qu'il nous faut donc faire est explorer et creuser encore davantage nos intuitions afin de parvenir à une intuition qui semblerait répondre adéquatement au problème que l'on considère, ou encore confronter nos intuitions entre elles en espérant parvenir à un ensemble cohérent. La première approche est typique de l'intuitionnisme éthique, et la seconde est typique de la méthode de l'équilibre réfléchi. J'aborderai ces approches distinctes aux deux chapitres suivants.

Je pourrais poursuivre cette analyse plus longuement, mais je pense avoir apporté suffisamment d'éléments à ce point-ci pour présenter ma réponse au problème original en ce qui attrait aux « théories tacites ». Ma réponse est que les théories tacites sont généralement une source acceptable pour nos intuitions morales. Avant de défendre cette position plus en détails, je dois d'abord présenter la troisième source possible décrite par Cummins, à savoir les théories *explicites*. Je reviendrai en conclusion de chapitre à la question des théories tacites.

3.3.3. Les théories explicites comme sources de nos intuitions

Le problème de l'origine théorique (explicite) de nos intuitions est fort simple : une intuition qui résulterait de l'intériorisation d'une théorie particulière ne peut être utilisée afin de justifier ladite théorie. Par exemple, un philosophe qui est un ardent défenseur du conséquentialisme et qui s'est toujours efforcé de raisonner sur des enjeux moraux en termes conséquentialistes aura fort probablement des intuitions favorables au conséquentialisme lorsqu'il considérera des dilemmes de type « trolley problems ». Or, l'intuition de cet individu ne saurait être vue comme une sorte de *preuve* de la préférabilité de l'option conséquentialiste dans le dilemme. C'est pour cette raison, c'est à dire pour éviter ce type de biais théoriques, que l'on s'intéresse normalement aux intuitions de non-experts en philosophie, ou plus exactement, aux intuitions d'individus non-experts mais tout de même *compétents*. Les intuitions « pré-théoriques » de non-experts compétents sont souvent considérées plus importantes que les intuitions d'experts.

C'est là une idée qui peut paraître étrange, et il est bon de s'y attarder davantage. En effet, s'il semble peu controversé de vouloir rechercher des intuitions ne résultant pas de biais théoriques, soutenir que les intuitions de non-experts soient préférables aux intuitions d'experts ne va pas de soi. Dans d'autres domaines, une telle idée serait tout simplement farfelue! Gopnick et Scheitzgebel donnent l'exemple de la physique théorique et des échecs:

« Twentieth-century physics, for example, has notoriously shown the inaccuracy of our everyday intuitions as hypotheses about the nature of space and time, light, and the microscopic world. The well-cultivated intuitions of a theoretical physicist may be more accurate in her intuitive judgments. [...] As an other example, consider that many of us have essentially no intuitions about the quality of different chess moves – we can only evaluate moves by reasoning about their likely outcomes – while Gary Kasparov surely has quite accurate intuitions in this regard. Our intuitions can develop substantially as our expertise grows » (Gopnick et Scheitzgebel , 1998, 79).

Un premier élément de réponse que l'on peut offrir à Gopnick et Scheitzgebel est que les intuitions du sens commun dans le cas des échecs proviennent d'individus pas ou très peu compétents dans le domaine, alors que les intuitions du sens commun dans le cas de la physique, comme Gopnick et Scheitzgebel le disent eux-mêmes, sont tout simplement erronées, dans la mesure où les découvertes en physique théorique au cours du dernier siècle ont largement démontré que nos intuitions physiques ordinaires ne correspondent pas à la réalité immensément complexe de l'univers physique. Or, de tels constats ne semblent pas s'appliquer aux intuitions philosophiques puisque, contrairement à l'exemple des échecs, les non-philosophes sont généralement capables de comprendre les enjeux soulevés en philosophie, et contrairement à l'exemple de la physique, les intuitions philosophiques n'ont pas été démontrées erronées – à tout le moins dans le domaine normatif.

De telles explications ne sauraient cependant clore le débat. On peut en effet soulever plusieurs autres objections à l'idée que les intuitions de non-experts doivent être jugées préférables aux intuitions d'experts. D'abord, comme il a été dit à la section 3.2, plusieurs soutiennent que la compétence intellectuelle est une denrée malheureusement plus rare qu'on aimerait le croire, comme le démontrent les nombreuses études en économie comportementale décrites à la section 3.1.1. Pour cette raison, l'idée d'accorder plus de poids aux intuitions de « non-experts » paraîtra peu attrayante. On peut cependant répondre qu'un tel argument rate sa cible, dans la mesure où il ne démontre pas que les intuitions d'experts sont préférables aux intuitions de non-experts compétents, mais simplement qu'il n'existe que très peu de non-experts véritablement compétents.

Un autre argument, soulevé par Gopnick et Scheitzgebel (1998) et Cummins (1998), concerne cette idée que les intuitions de non-experts puissent s'avérer complètement erronées, comme le démontre l'exemple de la physique théorique. Non seulement cet exemple montre-t-il la possibilité que nos intuitions pré-théoriques ne reflètent guère la réalité, mais il soulève également l'idée que nos intuitions soient tout simplement *inutiles*. En effet, on ne peut véritablement savoir que nos intuitions sont fiables que si elles ont été étudiées et testées empiriquement; or, si ce qui compte ultimement est la démonstration empirique, en quoi le renvoi aux intuitions est-il autre chose qu'une perte de temps? C'est ce qui amène Cummins à soutenir que nos intuitions sont inutiles, puisque les seules intuitions fiables sont celles qui sont « calibrées », c'est à dire celles qui sont supportées par une théorie scientifique adéquate. Or, dès lors qu'une intuition est calibrée, elle devient par le fait même inutile :

« [E]ven if philosophical intuition *can* be calibrated, it never *is* calibrated, because philosophers could have no possible use for intuition in a context in which the relevant theory was well enough settled to form the basis of a credible calibration test. Philosophical theory in such good shape is ready to bid the Socratic midwife farewell and strike out on its own in some other department. Philosophical intuition, therefore, is epistemologically useless, since it can be calibrated only when it is not needed. Once we are in a position to identify artifacts and errors in intuition, philosophy no longer has any use for it. But if we are *not* in a position to do this, philosophy should not have any faith in it » (Cummins, 1998,118).

Le problème avec cet argument, cependant, est qu'il n'est véritablement menaçant que pour une fraction des intuitions philosophiques, à savoir ces intuitions qui *peuvent* être validées ou réfutées empiriquement. Si cela semble être possible pour les intuitions portant sur la nature des idées, ou sur la conscience, ou autre objet de l'esprit, au moins une grande partie des intuitions qui intéressent les philosophes portent sur des enjeux *normatifs*. Or, il est difficile de voir comment des enquêtes empiriques pourraient parvenir à réfuter des intuitions normatives. Comment, par exemple, des recherches scientifiques pourraient parvenir à la conclusion que quelque chose est moralement bien ou mal, obligatoire ou surrogatoire, admirable ou condamnable? Bref, s'il est clair que la philosophie peut grandement bénéficier de découvertes empiriques de toutes sortes, les questions normatives ne sauraient obtenir de réponses satisfaisantes sur le simple terrain empirique.

Un troisième argument semble cependant mieux atteindre sa cible. Si on retourne à l'explication offerte concernant l'exemple des échecs, on constate que celle-ci n'est pas pleinement satisfaisante, et que l'analogie des échecs soulève une question tout à fait légitime concernant la

pratique philosophique de renvoi aux intuitions de non-experts. L'argument initial consistait à dire que les échecs sont un exemple flagrant de domaine où les intuitions d'experts sont nettement plus valables que les intuitions de non-experts, et la réponse à cet argument était que l'argument ratait sa cible puisque le « sens commun » n'a que pas ou très peu d'intuitions échéphiles, et ne peut donc pas être jugé avoir une véritable *compétence* dans le domaine. Or, cette réponse n'est pas pleinement satisfaisante dans la mesure où le constat initial s'applique même dans les cas de non-experts véritablement compétents; les intuitions échéphiles d'experts demeurent largement préférables aux intuitions de non-experts compétents.

Je pense être un bon exemple de non-expert compétent dans le domaine échéphile. Je joue aux échecs depuis la jeune enfance et j'ai participé à quelques tournois officiels, ayant plutôt bien performé dans certains d'entre eux. Je pense donc pouvoir légitimement soutenir avoir une compétence dans le domaine. Je ne suis cependant en aucun cas un expert, ayant une cote officielle bien modeste. Or, si le modèle de la philosophie s'appliquait aux échecs, mes intuitions échéphiles devraient en principe être jugées plus importantes que celles d'un Grand Maître International comme Gary Kasparov. Une telle idée suscitera, je l'espère, un rire moqueur tout à fait légitime chez tout individu ayant une connaissance au moins élémentaire des échecs. Il ne fait en effet aucun doute que la valeur des intuitions échéphiles est proportionnelle à la qualité (connaissances et talent) du joueur. La question paraît donc légitime : pourquoi n'en irait-il pas de même dans le domaine de la philosophie? Pourquoi doit-on se soucier des intuitions de non-experts plutôt que de s'en tenir aux intuitions d'experts?

La réponse qui me semble la meilleure consiste en une analogie linguistique²⁷. L'idée est que les intuitions du sens commun auraient en philosophie une valeur analogue aux intuitions du sens commun en ce qui concerne la grammaticalité ou l'applicabilité de termes langagiers. Une telle analogie revient à rejeter la pertinence de la distinction entre compétence et expertise, en soutenant que toute personne qui possède une compétence (linguistique, morale ou autre) *normale* devient par le fait même une référence fiable dans le domaine. Si cette analogie est valable, on comprend mieux alors pourquoi on souhaite tester nos intuitions philosophiques sur des non-philosophes, puisque ceux-ci devraient être en mesure d'attester la validité de propositions philosophiques s'ils détiennent une compétence normale. La pertinence de l'analogie semble cependant reposer largement sur l'existence de théories *tacites*, telle une grammaire innée en linguistique, sans quoi les intuitions du sens commun proviendraient essentiellement de l'éducation (linguistique, morale ou autre), et leur

27 Pour une discussion fort intéressante de la pertinence d'une analogie linguistique en philosophie morale, voir les sections 3, 3.1, 3.2 et 3.3 dans Sinnott-Armstrong (2008).

utilisation à des fins de justification reviendrait encore une fois à une forme de chauvinisme (voir section 3.3.1). Par ailleurs, on peut expliquer le contre-exemple des échecs en se référant à l'absence de théories tacites dans ce domaine : si les intuitions échéphiles de non-experts n'ont pas la valeur des intuitions linguistiques ou philosophiques de non-experts, c'est parce que ces premières résultent entièrement de l'apprentissage et de l'approfondissement de règles *arbitraires*, alors que les secondes reposeraient (*ex hypothesi*) sur certains principes innés, et ne seraient en cela pas arbitraires.²⁸

On paraît donc effectuer une boucle en retournant au problème initial de l'existence et de la pertinence de théories tacites. J'avais conclu la section 3.3.2 en promettant de revenir sur cette question. C'est ce que je ferai après avoir d'abord offert une rapide récapitulation du chapitre.

3.4. Conclusion

J'ai analysé ce qui me semble être les trois principales sources de doutes concernant la fiabilité des intuitions. Dans un premier temps, j'ai présenté différentes études scientifiques que certains jugent constituer une menace sérieuse à la fiabilité de nos intuitions, à savoir les études en économie comportementale concernant les évaluations irrationnelles d'individus et les études en science cognitive concernant la typicalité, la structure probabiliste et l'absence de conditions nécessaires et suffisantes derrière nos concepts. J'ai conclu que ces études devaient nous amener à cultiver la prudence dans l'évaluation de nos intuitions, mais qu'elles ne parvenaient guère cependant à discréditer la pratique de renvoi aux intuitions en philosophie. Dans un deuxième temps, j'ai considéré le problème du désaccord intuitif tel qu'étudié empiriquement par le mouvement de philosophie expérimentale. Après avoir soulevé plusieurs doutes quant à la méthodologie des expérimentaux, j'ai conclu que l'existence de désaccords intuitifs ne posait pas véritablement problème s'il existait réellement un « noyau universel » de croyances en éthique, épistémologie ou autre domaine pertinent pour la philosophie. Dans un troisième temps, j'ai abordé le problème de l'origine des intuitions. En suivant Robert Cummins (1998), j'ai distingué trois grandes sources possibles de nos intuitions, à savoir l'éducation, les théories tacites et les théories explicites. La première source est problématique lorsque nos intuitions résultent de l'intériorisation

28 On pourrait répondre que les échecs se fondent sur la logique, et que la compétence logique a vraisemblablement une racine innée. Néanmoins, les intuitions échéphiles sont largement constituées par des règles arbitraires, en particulier la façon dont les pièces bouge, et ces règles ne sont d'aucune façon des « extensions » de principes logiques innés.

de normes sociales *arbitraires*, puisque utiliser de telles intuitions à des fins de justification revient à une forme de chauvinisme. Il importe dès lors de montrer que les normes morales d'où proviennent certaines de nos intuitions morales ne sont pas arbitraires. Or, la meilleure manière de concevoir ce caractère non-arbitraire, à mon sens, est de montrer que ces intuitions morales reflètent des théories tacites innées. Je conclus le présent chapitre par une défense de cette idée.

Si on retourne à la section 3.3.2, j'avais montré que Peter Singer lui-même reconnaissait l'inévitabilité du recours aux intuitions, puisqu'il admettait que des propositions du type « cinq morts est pire qu'un mort » sont connues intuitivement, et qu'on ne peut faire l'économie de propositions de ce type en philosophie morale. Mais d'où proviennent des connaissances du type « cinq morts est pire qu'un mort »? En guise de réponse, on peut dire que tout individu qui réfléchit sérieusement à de telles questions arrivera à des conclusions similaires. Cependant, la proposition « cinq morts est pire qu'un mort » n'est pas *déduite* dans un processus de raisonnement, mais apparaît plutôt évidente en elle-même. On pourrait bien sûr chercher à réduire cette proposition à des propositions plus élémentaires, en disant par exemple qu'elle résulte de la combinaison des deux propositions suivantes : « moins de mal est préférable à plus de mal » + « la mort est (*prima facie*) mal ». Mais comme le reconnaît Singer lui-même, ces propositions plus élémentaires nous apparaissent autant sinon plus évidentes en elle-mêmes que la proposition initiale. Ces propositions élémentaires acquièrent en effet notre approbation spontanée et ne nous paraissent nécessiter aucune justification au-delà de cette approbation spontanée.

Ainsi, dire que la proposition « cinq morts est pire qu'un mort » peut être reconnue par tout individu qui réfléchit sérieusement à ce type de questions ne fait au final que nous amener à la question : pourquoi une telle proposition nous apparaît-elle évidente en elle-même? Pourquoi obtient-elle notre approbation spontanée? Il est difficile de croire que cette proposition nous apparaisse évidente en elle-même simplement parce qu'elle nous a été enseignée en bas âge. S'il est clair que ce type de proposition constitue une norme sociale qui est transmise culturellement, elle n'est pas pour autant arbitraire. On s'attendrait en fait à ce qu'un individu reconnaisse la validité de cette proposition même si on ne la lui avait jamais enseignée.²⁹ Si l'on accepte cette idée, on arrive alors à se demander qu'est-ce qui permet à un individu de constater par lui-même qu'une proposition du type « cinq morts est pire qu'un mort » est évidente en elle-même? À mon sens, seules deux explications sont possibles : ou bien notre faculté intuitive *produit* des propositions évidentes en elles-mêmes, ou bien cette faculté nous permet de *percevoir* ou *découvrir* des propositions

29 Certains relativistes radicaux seraient sûrement en désaccord avec cette idée, mais je préfère ne pas aborder ce type d'objections ici.

évidentes en elles-mêmes. Si la première explication est exacte, ces propositions évidentes proviennent de théories tacites innées et sont ainsi évidentes *pour des gens comme nous*. Si la seconde explication est exacte, ces propositions évidentes proviennent d'une réalité axiologique qui « survient » (« supervenes ») à notre réalité matérielle, et ces propositions sont ainsi évidentes *de manière pleinement objective*, plutôt qu'être de simples vérités anthropocentriques.

Le choix final est donc entre une explication réaliste et une explication innéiste.³⁰ Tel qu'expliqué précédemment, je n'ai pas le luxe d'analyser ici en profondeur la plausibilité du modèle réaliste. Je me permettrai tout de même les remarques suivantes : le modèle réaliste est clairement plus lourd métaphysiquement que le modèle qui renvoie à des théories tacites innées, et c'est pourquoi je penche davantage vers ce second. Je soupçonne également que les deux modèles, en dépit de leur différence métaphysique, permettent ultimement d'offrir des types de réponses identiques face à des enjeux moraux : on peut condamner des atrocités morales et justifier des actes vertueux aussi bien en renvoyant à une réalité axiologique objective qu'à des théories tacites universelles. S'il est indéniable que le modèle réaliste permette d'atteindre une objectivité morale plus complète, la différence *pratique* entre les deux modèles m'apparaît tout simplement nulle : qu'il y ait ou non des vérités morales pleinement objectives au sens réaliste, le vicieux continuera ses actes vicieux et le vertueux, ses actes vertueux. Dans tous les cas, on pourra toujours condamner le vicieux et louer le vertueux. Comme on dit en anglais : *same difference!*

Je termine ce chapitre en répondant directement à la question de départ : les intuitions sont-elles fiables? Pas toujours, mais elles le sont suffisamment pour que l'on puisse travailler avec elles, en gardant cependant toujours une certaine prudence à leur endroit. La possibilité d'utiliser les intuitions (avec prudence) est par ailleurs une bonne nouvelle, puisque les intuitions sont clairement inévitables lorsque l'on souhaite aborder des enjeux normatifs.

Ayant ainsi clarifié la question de la nature (chapitre 2) et de la fiabilité (chapitre 3) des intuitions, je peux à présent explorer la troisième grande question de mon mémoire, à savoir le rôle des intuitions en philosophie morale contemporaine. Je commence au chapitre suivant avec ce que j'appelle la méthodologie standard.

30 On pourrait être tentés de répondre que les propositions morales qui m'intéressent sont simplement des vérités conceptuelles. Cependant, ce type de réponse ne fait à mon sens que repousser le problème, puisqu'on doit alors se demander pourquoi nous avons les concepts et les définitions que nous avons. Par exemple, dire que la proposition « le meurtre est mal » est une vérité conceptuelle nous amène à nous demander pourquoi nous incluons dans la définition de « meurtre » le fait d'être mal? Est-ce purement arbitraire ou *ad hoc*, ou cela reflète-t-il une certaine nécessité, et de quel ordre serait cette nécessité? Est-ce un état du monde (i.e., un fait moral) qui nous contraint à définir ainsi le concept de meurtre, ou est-ce plutôt une contrainte qui renvoie à notre constitution d'être humain? On en revient ainsi aux explications réalistes ou innées.

4. MÉTHODOLOGIE STANDARD : EXPÉRIENCES DE PENSÉE ET ÉQUILIBRE RÉFLÉCHI

La méthodologie standard en philosophie morale contemporaine³¹ consiste, outre en l'analyse critique d'arguments, en la combinaison des deux éléments suivants : (1) des expériences de pensée variant en sophistication et complexité, et (2) la méthode de l'équilibre réfléchi, qui vise à ordonner et à systématiser nos intuitions avec des principes moraux plus abstraits. Les deux éléments sont étroitement associés, dans la mesure où l'on cherche d'abord à identifier nos intuitions sur un sujet en ayant recours à des expériences de pensée et que l'on se sert ensuite de la méthode de l'équilibre réfléchi afin de systématiser les intuitions dévoilées par les expériences de pensée

Il s'agit bien sûr d'une démarche qui a son lot d'opposants, et la première critique qui est généralement soulevée est que cette démarche ne peut être satisfaisante pour la simple raison que nos intuitions ne sont pas fiables. Or, c'est là une objection que je ne saurais considérer ici, puisque j'ai consacré tout le chapitre 3 à la question de la fiabilité des intuitions, ayant conclu que les intuitions *peuvent* être fiables, bien qu'il faille cultiver une certaine prudence à leur endroit. Je me concentrerai donc dans ce chapitre sur les autres critiques soulevées à l'endroit de la méthodologie contemporaine.

Le chapitre sera divisé assez simplement entre chacun des deux éléments de la méthodologie standard, en commençant par les expériences de pensée (4.1), et en continuant avec la méthode de l'équilibre réfléchi (4.2). Puisque plusieurs des questions concernant les expériences de pensée, notamment des questions sur leur nature et leur valeur épistémique, sont similaires aux questions concernant les intuitions, et que ces questions ont largement été traitées aux chapitre 2 et 3, mon analyse des expériences de pensée sera plus brève que celle de la méthode de l'équilibre réfléchi. Je conclurai en expliquant pourquoi la méthode de l'équilibre réfléchi *large* semble être incontournable en philosophie morale (4.3).

4.1. Expériences de pensée

Les expériences de pensée sont des exercices intellectuels qui consistent en l'élaboration de scénarios variant en complexité et en sophistication dont le but est de susciter des jugements

31 Il est possible que cette méthodologie soit standard non seulement pour la philosophie morale, mais pour la quasi-totalité des champs philosophiques. C'est là une thèse que je ne peux cependant démontrer, et je n'ai guère à le faire puisque je souhaite principalement me concentrer sur le domaine moral.

spontanés (intuitions) servant à tester des principes ou théories scientifiques ou philosophiques. Bien que plusieurs taxonomies existent, celle proposée par Karl Popper (1959) paraît la plus efficace et répandue : Popper distingue entre des expériences de pensée critiques (utilisées *contre* une théorie), apologétiques (utilisées *en faveur* d'une théorie) et heuristiques (utilisées *pour illustrer* une théorie). Les expériences de pensée varient également dans leur degré de réalisme; si plusieurs se présentent comme de simples exemples tirés de la vie de tous les jours, d'autres sont totalement invraisemblables et extravagantes.

Avant de me pencher sur le problème de la nature et de la valeur épistémique des expériences de pensée (4.1.2), je donne quelques exemples classiques d'expériences de pensée philosophiques à la section suivante.

4.1.1. Les grands classiques

Le recours aux expériences de pensée se retrouve dans pratiquement toutes les branches de la philosophie.³² On en retrouve des exemples nombreux en éthique, épistémologie, métaphysique, esthétique, philosophie du langage et philosophie de l'esprit. Bien que mon mémoire porte principalement sur les intuitions en philosophie morale, il est bon de prendre une perspective plus large et de présenter quelques exemples classiques tirés de différentes branches de la philosophie.

Le grand classique en philosophie morale est une famille d'expériences de pensée mettant en vedette un train à la dérive menaçant de tuer des individus – que l'on qu'on appelle les « trolley problems ». Initialement développés par Philippa Foot (Foot, 1978), ces scénarios imaginaires sont généralement utilisés afin de montrer que la théorie conséquentialiste émet des recommandations morales contraires à nos intuitions ou convictions profondes. Ces scénarios ont été repris et modifiés non seulement par des philosophes, mais également par plusieurs psychologues et autres scientifiques tels Joshua Greene (Greene et al., 2001) et Marc Hauser (Hauser, 2007). L'expérience de pensée initialement présentée par Philippa Foot était décrite de la manière suivante :

« A trolley is running out of control down a track. In its path are five people who have been tied to the

32 J'utilise l'adverbe « pratiquement » parce que certains affirment que les branches les plus étroitement rapprochées des sciences empiriques, par exemple la philosophie de la physique, ont complètement délaissé le recours aux intuitions à des fins de justification. Robert Cummins, un anti-intuitionniste, affirme par exemple : « Philosophers of physics interested in space and time do not consult their intuitions any more, they ask how we must understand space and time if the physical theories that appeal to them are to be true and explanatory. I do not know if something analogous will replace intuition in every branch of philosophy, but *something* had better replace it. It cannot support any conclusion worth drawing. » (Cummins, 1998, 126).

track by a mad philosopher. Fortunately, you could flip a switch, which will lead the trolley down a different track to safety. Unfortunately, there is a single person tied to that track. Should you flip the switch or do nothing? » (Foot, 1978, 20).

Une expérience de pensée sans doute encore plus célèbre que celle de Foot est celle développée par Hilary Putnam concernant les habitants d'une fictive « Terre jumelle » (*twin Earth*). Introduite dans « Meaning and Reference » (Putnam, 1975), l'expérience de pensée est utilisée par Putnam afin de défendre une théorie sémantique « externaliste », théorie selon laquelle – dans les mots de Putnam – « meanings just ain't in the head! » (*ibid.*, 227). En nous amenant à imaginer une planète identique à tous les niveaux à la Terre, à la seule exception que ce qu'on appelle de l'eau a sur cette autre planète une composition chimique différente, Putnam cherche à montrer que nos intuitions pré-théoriques concernant la référence et la signification correspondent à la théorie externaliste qu'il défend. Cette expérience de pensée a été reprises par plusieurs scientifiques. Alison Gopnik et Eric Schwitzgebel décrivent l'utilité de cette expérience de pensée pour le monde scientifique :

« The intuitions described by Putnam, Kripke, and others after them are not just exotic philosophical intuitions – Frank Keil (1989) and Susan Gelman (Gelman and Wellman 1991) have shown that they may be shared even by fairly young children, and cognitive psychologists like Murphy and Medin (1985) and Rips (1989) have shown that they are shared by ordinary adults and play an important role in adult action and judgment. This intuition is an important piece of data for empirical psychology; Keil and the rest are treating the Twin Earth intuitions evidentially » (Gopnik et Schwitzgebel 1998).

Enfin, si l'on pouvait imaginer un Top 3 des plus célèbres expériences de pensée cherchant à susciter des intuitions pré-théoriques, Edmund Gettier trônerait sans doute au sommet et déclasserait Putnam et Foot. Les deux fameux contre-exemples présentés par Gettier dans son très court article « Is Justified True Belief Knowledge? » (Gettier, 1963), ont été centraux aux débats en épistémologie contemporaine, en plus d'avoir donné lieu à plusieurs recherches scientifiques, dont certaines menées par des philosophes expérimentaux (notamment Weinberg, Nichols et Stich, 2001). Les deux scénarios développés par Gettier dans l'article visent essentiellement à montrer que contrairement à la définition classique de « connaissance » remontant à Platon, une croyance peut être vraie et justifiée sans pour autant véritablement constituer une « connaissance ». C'est du moins ce que semble révéler notre réaction intuitive face aux deux exemples offerts par Gettier, qui mettent en scène un certain Smith qui détient des croyances par rapport à un certain Jones qui

s'avèrent être vraies et justifiées, mais toujours pour les *mauvaises raisons*. De nombreux philosophes ont cherché à répondre au problème soulevé par Gettier, ce qui a donné naissance à une panoplie de scénarios loufoques à la Jones et Smith :

« Imaginary cases are described, involving Brown and his travels in Spain, Nogot, Havit and their vehicles, Tom Grabit and his kleptomaniac proclivities at the library, gypsy lawyers, Norman the clairvoyant, barn facades in the countryside, and a host of others. In each of the cases described, there is a good deal of agreement about whether, under the described conditions, a subject knows, or is justified in believing, something be the case. Intuitions about these cases are then used to clarify the conditions under which various epistemic notions rightly apply » (Kornblith, 1998, 130).

Des centaines d'expériences de pensée de ce type ont été développées au cours des années en philosophie. On peut ajouter à la liste des exemples célèbres les classiques platoniciens de l'allégorie de la caverne et de l'anneau de Gygès, le violoniste branché à nos reins pendant 9 mois décrit par Judith Thomson pour illustrer l'acceptabilité intuitive du droit à l'avortement (Thomson, 1971), le cas du bébé qui se noie dans une flaque d'eau développé par Peter Singer (Singer, 1972) de même que la position originelle derrière le voile d'ignorance développée par l'économiste John Harsanyi (Harsanyi, 1953) et popularisée par John Rawls (Rawls, 1971). Tous cherchent à susciter des réactions intuitives afin de tester ou justifier une théorie ou principe. Ce type d'exercice intellectuel n'est par ailleurs pas l'adage exclusif de la philosophie ou de la psychologie; on en retrouve également de célèbres exemples dans d'autres sciences, à commencer par la physique théorique, où le fameux « chat de Schrödinger » a été utilisé pour illustrer un paradoxe associé à l'« École de Copenhague » en mécanique quantique et les expériences de pensée développées par Albert Einstein impliquant un ascenseur tombant dans le vide ou un individu voyageant sur un faisceau de lumière.

La grande question qui intéresse les philosophes est de savoir quelle est la valeur épistémique d'expériences de pensée de ce type. Ce type d'exercice intellectuel permet-il d'acquérir des connaissances *nouvelles*? C'est la question centrale qui guide la prochaine section.

4.1.2. Nature et valeur épistémique

On peut distinguer entre trois grandes familles de théories abordant chacune les expériences de pensée de manière fort distincte: les théories rationalistes, les théories empiristes et les théories

sceptiques.³³ Concernant ces dernières, outre Duhem (1913), il n'existe pratiquement aucun sceptique *radical* en ce qui attrait aux expériences de pensée dans le domaine scientifique, c'est à dire qu'aucun individu ne juge les expériences de pensée complètement illégitimes ou inutiles. Cela s'explique sans doute par le fait que certaines sciences, telle la physique théorique, ont grandement progressé grâce aux expériences de pensée et qu'il est donc difficile de complètement rejeter la pertinence de ce type d'exercice. Lorsqu'il s'agit du domaine philosophique plutôt que scientifique cependant, les sceptiques radicaux, bien que rares, existent, plusieurs s'identifiant d'ailleurs au mouvement de philosophie expérimentale (voir section 3.2). Cependant, comme le soulignent Brown et Fehige dans l'encyclopédie en ligne *Stanford Encyclopedia of Philosophy*, la plupart des objections soulevées par ces sceptiques des expériences de pensée en philosophie sont ultimement des objections qui concernent la fiabilité des intuitions. Plutôt que d'aborder à nouveau dans cette section des arguments largement discutés au chapitre précédent, je me concentrerai plutôt ici sur le débat contemporain principal qui oppose la théorie empiriste défendue par John Norton (1996, 2004) et la théorie rationaliste défendue par James Brown (1993, 2011).

Brown et Norton ont en commun l'idée que les expériences de pensée sont tout à fait légitimes et permettent réellement d'acquérir des connaissances dans les domaines philosophiques et scientifiques. Leur rejet commun du scepticisme semble cependant être le seul véritable point de convergence de leurs théories. En effet, même l'idée d'acquisition de connaissances est interprétée de manière fort différente par les deux philosophes : alors que Brown soutient qu'une expérience de pensée peut permettre, par la voie de notre faculté intuitive, d'acquérir des connaissances *nouvelles* sur la nature du monde, Norton soutient au contraire qu'une expérience de pensée ne fait en réalité qu'utiliser des données déjà connues par l'individu, ces données ayant toutes été acquises empiriquement. Dans la mesure où elles n'apportent aucune donnée nouvelle et ne font que réorganiser ou généraliser nos données actuelles, nos expériences de pensée se comportent, selon Norton, exactement comme des *arguments* :

« [Thought experiments] are merely picturesque arguments and in no way remarkable epistemologically.

33 Je me concentre dans cette section entièrement sur les débats *contemporains* concernant la nature et la valeur épistémique des expériences de pensée. Il est tout de même bon de situer les débats contemporains dans un contexte historique plus large; Brown et Fehige (2011) soutiennent, dans l'encyclopédie en ligne *Stanford Encyclopedia of Philosophy*, qu'il y a quatre grandes phases historiques de débats sur les expériences de pensée : une première phase au 18^e et 19^e siècle, animée principalement par Lichtenberg, Novalis et Orsted; une deuxième phase au début du 20^e siècle, animée surtout par Duhem, Mach et Meinong; une troisième phase dans la deuxième moitié du 20^e siècle, stimulée principalement par les travaux de Kuhn, Popper et Koyré; et enfin, la phase actuelle qui débute vers la fin des années 1980 et qui est grandement centrée autour de l'opposition entre Brown et Norton.

[...] They can do no more than can ordinary thinking with its standard tools of assumption and argument. They open no new channels of access to the physical world. Thought experiments are just what you thought » (Norton, 1996, 333-366).

S'il soutient que les expériences de pensée n'ont rien de remarquable *épistémologiquement*, Norton ajoute cependant qu'elles constituent un outil intellectuel puissant et précieux pouvant donner des résultats remarquables, comme le démontrent les exploits d'Albert Einstein dans ce domaine : « Generations of scientists have stood in awe of Einstein's great achievements and have tried to mimic his scientific voice. And when they speak in his voice, they have learned his language of thought experiments » (*ibid.*, 365-366).

À l'inverse, Brown ne craint guère de souligner qu'il y a une épistémologie véritablement remarquable derrière nos expériences de pensée. Pour Brown, les expériences de pensée mènent à une « saisie intuitive » qui peut permettre d'avoir un accès direct aux lois de la nature :

« This *a priori* knowledge is gained by a kind of perception of the relevant laws of nature which are, it is argued, interpreted realistically. Just as the mathematical mind can grasp (some) abstract sets, so the scientific mind can grasp (some of) the abstract entities which are the laws of nature » (Brown, 1993, 77)

Contrairement à Norton, donc, Brown soutient que certaines expériences de pensée peuvent réellement apporter des données nouvelles et mener à des croyances qui n'étaient d'aucune façon détenues, inconsciemment ou autre, par l'individu. En ce sens, Brown affirme que certaines expériences de pensée ne peuvent être réduites à de simples arguments, et il donne en appui à sa thèse une série d'expériences de pensée tirées des mathématiques et de la physique, affirmant que ces exemples ne peuvent être reconstruits sous la forme d'arguments (*ibid.*) Norton a vu dans cette affirmation de Brown un défi qu'il s'est empressé de relever en reprenant quatre des exemples d'expériences de pensée décrits par Brown afin de montrer que, contrairement à ce que Brown soutient, de telles expériences de pensée *peuvent* réellement être reconstruites sous la forme d'arguments (Norton, 1996).

Il n'est guère nécessaire d'entrer plus en profondeur dans le débat entre Brown et Norton pour voir que l'opposition entre les deux philosophes se ramène ultimement à une opposition entre réalisme et anti-réalisme. En effet, la différence essentielle entre les deux théories n'est pas celle de savoir s'il est possible ou non de réduire des expériences de pensée à des arguments, ce qui est clairement un débat secondaire par rapport au débat principal, qui concerne lui l'existence ou non de

« lois de la nature » conçues au sens platonicien comme des propriétés abstraites accessibles par intuition. Norton critique d'ailleurs fréquemment Brown en raison du caractère platonicien de sa théorie, jugeant qu'une telle théorie souffre d'un important problème de parcimonie ontologique.

« [In Brown's] Platonic thought experiments, the outcome is supposed to arise almost miraculously from an act of Platonic perception quite unlike mundane operation of ordinary inference. If it turns out – as the reconstruction thesis asserts – that the same outcome can be achieved by ordinary means, we may well begin to wonder if we need the mysteries of Platonic perception to explain thought experiments » (*ibid.*, 339).

Brown réplique à cette critique en renvoyant à un principe d'inférence à la meilleure explication : « Whatever we eliminate by employing the principle of parsimony we can easily reintroduce by an inference to the best explanation (see Meixner, 2000). And this is exactly what a Platonist contends his or her Platonism about thought experimenting to be [...] » (Brown et Feige, 2011). Une réplique que l'on pourrait cependant offrir à Brown est que le principe d'inférence à la meilleure explication ne saurait légitimement se *substituer* au principe de parcimonie ontologique, puisque les deux principes ne sont pas opposés mais agissent en réalité conjointement : c'est précisément le manque de parcimonie ontologique qui amène quelqu'un comme Norton à douter que le modèle platonicien constitue la meilleure explication de la nature des expériences de pensée. Je préfère cependant ne pas aborder plus en profondeur le débat sur le réalisme, pour des raisons soulevées au chapitre précédent. Mon but ici est simplement de montrer que le débat contemporain sur les expériences de pensée se réduit en grande partie à un débat sur le réalisme. Norton et Brown ne sont bien sûr pas les seuls philosophes à aborder le problème de la nature et de la valeur épistémique des expériences de pensée, et il existe plusieurs autres options théoriques. Ces théories alternatives ont en commun leur rejet à la fois du platonisme et du scepticisme. Sans vouloir entrer trop dans les détails, je souligne rapidement les principales théories alternatives.³⁴

Certains philosophes défendent un modèle rationaliste qui, comme Brown, souligne le rôle central joué par l'intuition dans les expériences de pensée. Contrairement à Brown cependant, ces philosophes interprètent les intuitions en termes strictement naturalistes (e.g., Gendler, 1998, Brendel, 2004, Fehige, 2009). Du côté empiriste, deux modèles théoriques importants sont le « constructivisme conceptuel » (e.g., Kuhn 1964) et « l'expérimentalisme » (Mach 1897, Sorensen

³⁴ L'information concernant ces théories alternatives provient de l'encyclopédie en ligne *Stanford Encyclopedia of Philosophy*. Notons par ailleurs que ces théories portent principalement sur les expériences de pensée *scientifiques*.

1992, Buzzoni 2008). Le premier modèle est largement construit sur la base de la théorie développée par Kuhn en philosophie des sciences, et il conçoit ainsi le rôle des expériences de pensée de manière analogue au rôle des révolutions scientifiques :

« [A] well-conceived thought experiment can bring on a crisis or at least create an anomaly in the reigning theory and so contribute to paradigm change. Thought experiments can teach us something new about the world, even though we have no new empirical data, by helping us to re-conceptualize the world in a better way. » (Brown et Feige, 2011).

Quant à l'expérimentalisme, comme son nom laisse deviner, il s'agit d'un modèle qui conçoit les expériences de pensée comme étant de véritables expériences, ou plus exactement une sous-classe d'expériences scientifiques. Buzzoni (2008) défend l'idée d'une « unité dialectique » d'expériences de pensée et d'expériences scientifiques réelles, affirmant que les deux éléments sont indissociables à l'entreprise scientifique : « Without thought experiments there wouldn't be real-world experiments because we would not know how to put questions to nature; without real-world experiments there wouldn't be answers to these questions » (Brown et Feige, 2011).³⁵

Je pourrais analyser plus en détails ces différents modèles théoriques afin de voir lequel paraît le plus plausible afin de rendre compte de la nature des expériences de pensée, mais je préfère consacrer davantage d'attention dans ce chapitre à l'autre élément de la méthodologie standard, à savoir la méthode de l'équilibre réfléchi. Ce choix s'explique essentiellement par le fait que la méthode de l'équilibre réfléchi soulève davantage d'enjeux concernant la *pratique* philosophique que ne le fait le débat sur la nature des expériences de pensée, et c'est pourquoi je souhaite me concentrer davantage à cette première. Je me tourne donc immédiatement vers la méthode de l'équilibre réfléchi et les différents enjeux qu'elle soulève.

4.2. Équilibre réfléchi

On associe généralement la méthode de l'équilibre réfléchi à John Rawls, qui, dans *A Theory of Justice* (1971), introduit par la première fois le concept d'équilibre réfléchi et en décrit les principales caractéristiques.³⁶ L'association de cette méthode à Rawls est légitime, d'autant plus que,

35 Contrairement aux autres modèles théoriques, l'expérimentalisme s'applique mal aux expériences de pensée *philosophiques*, la théorie étant développée dans une perspective scientifique.

36 Le passage où le terme est introduit pour la première fois se trouve à la page 20 de *A Theory of Justice*: « This state of affairs I refer to as reflective equilibrium. It is an equilibrium because at last our principles and judgments

vingt ans avant la publication de *A Theory of Justice*, Rawls délimitait déjà les contours de ce qui allait devenir la méthode de l'équilibre réfléchi dans un article paru dans *Philosophical Review* intitulé « Outline of a Procedure for Ethics » (Rawls, 1951).

Bien qu'il soit exact de prêter la paternité de la méthode de l'équilibre réfléchi principalement à Rawls, d'autres avaient cependant tracé la voie avant lui. Notons d'abord que les idées méthodologiques de Rawls étaient influencées par un autre philosophe contemporain, Henry Nelson Goodman, qui défendait déjà un procédé d'ajustement mutuel entre principes et jugements pour des inférences déductives et inductives.³⁷ Goodman décrit notamment sa méthode dans un passage souvent cité de son ouvrage *Fact, Fiction, and Forecast* (1955) :

« This looks flagrantly circular. I have said that deductive inferences are justified by their conformity to valid general rules, and that general rules are justified by their conformity to valid inferences. But this circle is a virtuous one. A rule is amended if it yields an inference we are unwilling to accept; an inference is rejected if it violates a rule we are unwilling to amend. The process of justification is the delicate one of making mutual adjustments between rules and accepted inferences; and in the agreement thus achieved lies the only justification needed for either. » (Goodman 1955, 66-67).

On retrouvait également des échos de ce qui allait devenir la méthode de l'équilibre réfléchi bien avant Rawls et Goodman. On retrouve par exemple dans certains passages de W. D. Ross la défense d'une méthode qui se rapproche en partie de l'équilibre réfléchi, notamment dans le passage suivant tiré de *Foundations of Ethics* (1939) :

« We must start with the opinions that are crystallized in ordinary language and ordinary ways of thinking, and our attempt must be to make these thoughts, little by little, more definite and distinct, and by comparing one with another to discover at what points each opinion must be purged of excess and misstatement till it becomes harmonious with other opinions which have been purified in the same way » (Ross, 1939, 3).

Rawls voyait pour sa part une parenté entre sa méthode et celle utilisée par Henry Sidgwick dans *The Methods of Ethics* (1874), une parenté que Peter Singer a cependant vertement critiqué dans « Sidgwick and Reflective Equilibrium » (Singer, 1974), affirmant notamment que Sidgwick, contrairement à Rawls, accordait peu d'importance aux jugements bien pesés. Enfin, certains sont

coincide; and it is reflective since we know to what principles our judgments conform and the premises of their derivation. » (Rawls, 1971, 20)

37 Rawls fait référence à la méthode de Goodman dans une note de bas de page, à la page 20 de *A Theory of Justice*.

allés jusqu'à remonter à Aristote afin de trouver la véritable source de la méthode de l'équilibre réfléchi. C'est le cas de Michael DePaul et William Ramsey, qui conçoivent la méthode rawlsienne comme l'instance la plus raffinée de ce qu'ils nomment l'« approche aristotélicienne » :

« [Aristotle] sought an account that agrees with our intuitive judgments *for the most part*, and, at least ideally, provided some understanding of why we make the intuitive judgments that disagree with the account. Many analytic philosophers employ essentially the same method. They "test" their theories against the "data" of what we say about real and hypothetical examples, that is against, out intuitive judgments. But they hold intuitive judgments to be revisable, both in the face of apparently contradictory intuitive judgments about analogous cases and in the face of theories that are well established or intuitively plausible on their own. » (DePaul et Ramsey, 1998, Viii).

Bref, s'il est vrai que la paternité de la méthode revient en grande partie à Rawls, d'autres avaient partiellement tracé la voie avant lui. La méthode a également continué de se raffiner *après* lui, ayant notamment donné naissance à la méthode de l'équilibre réfléchi *large*, popularisée principalement par Norman Daniels, de même qu'à différents autres modèles cherchant à perfectionner la méthode rawlsienne de diverses manières. Je présente dans cette section ces principaux modèles et leurs faiblesses, à commencer par la méthode standard développée par Rawls (4.2.1), puis la méthode de l'équilibre réfléchi *large* développée par Daniels (4.2.2) et enfin d'autres modèles développés à partir de la méthode standard (4.2.3).

4.2.1. La méthode rawlsienne

La méthode de l'équilibre réfléchi n'est qu'un parmi de nombreux éléments que déploie Rawls dans *A Theory of Justice* afin de fonder les principes de la justice comme équité. L'impressionnant appareillage conceptuel de Rawls inclut notamment l'idée de contrat social, de position originelle, de voile d'ignorance, de société bien ordonnée, de structure de base de la société et de personnes libres et égales. Avec la publication de *Political Liberalism* (1993) viendront s'ajouter les idées de consensus par recoupement, de justification publique, de raison publique, de doctrines englobantes et de pluralisme raisonnable. Autant que cela est possible, je souhaite ici détacher la méthode de l'équilibre réfléchi du réseau conceptuel rawlsien afin de l'aborder sous un angle plus restreint, à savoir comme une simple méthode de justification de théories et principes moraux. S'il est vrai que l'équilibre réfléchi a une fonction bien précise à l'intérieur du réseau conceptuel rawlsien, on peut néanmoins l'extraire de ce réseau sans en affecter l'essence.

On peut distinguer trois grandes étapes de la méthode rawlsienne. On débute d'abord par la prise en considération de nos intuitions ou jugements initiaux concernant l'enjeu abordé, à partir desquels on cherche à identifier une classe de jugements que Rawls nomme les jugements « bien pesés ». On parvient à identifier nos jugements bien pesés en se débarrassant simplement de tous les jugements dans lesquels nous n'avons pas pleinement confiance ou que nous soupçonnons d'être affectés par des conditions défavorables :

« For example, we can discard those judgments made with hesitation, or in which we have little confidence. Similarly those given when we are upset or frightened, or when we stand to gain one way or the other can be left aside. All these judgments are likely to be erroneous or to be influenced by an excessive attention to our own interests. Considered judgments are simply those rendered under conditions favorable to the exercise of the sense of justice, and therefore in circumstances where the more common excuses and explanations for making a mistake do not obtain » (Rawls, 1971, 47-8).

Une fois cette épuration préliminaire complétée, on passe à la deuxième étape de la démarche, qui consiste à identifier à partir de nos jugements bien pesés certains principes moraux plus abstraits ou généraux. Rawls propose ici une analogie avec la démarche utilisée en linguistique, où l'on cherche à identifier des règles et principes grammaticaux à partir du langage ordinaire. S'inspirant de Noam Chomsky (1965), Rawls conçoit l'existence d'une compétence morale analogue à la compétence grammaticale qui permettrait aux individus normalement constitués de formuler des jugements moraux à partir de principes inconscients qui sous-tendraient leurs jugements particuliers. Ainsi, après avoir identifié nos jugements bien pesés, notre tâche consiste à trouver les principes moraux plus généraux qui se cachent derrière nos jugements. Ces principes caractérisent notre capacité morale, de la même manière que les principes grammaticaux caractérisent notre capacité linguistique :

« Now one may think of moral philosophy at first [...] as the attempt to describe our moral capacity. [...] This enterprise is very difficult. For by such a description is not meant simply a list of the [moral] judgments [...] that we are prepared to render, accompanied with supporting reasons when these are offered. Rather, what is required is a formulation of a set of principles which, when conjoined to our beliefs and knowledge of the circumstances, would lead us to make these judgments with their supporting reasons were we to apply these principles conscientiously and intelligently. [...] A useful comparison here is with the problem of describing the sense of grammaticalness that we have for the sentences of our native language. In this case the aim is to characterize the ability to recognize well-formed sentences by formulating clearly expressed principles which make the same discrimination as the native speaker. [...]

A similar situation presumably holds in moral philosophy. [...] A correct account of moral capacities will certainly involve principles and theoretical constructions which go much beyond the norms and standards cited in everyday life [...] » (Rawls, 1971, 46-47).³⁸

Après avoir séparé le bon grain de l'ivraie dans nos jugements et intuitions et après avoir identifié les principes moraux plus abstraits qui sous-tendent ces jugements et intuitions, on peut passer à la troisième étape qui consiste en un processus d'ajustements et de révisions entre jugements particuliers et principes abstraits visant à atteindre un équilibre harmonieux. C'est cette troisième étape qui est la plus étroitement associée à la méthode de l'équilibre réfléchi. Dans ce processus d'ajustements et de révisions, ni les jugements ni les principes ne se voient accorder une quelconque primauté ou supériorité du simple fait d'être un jugement ou un principe. Un jugement particulier peut nous amener à reconsidérer l'ensemble d'un système moral si ce jugement nous apparaît suffisamment important; à l'inverse, un principe peut nous amener à rejeter certains jugements particuliers qui ne s'y conforment pas si l'on juge ce principe suffisamment important. L'objectif est donc d'évaluer le poids relatif de chaque principe et jugement dans l'atteinte d'un tout harmonieux, comme le résume bien Michael DePaul : « [N]either particular judgment nor theoretical principle is always favored. Whenever conflicts emerge, the philosopher must reflect on the connection among her beliefs and determine what to revise on the basis of what, all things considered, seems to her most likely to be correct » (DePaul, 1998, 295).

Par ce long processus de révisions et d'ajustements, on espère parvenir à atteindre un état où nos différents jugements et principes sont en harmonie les uns avec les autres, c'est à dire où aucun d'entre eux n'entre en contradiction avec un autre. Lorsque notre maison intellectuelle est finalement en ordre, pour reprendre l'expression de Richard Foley,³⁹ on atteint alors un premier équilibre réfléchi. Cet équilibre n'est cependant que provisoire et peut toujours être modifié si l'on arrive à considérer de nouveaux principes ou jugements qui ne nous étaient pas accessibles auparavant. Idéalement, une personne devrait chercher à étendre ses connaissances afin de prendre en considération le plus grand nombre de principes et de théories morales possible. Plus le nombre et l'importance des théories et principes moraux considérés seront élevés, plus notre équilibre réfléchi

38 J'ai remplacé ici les endroits où Rawls parle de « justice » par des termes référant à la moralité. Ce dont Rawls discute dans ce passage s'applique d'abord et avant tout aux capacités morales et aux principes moraux qui la caractérisent, et il choisit ensuite de recentrer son enquête sur les questions de justice qui l'intéresse plus spécifiquement. Le passage suivant illustre ce point : « Now one may think of moral philosophy at first [...] as the attempt to describe our moral capacity; or, in the present case, one may regard a theory of justice as describing our sense of justice » (*ibid.*, 46)

39 « Putting our intellectual house in order » (Foley, 1998, 246)

gagnera en valeur.

Dans la mesure où notre équilibre réfléchi est toujours sujet à révision à la suite de nouvelles connaissances, on n'atteint jamais avec la méthode rawlsienne un état de certitude pouvant nous amener à prétendre posséder *la* bonne théorie morale ou théorie de la justice. La méthode rawlsienne est ainsi clairement faillibiliste. C'est là le principal élément qui distingue sa méthode de l'approche intuitionniste, qui elle conçoit certains principes abstraits fondamentaux comme étant évidents en eux-mêmes et cherche à développer un système éthique à partir de ces principes évidents.⁴⁰ À l'opposé de cette méthode, Rawls soutient que même les jugements ou principes dans lesquels nous avons le plus confiance et qui sont harmonie avec nos autres principes et jugements peuvent s'avérer erronés et sujets à modification.

Certains ont reproché à Rawls ce caractère faillibiliste de sa méthode, soulignant par exemple que certains principes moraux élémentaires – par exemple le principe « il est mal d'imposer gratuitement de la souffrance à autrui » – ne paraissent aucunement révisibles. La réplique de Rawls et des ses partisans a principalement consisté à souligner que la méthode de l'équilibre réfléchi n'était guère différente de la démarche scientifique traditionnelle. En science, toute théorie ou explication est sujette à révision, et plusieurs croyances qui étaient autrefois jugées évidentes – par exemple la croyance que la terre est plate – ont fini par être rejetées suite à des observations nouvelles. On voit mal ainsi pourquoi le faillibilisme de la méthode rawlsienne devrait poser plus de problèmes qu'il n'en pose en sciences.

Une autre critique importante qui est adressée à la méthode est qu'en systématisant les jugements et principes comme elle le propose, elle parvient seulement à *décrire* la conception morale propre à un individu et ne permet guère de démontrer si cette conception morale est ou non *acceptable*. La méthode rawlsienne, ses critiques soutiennent, est un exercice en *anthropologie* morale plutôt qu'en *philosophie* morale. Certains vont même jusqu'à soutenir qu'une telle méthode n'a au plus qu'un intérêt *biographique*. C'est là un des nombreux reproches de Peter Singer à la méthode rawlsienne :

« A normative ethical theory, however, is not trying to explain our common moral intuitions. It might reject all of them, and still be superior to other normative theories that better matched our moral judgments. For a normative moral theory is not an attempt to answer the question “Why do we think as we do about moral questions?” [I]t is obvious that the question [...] may require a historical, rather than a

40 La section 8 de *A Theory of Justice* est entièrement consacrée au rejet de l'approche intuitionniste. Je reviendrai sur cette critique rawlsienne de l'intuitionnisme au chapitre 5.

philosophical, investigation. » (Singer, 2005, 345).

En réponse à ce type de critiques, Rawls peut placer l'accent sur son impératif d'étendre autant que possible les théories et principes moraux pris en considération dans l'équilibre réfléchi. Puisque l'individu ne peut se contenter de simplement rechercher l'équilibre dans ses jugements et principes actuels, mais qu'il doit également s'éduquer afin de prendre en considération une pluralité de nouvelles théories et principes moraux, il semble erroné d'accuser la méthode rawlsienne d'être un exercice d'anthropologie ou de biographie morale plutôt que de philosophie morale.

Une telle réponse n'est cependant guère suffisante pour défaire la critique soulevée par Singer et autres. Car s'il est vrai que la méthode rawlsienne exhorte les individus à aller au-delà de leurs jugements et principes moraux actuels, la méthode demeure néanmoins une forme d'anthropologie morale dans la mesure où les nouvelles théories morales qui seront explorées par les individus ne pourront ultimement qu'être évaluées à l'aune de leur intuition personnelle. Même en voulant limiter autant que possible le rôle de l'intuition dans la détermination des principes de justice – ce que Rawls décrit comme son objectif dans la section 8 de *A Theory of Justice* – le renvoi à l'intuition est tout simplement inévitable dès lors qu'on aborde le domaine moral, ce que Rawls reconnaît d'ailleurs lui-même : « There is no reason to suppose that we can avoid all appeals to intuition, of whatever kind, or that we should try to. » (Rawls, 1971, 44).

À cela, Rawls pourra répliquer que si le renvoi aux intuitions condamne réellement une méthodologie morale à une forme d'anthropologie, il s'agit alors d'un constat qui s'applique nécessairement à toutes les méthodologies morales, puisque le renvoi aux intuitions est tout simplement inévitable. C'est là un argument que soulève notamment Philip Stratton-Lake :

« Reference to what we really think would, of course, be inappropriate in sciences, for here we have a much more reliable method of discovering non-normative truths about the world. But we cannot establish moral truths by empirical research, or experiment [...]. The only way of establishing moral truth is by reflection on what we really think once the relevant concepts and principles have been clarified, distinguished from each other, and systematized. » (Stratton-Lake, 2002, 26).

Il existe cependant une autre manière de tenter de répondre à la critique soulevée par Singer, qui consiste non pas à souligner l'inévitabilité de ce qu'on peut appeler le caractère anthropologique de la méthodologie morale, mais plutôt à proposer une alternative à cette anthropologie. C'est ce que fait Norman Daniels en développant la méthode de l'équilibre réfléchi *large*. J'aborde cette

méthode à la section suivante.

4.2.2. L'équilibre réfléchi large

Norman Daniels reconnaît l'importance de la critique soulevée (entre autres) par Peter Singer. Daniels admet que le fait de limiter la méthode de l'équilibre réfléchi à la recherche d'un équilibre entre principes et jugements ne peut au final que mener à une forme d'anthropologie morale, dans la mesure où l'équilibre entre principes et jugements vient *décrire* mais non *justifier* la conception morale détenue par l'individu :

« In moral philosophy we are interested in understanding how to justify moral principles and choose between competing sets of principles. When the linguist produces a grammar which seems to characterize a syntactic competency [...], his descriptive task, at least, is done. Moreover, he has no further *justificatory* and *prescriptive* task of trying to show why one descriptively [...] adequate grammar is to be preferred over another. Narrow equilibrium leaves us in the position of the descriptive linguist [...] » (Daniels, 1980a, 25).

Afin d'éviter que la méthode rawlsienne ne se réduise à une simple description de la vision morale propre à chaque individu, Daniels propose une interprétation particulière de Rawls⁴¹ qui l'amène à distinguer entre deux types d'équilibre réfléchi : l'équilibre *étroit* et l'équilibre *large*. Selon Daniels, l'équilibre réfléchi étroit n'est qu'une version sophistiquée de l'intuitionnisme éthique et ne permet guère de justifier une théorie morale. La différence principale entre les deux méthodes est la suivante : alors que l'équilibre étroit prend en considération deux éléments, à savoir les jugements et principes moraux, l'équilibre large, lui, prend en considération trois éléments : (a) les jugements moraux, (b) les principes moraux et (c) les *théories d'arrière-fond* (« background theories »). La méthode défendue par Daniels préserve ainsi les trois étapes de la méthode rawlsienne décrite à la

41 Il n'est pas clair que Rawls endosse réellement ce que Daniels conçoit comme équilibre large. S'il est vrai que Rawls a lui-même adopté cette terminologie dans des écrits ultérieurs, l'idée que Rawls avait de la distinction entre l'équilibre étroit et large semblait différer de celle de Daniels. Pour Rawls, l'équilibre devient large non pas lorsqu'il inclut les théories non-morales (ou « théories d'arrière-fond »), mais plutôt lorsqu'il s'étend au plus grand nombre de théories et conceptions morales possible. Cette idée, qui était déjà présente dans *A Theory of Justice* (p.49), est reprise par Rawls en fin de vie dans *Justice as Fairness, A Restatement* (Rawls, 2001) : « This suggests that we regard as wide [...] that reflective equilibrium reached when someone has carefully considered alternative conceptions of justice and the force of various arguments for them. More exactly, this person has considered the leading conceptions of political justice found in our philosophical tradition [...]... and has weighed the force of the different philosophical and other reasons for them » (*ibid*, 31). Daniels répondrait cependant que la référence de Rawls au besoin d'analyser les raisons et arguments philosophiques des différentes conceptions de la justice est précisément ce qui crée le lien avec les théories d'arrière-fond.

section précédente – l'identification des jugements bien pesés, l'identification des principes sous-jacents et les ajustements et révisions entre principes et jugements bien pesés – mais ajoute une quatrième étape, qui consiste en l'identification et l'évaluation des théories d'arrière-fond propres aux différentes conceptions morales considérées.

« We do *not* simply settle for the best fit of principles with judgments, however, which would give us only a *narrow* equilibrium. Instead, we advance philosophical arguments intended to bring out the relative strengths and weaknesses of the alternative sets of principles (or competing moral conceptions). These arguments can be construed as inferences from some set of relevant background theories. [...] We can imagine the agent working back and forth, making adjustments to his considered judgments, his moral principles, and his background theories. In this way he arrives at an equilibrium point that consists of the ordered triple (a), (b), (c). » (*ibid.*, 1979, 258-259).

Les théories d'arrière-fond dont il est question sont par exemple une théorie de la personne, une théorie du choix rationnel, une théorie du rôle de la moralité en société, de même qu'un ensemble de différentes théories sociales. Bien que cet exemple ne soit pas considéré par Daniels lui-même, on peut également concevoir certaines théories métaphysiques, religieuses ou existentielles (concernant le sens de la vie) comme constituant des théories d'arrière-fond pouvant affecter notre conception morale.⁴² Daniels souligne que les théories d'arrière-fond peuvent aussi bien être des théories morales que des théories non-morales. Il s'empresse par ailleurs de préciser que bien que ces théories soient dites « d'arrière-fond », elles ne doivent pas être comprises de manière fondationnaliste. En effet, ces théories ne sont pas un fondement pour nos principes et jugements, tel que l'identification des bonnes théories d'arrière-fond permettrait de justifier nos principes et jugements qui reposent sur elles. Daniels conçoit plutôt ces théories d'arrière-fond comme étant un élément supplémentaire que l'on doit placer en équilibre avec les deux autres éléments; si les théories d'arrière-fond peuvent nous amener à réviser nos principes et jugements qui s'y opposent, l'inverse est aussi vrai : nos théories d'arrière-fond peuvent être révisées si elles entrent en conflit avec certains jugements ou principes qui nous paraissent suffisamment importants. Bref, aucun élément ne sert de fondement dans la démarche décrite par Daniels :

« The acceptability of deep theories in (c) can force the revision of initial considered judgments in (a). Considered moral judgments [...] [can] constrain the acceptability of theories in (c) [and] are also

42 Par exemple, pour plusieurs individus, des impératifs moraux reposent ultimement sur une volonté divine; s'il s'avérait qu'il n'existait aucune volonté divine, un tel fait devrait certainement amener ces individus à reconsidérer leur conception morale.

revisable in light of many considerations that affect the acceptability of the theories in (c) and their equilibrium with principles in (b) and other judgments in (a). No considered judgments are foundational in any sense; all are subjected to great pressures for revision and are accordingly revised » (*ibid.*, 1980a, 26).

On peut se demander quel est l'avantage de la méthode de l'équilibre réfléchi large par rapport à l'alternative étroite si la première ne fait qu'ajouter un ingrédient supplémentaire qui doit être placé en harmonie avec les autres ingrédients. Plus spécifiquement, en quoi l'ajout des théories d'arrière-fond permet-elle de répondre à la critique soulevée par Singer selon laquelle la méthode de l'équilibre réfléchi ne fait que *décrire* notre conception morale et constitue en cela une forme d'anthropologie morale? La réponse se trouve essentiellement dans le fait que la méthode décrite par Daniels incorpore des éléments *non-moraux*, qui sont eux capables de justification ne dépendant pas de l'intuition individuelle. Par exemple, une conception morale qui incorporerait des croyances religieuses erronées ou des préjugés concernant certains groupes jugés inférieurs pourrait mener à une réfutation sur une base factuelle. On voit ainsi comment la méthode de l'équilibre réfléchi large offre également une réponse au problème du désaccord moral : entre deux équilibres réfléchis antagonistes, il est possible de démontrer que l'un est supérieur à l'autre si l'on parvient à identifier des théories d'arrière-fond erronées qui sont incorporées uniquement dans l'un des deux équilibres. C'est là un des avantages de la méthode que souligne Daniels : « A [...] benefit of wide equilibrium is that [...] disagreements about theories may be more tractable than disagreements about moral judgments and principles. Consequently, if the moral disagreements can be traced to disagreements about theory, greater moral agreement may result » (*ibid.*, 1979, 262).

La méthode possède par ailleurs deux autres vertus importantes, à savoir le fait qu'elle ne soit ni trop permissive ni trop conservatrice. Concernant la permissivité, on souhaite évidemment éviter qu'une méthode permette de justifier des conceptions morales extrêmes ou farfelues. Or, comme le souligne Michael DePaul, la méthode de l'équilibre réfléchi large⁴³ constitue un filtre efficace qui permet de ne pas laisser passer des conceptions morales indésirables :

« It is not [...] all that easy to accept any old screwy thing in reflective equilibrium. To do so one must be willing to make suitably screwy adjustments throughout one's entire system of beliefs. So it is not surprising that we do not very often find people who have a coherent but wacky system of beliefs. When we discover someone who holds some bizarre belief, what we usually find is that she also holds, or is

43 Bien que DePaul ne précise pas dans le passage cité qu'il parle de l'équilibre réfléchi *large*, il ne faut aucun doute en lisant l'article que DePaul considère la version large et non étroite.

disposed to accept, other more “normal” beliefs that we can use to persuade her to revise her strange belief. This is to say, what we usually find is that people who hold screwy beliefs have not reached a point of reflective equilibrium [...] » (DePaul, 1998, 297-8).

Si DePaul est encouragé par le fait que la méthode de l'équilibre réfléchi large permette de discréditer des théories farfelues en soutenant qu'elles ne sont tout simplement pas en état d'équilibre réfléchi, Stephen Stich y voit pour sa part une faille importante de la méthode. Stich affirme en effet que cette caractéristique de la méthode fait en sorte qu'il est pratiquement impossible d'offrir de contre-exemples, puisque l'on peut discréditer *n'importe quelle* conception morale en soutenant qu'elle n'a pas atteint un équilibre large : « [C]ounterexamples are hard to come by because it is so hard to show that anything is in WRE⁴⁴ for anyone. (“Would she continue to accept the rule if she thought her epistemological and metaphysical views and came to some stable equilibrium view?” Well, God knows.) » (Stich, 1998, 101).

En réponse à Stich, on peut souligner que la difficulté de démontrer qu'une conception morale est véritablement en équilibre réfléchi large est une conséquence inévitable du caractère faillibiliste de la méthode. Or, malgré ses limites, le faillibilisme a l'avantage de ne jamais rien prendre pour acquis et d'éviter ainsi la sorte de conservatisme moral que dénonce entre autres Peter Singer. D'ailleurs, si Singer se voulait fort critique à l'endroit de la méthode décrite par Rawls, il se montre cependant beaucoup plus sympathique à l'endroit de l'interprétation offerte par Daniels, précisément parce que la méthode décrite par Daniels semble échapper au conservatisme que Singer condamne :

« Admittedly, it is possible to interpret the model of reflective equilibrium so that it takes into account any grounds for objecting to our intuitions, including those that I have put forward. Norman Daniels has argued persuasively for his “wide” interpretation of reflective equilibrium. If the interpretation is truly wide enough to countenance the rejection of all our ordinary moral beliefs, then I have no objection to it. » (Singer, 2005, 347).

Une autre critique que l'on pourrait être tenté d'opposer à la méthode de l'équilibre réfléchi large est qu'elle tente ultimement de faire reposer le moral sur le non-moral – un péché qui est bien sûr décrié par la plupart des philosophes. Ce qui sous-tend une telle critique est une caractéristique de la méthode décrite plus haut, à savoir le fait qu'elle ouvre la porte à une forme de justification factuelle plutôt que simplement intuitive en accordant une place importante aux théories non-

44 Wide Reflective Equilibrium (WRE).

morales. Or, on comprend rapidement pourquoi cette critique rate sa cible, pour la simple raison que la justification factuelle ne vient pas remplacer la justification intuitive mais vient plutôt s'ajouter à elle. Cela nous renvoie bien sûr à la nature cohérentiste de la démarche : les éléments non-moraux pris en considération dans l'équilibre réfléchi large ne constituent pas la source ou le fondement des éléments moraux. Tous les éléments – jugements moraux, principes moraux et théories d'arrière-fond morales ou non-morales – entrent dans une relation d'influences réciproques où aucun élément n'est jugé posséder une quelconque primauté.

Cette explication mène cependant à un autre type de problème : en voulant éviter le problème de la réduction du moral au non-moral, on semble soulever le problème de la *circularité* de la méthode de l'équilibre réfléchi large. En effet, pour éviter que les théories non-morales constituent le fondement justificateur de la méthode, Daniels soutient que ces théories peuvent elles-mêmes être contraintes par des éléments moraux. Or, ces éléments moraux ne peuvent au final que se présenter sous la forme d'intuitions morales – soit des intuitions épurées se présentant comme des « jugements moraux bien pesés » (a), soit des intuitions plus abstraites se présentant comme des « principes moraux » sous-jacents aux jugements bien pesés (b). Bref, pour éviter que nos théories d'arrière-fond (c) jouent un rôle fondationnel, on en arrive au final à affirmer qu'elles sont contraintes par nos jugements bien pesés et nos principes moraux – qui sont pourtant eux-mêmes dits contraints par nos théories d'arrière-fond. On semble se retrouver ici devant une démarche circulaire; pourquoi devrait-on juger ce cercle vertueux plutôt que vicieux?

Daniels répond à cette objection en développant ce qu'il nomme la « contrainte d'indépendance » (« the independence constraint »). Il décrit cette contrainte de la manière suivante :

« The independence constraint [...] for wide equilibrium in general applies in this way : *if there are any considered moral judgments (a') which may constrain [background] theory acceptability, they must to a significant extent be disjoint from the considered moral judgments (a) which constrain [the initial narrow] equilibrium. Notice that both (a) and (a') are in the same wide equilibrium with the principles, but they constrain their acceptability in different ways* » (*ibid.*, 1980b, 89).

Daniels cherche donc à éviter le problème du cercle vicieux en soutenant que les principes ou jugements moraux qui contraignent les théories d'arrière-fond *ne sont pas les mêmes* que ceux qui sont utilisés pour atteindre l'équilibre initial *étroit* entre principes et jugements – qui est atteint lors de la troisième étape de la démarche rawlsienne. En schématisant, on peut dire que Daniels

rejette le modèle de justification 1 ci-bas et endosse le modèle de justification 2 :

Modèle 1 : $c \leftrightarrow (a \leftrightarrow b)$

Modèle 2 : $(a' \leftrightarrow b') \rightarrow c \rightarrow (a \leftrightarrow b)$ ⁴⁵

Les flèches renvoient ici à une relation d'influence ou de contrainte.⁴⁶ Les symboles (a') et (b') renvoient à des jugements et principes moraux différents de ceux qui sont utilisés à l'étape initiale de recherche d'équilibre étroit, étape qui est décrite par la relation $(a \leftrightarrow b)$. On voit que le modèle 1 suppose une circularité *complète* entre jugements, principes et théories, alors que le modèle 2 évite une telle circularité complète en introduisant une distinction entre deux types de jugements et de principes.

Le modèle 2 n'est bien sûr pas sans problèmes. Le problème le plus évident est de savoir s'il y a ou non des éléments qui contraignent (a') et (b'). La contrainte réciproque qu'ils exercent entre eux (\leftrightarrow) suffit-elle réellement? Cela serait surprenant, considérant qu'une telle contrainte réciproque n'était pas jugée suffisante dans l'équilibre initial entre (a) et (b). Daniels, tout comme Rawls, mentionne par ailleurs à plusieurs reprises qu'aucun élément n'est exempt de modification éventuelle et que tous sont révisables. On est alors en droit de se demander quel élément pourrait amener une modification à (a') et (b'). S'agit-il d'une autre théorie d'arrière-fond (c')? S'agit-il d'autres jugements ou principes moraux [(a'') et (b'')]? Peu importe la réponse, on ne parviendra guère à échapper au trilemme de Münchhausen, selon lequel la justification doit au final mener soit à une régression à l'infini, soit à une circularité, soit à une dépendance sur des propositions non-justifiées.

La branche du trilemme favorisée par Daniels et Rawls⁴⁷ est clairement la deuxième, à savoir celle de la justification circulaire. D'autres ont cependant préféré interpréter la méthode de

45 Daniels ne mentionne que les jugements bien pesés (a') comme pouvant contraindre les théories d'arrière-fond, mais on peut supposer que les principes moraux (b') contraignent également les théories d'arrière-fond, dans la mesure où ces principes moraux ne sont qu'une abstraction de jugements bien pesés.

46 On pourrait penser que les flèches renvoient plutôt à une relation de *justification* entre les éléments. Cependant, dans l'esprit de Daniels (et de Rawls), la justification ne se situe pas au niveau des relations particulières qui unissent les éléments, mais plutôt au niveau de la cohérence d'*ensemble*. À cet égard, voir la citation de Rawls à la note suivante. Merci à Christine Tappolet d'avoir soulevé la possibilité que ces flèches renvoient à une relation de justification.

47 Plusieurs passages de *A Theory of Justice* confirme la démarche cohérentiste de Rawls, à commencer par le passage suivant : « I have not proceeded then as if first principles, or conditions thereon, or definitions either, have special features that permit them a peculiar place in justifying a moral doctrine. They are central elements and devices of theory, but justification rests upon the entire conception and how it fits in with and organizes our considered judgments in reflective equilibrium. As we have noted before, justification is a matter of the mutual support of many considerations, of everything fitting together into one coherent view. » (Rawls, 1971, 579)

l'équilibre réfléchi en termes fondationnalistes afin d'éviter le problème de la circularité. En effet, bien que la méthode ait principalement été conçue de manière cohérentiste, elle est selon plusieurs tout à fait compatible avec un modèle de justification fondationnaliste, comme le souligne Christine Tappolet :

« En tant que telle, la méthode de l'équilibre réfléchi peut aisément être adoptée par le fondationnalisme. Rien ne s'oppose, en effet, à ce que l'on accorde un statut épistémique privilégié à certaines des croyances impliquées, comme les croyances pondérées. Dès lors, ces croyances pourraient former le fondement de l'édifice de la justification. Comme les croyances pondérées sont susceptibles d'être abandonnées, il s'agirait évidemment d'un fondement faillible. Toutefois, le fondationnalisme peut aisément s'en contenter. [...] À ce titre, il est instructif de noter que certains fondationnalistes contemporains – on peut penser à Robert Audi, mais aussi à Michael DePaul – font une large place à la méthode de l'équilibre réfléchi. » (Tappolet, 2000, 81)

Outre Audi et DePaul que mentionne Tappolet, Roger Ebertz est un autre philosophe qui défend une interprétation fondationnaliste de la méthode de l'équilibre réfléchi. Dans « Is Reflective Equilibrium A Coherentist Model? » (1993), Ebertz reprend l'idée développée entre autres par DePaul dans « Reflective Equilibrium and Foundationalism » (1986) selon laquelle la méthode de l'équilibre réfléchi peut être associée au fondationnalisme dans la mesure où les deux positions portent sur des enjeux différents, quoique reliés : « Reflective equilibrium is a method of theory construction [while foundationalism is] a type of account of the epistemic status of our beliefs » (DePaul, 1986, 68).⁴⁸ Ebertz va cependant plus loin que DePaul et défend l'idée que la méthode de l'équilibre réfléchi *requière* un modèle de justification fondationnaliste – ou plus exactement, ce qu'il nomme « fondationnalisme modeste », qui soutient essentiellement que les fondations sont faillibles – puisqu'une interprétation cohérentiste rend le modèle peu plausible. Son argument consiste à dire que la cohérence ne suffit pas à elle seule à justifier les éléments en équilibre réfléchi, et qu'il faut que ces éléments soient jugés plausibles en eux-mêmes pour que la justification puisse être conférée. Les éléments jugés plausibles en équilibre réfléchi servent de *fondements provisoires* à toutes les étapes du processus, comme il le souligne dans le passage suivant :

48 Pour une défense de la position contraire, selon laquelle l'équilibre réfléchi est incompatible avec le fondationnalisme, voir notamment « Relativism and Wide Reflective Equilibrium » (Nielsen, 1993), tout spécialement les pages 317 et 318.

« Part of the reflective process is the continual asking of the question 'Do these principles really fit what seems to me to be just? Right? Etc.?' [...] [T]here is a kind of test by which principles can always be tested – the test of whether they fit the considered moral judgments we are committed to *at this point in the reflective process*. If they do not they are not justified. Thus considered moral judgments do play a special justificatory role in reflective equilibrium, even after a radical shift » (Ebertz, 1993, 203-204)

En l'absence de plausibilité indépendante des éléments, donc, le fait que ces éléments soient en équilibre ne permettrait pas d'atteindre une quelconque justification. C'est là une idée qui est également largement défendue par Richard Brandt dans *A Theory of the Good and the Right* (1979), qui soutient qu'aucune justification n'est possible en l'absence du critère suivant : « unless some of the beliefs are initially credible – and not merely initially believed – for some reason other than their coherence, say, because they state facts of observation » (Brandt, 1979, 147)

Je n'entre pas plus en détails dans le débat opposant les interprétations cohérentistes et fondationnalistes de la méthode. Il suffisait à mon propos de mentionner qu'une interprétation fondationnaliste de la méthode peut potentiellement répondre à certaines critiques soulevées à l'endroit de celle-ci, par exemple la critique selon laquelle plus de cohérence ne signifie aucunement plus de plausibilité. J'expliquerai en conclusion de chapitre pourquoi la méthode de l'équilibre réfléchi large, qu'elle soit ou non interprétée en termes fondationnalistes, me semble être une méthode incontournable en philosophie morale.

Avant de conclure, je présente dans la section suivante quelques autres modèles développés à partir de la méthode standard développée par Rawls.

4.2.3. Autres modèles

Deux grandes stratégies ont mené au développement d'autres modèles à partir de la méthode standard développée par Rawls : la restriction des individus dont l'équilibre réfléchi est jugé légitime, et à l'opposé, l'élargissement de la méthode aux sociétés ou même à l'ensemble de l'humanité.

Dans le premier cas, on parle généralement de la méthode de l'équilibre réfléchi *des experts*. Cette version de la méthode a entre autres été défendue par Stich et Nisbett (1980), qui ont soutenu que lorsqu'on dit d'un principe inférentiel qu'il est justifié, ce qu'une telle justification suppose est qu'un tel principe passerait le test de l'équilibre réfléchi (étroit) des individus considérés comme des experts dans les domaines pertinents. Stich s'est cependant distancé de cette méthode plus tard,

soulignant entre autres qu'une communauté d'experts peut avoir des positions complètement erronées en raison de multiples facteurs :

« [...] Unless experts are picked out in a question-begging way (e.g., those people whose inferential practices are in fact justified), it seems entirely possible for the expert community, under the influence of ideology, recreational chemistry, or evil demons, to end up endorsing some quit nutty set of rules »
(Stich, 1998, 102)

Un autre problème avec la méthode de l'équilibre réfléchi d'experts, qui est entre autres souligné par Conee et Feldman (1983), renvoie à la difficulté d'identifier les véritables experts. Des individus ou communautés peuvent en effet reconnaître comme experts des groupes d'individus totalement différents, et les critères nous permettant de trancher entre ces différents candidats au statut d'experts ne sont pas clairs. En poussant le raisonnement plus loin, on se retrouve devant un véritable paradoxe lorsqu'on réalise que les individus à même d'identifier les critères devant mener à l'identification des experts doivent eux-mêmes détenir une forme d'expertise – ce qui nous mène bien sûr à une régression à l'infini.

Le problème de l'identification des experts paraît encore plus important lorsqu'on se concentre sur le phénomène moral : qui peut-on tenir comme experts dans le domaine de la moralité? L'expertise morale doit-elle être interprétée sur le modèle de l'agent vertueux? Si oui, doit-on considérer davantage les vertus intellectuelles ou les vertus pratiques? Si la première option paraît plus plausible, dans la mesure où l'on recherche non pas des individus qui *agissent* bien mais plutôt qui des individus qui sont aptes à *connaître* les bonnes positions morales, il s'agit toutefois d'une option qui a son lot de problèmes.⁴⁹ En effet, même en rétrécissant le champ des vertus recherchées aux vertus intellectuelles, il nous reste toujours à déterminer les critères de l'expertise. Est-ce le fait d'être très intelligent? Le fait d'être capable d'analyses véritablement impartiales et objectives? Le fait d'avoir beaucoup réfléchi sur des questions d'ordre moral? Ou peut-être même le fait d'avoir une formation académique substantielle en philosophie morale? Sans vouloir tomber dans la caricature – « X est un expert moral ssi X possède un doctorat en philosophie morale » – de telles questions ont le mérite de montrer la difficulté de l'identification d'experts.⁵⁰

Une autre option qui peut être envisagée est de limiter la légitimité de l'équilibre réfléchi non

49 On pourrait objecter qu'avoir des vertus intellectuelles n'est pas suffisant pour être un expert moral, puisqu'il faut avoir « le cœur à la bonne place » pour effectuer de bons jugements moraux. Il faudrait donc ajouter un critère émotionnel ou empathique en plus des critères strictement intellectuels.

50 Une piste de solution intéressante peut être celle développée par Ernest Sosa avec le modèle de l'épistémologie de la vertu. Voir « The Raft and the Pyramid » (Sosa, 1980).

pas aux experts mais bien aux individus jugés *compétents*. Bien que la tâche de définir la compétence est loin d'être sans problèmes, elle paraît néanmoins moins ardue que celle de définir l'expertise. Une telle stratégie, que l'on peut appeler la méthode de l'équilibre réfléchi *compétent*, peut être associée à Rawls lui-même, ou plus exactement au *jeune* Rawls. Dans l'article publié en 1951 où il développe pour la première fois l'idée qu'il baptisera vingt ans plus tard la méthode de l'équilibre réfléchi, Rawls propose une description de ce qu'il appelle la classe des juges moraux compétents (« competent moral judges ») en offrant quatre conditions jugées nécessaires à la compétence morale.⁵¹ En laissant de côté le problème de l'identification des critères nécessaires et suffisants de la compétence, la question qui est véritablement centrale est de savoir pourquoi il est préférable de limiter l'équilibre réfléchi aux individus compétents plutôt qu'aux experts. Après tout, le simple fait qu'il soit difficile (impossible?) d'identifier une classe d'experts de manière non-controversée ne signifie pas que l'équilibre réfléchi d'experts ne soit pas théoriquement préférable à l'équilibre réfléchi compétent. Je ne répondrai pas ici directement à cette question; j'ai abordé à la section 3.3.3 les raisons pouvant nous amener à préférer considérer les intuitions d'individus compétents plutôt que de se limiter à celles d'experts. Je me contenterai donc ici de renvoyer à cette section.

La deuxième stratégie qui mène à d'autres modèles de l'équilibre réfléchi consiste à se référer à des collectivités plutôt que simplement à des individus. La version la plus modérée de cette stratégie consiste à ajouter comme critère d'évaluation de nos jugements et principes moraux le fait que ces jugements et principes soient largement acceptés dans notre société. On pourrait par exemple discréditer dans le processus d'épuration initiale (étape 1 de la démarche rawlsienne) tous les jugements et principes qui sont clairement rejetés par notre société. Une telle méthode est généralement associée à Aristote, qui soutient entre autres dans les *Topiques* (1997) que l'on doit débiter toute enquête philosophique avec ce qu'il appelle les *endoxa*, c'est à dire les opinions jugées crédibles du fait qu'elles soient acceptées par tous, par la majorité ou par les sages (*Top.* i 1, 100b21–23). Une telle approche n'a pas beaucoup d'adeptes en philosophie contemporaine cependant⁵², principalement parce que le fait qu'une opinion soit acceptée par un grand nombre

51 Les juges compétents ont (1) un degré d'intelligence normal, (2) un caractère raisonnable (en un sens que Rawls décrit plus en détails), (3) une connaissance des choses du monde et des conséquences des actions que tout individu normalement intelligent devrait connaître et (4) une connaissance du phénomène des intérêts qui, lorsqu'entrant en conflit, peuvent mener au besoin de prendre des décisions morales (Rawls, 1951, 178-182). Rawls soutient que ces quatre conditions ne sont pas arbitraires : « [They are] those characteristics which, in the light of experience, show themselves as necessary conditions for the reasonable expectation that a given person may come to know something » (*ibid.*, 181).

52 Danels critique notamment l'idée d'ajouter un tel critère d'acceptabilité sociale comme étant artificiel : « Suppose

d'individus est loin d'être un gage de vérité.⁵³

Une autre version de cette stratégie qui accorde encore plus d'importance aux collectivités consiste à aborder celles-ci non pas comme offrant des critères d'acceptabilité de nos jugements et principes, mais plutôt comme étant *le lieu même* où s'exerce l'équilibre réfléchi. L'idée est ici de concevoir l'équilibre réfléchi comme s'exerçant non pas à l'intérieur d'un individu, mais plutôt à l'intérieur d'une communauté (e.g., la communauté scientifique), d'une société, d'une génération, ou même de l'humanité dans son ensemble. Cette stratégie de *délocalisation* ou de *collectivisation* de l'équilibre réfléchi est conçue par ses adeptes comme la meilleure façon de défendre la possibilité de l'atteinte de la vérité dans un quelconque domaine d'enquête. George Bealer défend entre autres une telle thèse :

« But, despite their fallibility, intuitions on my view nevertheless have a certain kind of strong modal tie to the truth. This tie is not “local”, however, since individual intuitions can be mistaken. Nor is the tie an ordinary holistic tie: I accept the possibility that some hypothetical subject's best efforts at the theoretical systemization of his intuitions might be mistaken. Rather, the tie is relativized; specifically, it is relativized to theoretical systemizations arrived at in relevantly high quality cognitive conditions. Such conditions might be beyond what individual human beings can achieve in isolation. It is plausible that we approximate such cognitive conditions only in sustained cooperation with others, perhaps over generations. And even here, it is an open question whether we will ever approximate them sufficiently closely. [...] I believe that, *collectively, over historical time, undertaking philosophy as a civilization-wide project*, we can obtain authoritative answers to a wide variety of central philosophical questions » (Bealer, 1998, 202-203).

Ce que propose Bealer, donc, est de concevoir le projet scientifique humain dans son ensemble comme le point de référence de la vérité. Il est clair que l'approche de Bealer n'est pas à proprement parler une *méthode* pour faire de la philosophie, mais plutôt une manière de concevoir la vérité et le type de relation qui unit l'intuition individuelle à cette vérité. En ce sens, cette approche est parfaitement compatible avec les différents modèles *localisés* de la méthode de l'équilibre réfléchi décrits précédemment. D'ailleurs, concevoir la vérité comme l'aboutissement du

we begin by admitting into the set of initial considered moral judgments only those judgments on which there is substantial consensus. [...] [T]he approach makes the wide equilibrium that emerges (if one does) much more a collective or social product from the start than does my approach, which is a quite unnatural idealization in this regard » (Daniels, 1979, 281-2).

53 Si on interprète cependant le critère d'acceptabilité sociale de manière légèrement différente, celui-ci peut sembler plus plausible. Il s'agirait de soutenir que pour qu'un jugement puisse légitimement être qualifié d'« intuitif », il faut qu'il puisse être reconnu comme tel par tout individu minimalement compétent – ce qui nous renvoie à la discussion sur le phénomène du désaccord intuitif de la section 3.2.

projet scientifique humain peut nous aider à accepter le caractère faillibiliste de la méthode de l'équilibre réfléchi; la déception découlant du fait que même notre meilleur effort individuel pour atteindre un équilibre réfléchi large ne nous permette jamais d'atteindre un état de certitude est compensée par le fait que cet effort individuel s'inscrit à l'intérieur d'une marche collective vers la vérité. Autrement dit, la perspective plus large de Bealer nous rappelle que l'effort du chercheur individuel n'est jamais vain.

Voilà donc les différents autres modèles de l'équilibre réfléchi. Bien que leur présentation ait été plutôt sommaire, je pense pouvoir conclure qu'aucun d'eux ne vient véritablement menacer le modèle généralement reconnu comme étant le plus sophistiqué, à savoir l'équilibre réfléchi large décrit par Daniels.⁵⁴ On retiendra tout de même de ces modèles l'importance d'avoir une compétence minimale pour appliquer la méthode de l'équilibre réfléchi, de même que la possibilité d'interpréter le caractère faillibiliste de la méthode en lien avec une perspective collectiviste de l'atteinte de la vérité.

4.3. Conclusion

Ce chapitre a porté sur la méthodologie standard en philosophie morale contemporaine. J'ai analysé successivement les deux éléments de cette méthodologie, à savoir l'usage des expériences de pensée et la méthode de l'équilibre réfléchi. Mon analyse de l'usage des expériences de pensée fut plutôt brève, principalement parce que la plupart des arguments utilisés pour et contre les expériences de pensée sont également des arguments utilisés pour et contre l'usage des intuitions en philosophie, sujet largement discuté au chapitre 3. J'ai néanmoins présenté les deux principaux modèles contemporains des expériences de pensée, à savoir le modèle empiriste de John Norton et le modèle rationaliste de James Brown, concluant que la différence essentielle entre les deux modèles renvoie à la différence entre réalisme et anti-réalisme.

La partie la plus longue du chapitre a porté sur la méthode de l'équilibre réfléchi. J'ai d'abord présenté le modèle standard de la méthode développée par Rawls, en soulignant certaines critiques

54 L'une des rares voix discordantes qui soutient que le modèle de l'équilibre large n'est guère supérieur au modèle de l'équilibre étroit est celle de Margaret Holmgren dans « The Wide and Narrow of Reflective Equilibrium » (1989). Holmgren soutient que la plupart des vertus présumées de l'équilibre large se retrouvent également dans l'équilibre étroit. L'équilibre étroit – qu'elle nomme « intuitionnisme sophistiqué » pour reprendre l'expression de Daniels – permet par exemple d'intégrer tout autant des théories non-morales : « Moral intuitionists have always acknowledged that nonmoral background theories can supply *subsumptive* premises in moral arguments, and NRE clearly accomodates their use in this role. What is important in NRE is that the *nomological* premisses of moral arguments are moral principles, and that the credibility of these principles is gauged by the extent to which they economically systematize our considered moral judgments » (Holmgren, 1989, 54).

qui lui ont été soulevées, notamment le fait que la méthode paraît uniquement décrire notre conception morale sans permettre de la justifier. J'ai ensuite présenté le modèle plus sophistiqué de l'équilibre réfléchi large, qui permet de répondre aux objections soulevées contre le modèle standard en plus de bénéficier d'autres vertus. J'ai également souligné la possibilité d'offrir une interprétation fondationnaliste plutôt que cohérentiste de la méthode. Enfin, j'ai présenté rapidement quelques modèles alternatifs de la méthode, alternatives qui viennent soit restreindre l'usage légitime de la méthode, soit l'étendre aux collectivités. J'ai conclu qu'aucune de ces alternatives ne venaient véritablement menacer le modèle de l'équilibre réfléchi large.

La méthode de l'équilibre réfléchi large ressort donc au final avec le statut de modèle le plus efficace afin de développer une théorie morale. Je souhaite cependant aller plus loin et conclure en soulignant que l'on peut également concevoir ce modèle comme étant un modèle *par défaut*. C'est là une thèse qui est admirablement défendue par Michael DePaul dans « Why Bother With Reflective Equilibrium? » (1998). La question que pose et à laquelle cherche à répondre DePaul dans cet article est : quelle alternative à la méthode de l'équilibre réfléchi large existe-t-il?⁵⁵ DePaul conclut qu'aucune véritable alternative acceptable n'existe, puisque tout modèle qui voudrait rejeter la quête d'un équilibre réfléchi large serait tout simplement irrationnel. En effet, puisque la méthode enjoint à prendre en considération *l'ensemble* de nos croyances et de les analyser, confronter entre elles et systématiser, elle constitue la seule démarche rationnelle pouvant être adoptée par un individu sérieux dans sa quête de savoir.

Même les sceptiques de l'usage des intuitions en philosophie, par exemple, acceptent nécessairement, par défaut, la méthode de l'équilibre réfléchi large, dans la mesure où leurs réserves à l'endroit des intuitions sont simplement des arguments supplémentaires qu'ils jugent devoir être pris en considération dans l'équilibre réfléchi large. C'est là un exemple que décrit DePaul lui-même pour illustrer l'inévitabilité de la méthode de l'équilibre réfléchi large :

« [...] [T]he critic might cite studies by cognitive psychologists showing that our intuitive judgments about the area in question are unreliable, and press the inquirer not to allow these intuitive judgments any weight at all in determining the theory she ends up accepting. In all likelihood, what the critic is thinking is that the inquirer is ignorant of the relevant studies and that she accepts many of the same forms of arguments and background epistemological views that he does, so that when she is made aware of the studies and the implications of these studies [...], she will accept the studies and implications just as the critic has. If this is what the critic is doing, then he really is not doing anything that conflicts with the

55 Comme il a été mentionné à la note 41, DePaul parle simplement d'équilibre réfléchi, mais il a clairement en tête l'équilibre *large* de Daniels.

method of reflective equilibrium. He is merely providing the inquirer with information acquired in ways the inquirer accepts which conflicts with some of what the inquirer believes and expecting the inquirer to make appropriate adjustments so that her system of beliefs is again coherent. This is nothing more nor less than what reflective equilibrium dictates » (DePaul, 1998, 304).

Bref, puisqu'un individu qui recherche honnêtement la vérité ne peut légitimement exclure un argument ou une explication qui lui semble raisonnable ou refuser d'entendre un tel argument ou explication, un tel individu est *ipso facto* un adepte de la méthode de l'équilibre réfléchi large. Rejeter cette méthode revient à ne pas rechercher sérieusement la vérité, puisqu'on ne peut rejeter la méthode qu'en effectuant l'une ou l'autre des trois choses suivantes : « (A) abandon reflection altogether, or (B) direct the inquirer to reflect, but to do so incompletely, that is, to leave certain beliefs, principles, theories, or what have you out of account, or (C) not allow the results of the inquirer's reflections to determine what the inquirer goes on to believe. » (*ibid.*, 301). Puisqu'aucune de ces options n'est rationnelle, on se retrouve avec le modèle de l'équilibre réfléchi large comme étant le modèle par défaut pour toute démarche philosophique ou scientifique. On peut comprendre cette inévitabilité en lien avec ce que Wilfrid Sellars décrivait comment étant le but ultime de la philosophie : « The aim of philosophy, abstractly formulated, is to understand how things in the broadest possible sense of the term hang together in the broadest possible sense of the term » (Sellars, 1963, 39-40).

Comprise en ce sens, la méthode de l'équilibre réfléchi large cesse d'être envisagée comme la rivale de la démarche philosophique dite de la « vieille école », à savoir la démarche intuitionniste. Si DePaul a raison, l'intuitionnisme doit être compris en harmonie avec les démarches décrites par Rawls et Daniels. C'est là une idée qui en fera sursauter plus d'un, à commencer par les deux principaux intéressés, qui ont toujours compris leur démarche en opposition à la démarche intuitionniste. Je consacre le prochain et dernier chapitre à cette démarche intuitionniste, et je chercherai entre autres à voir dans quelle mesure cette démarche peut être vue comme compatible avec la méthode de l'équilibre réfléchi.

5. INTUITIONNISME ÉTHIQUE

L'intuitionnisme éthique est souvent considéré comme la théorie « de la vieille école » en philosophie morale. Ce titre souligne bien sûr le caractère ancien de la théorie, qui fut développée et défendue avant l'utilitarisme et le kantisme, mais il est également utilisé de manière péjorative par certains de ses adversaires, qui cherchent ainsi à souligner son caractère « dépassé » ou « vieux jeu ».

La théorie intuitionniste a plus d'une fois été jugée dépassée à travers son histoire. Mais comme le phénix qui renaît de ses cendres, elle a à chaque fois connu un nouveau souffle. La théorie est apparue à la fin du XVII^e siècle et au début du XVIII^e siècle⁵⁶ en opposition à deux courants majeurs : la théorie hobbesienne fondant la moralité sur l'intérêt personnel éclairé (Hobbes, 1651) et la théorie du droit naturel défendue notamment par Samuel von Pufendorf (1660) et John Locke (1689)⁵⁷. Des moralistes britanniques tels Richard Cumberland (1672), Samuel Clark (1728), John Balguy (1728), Ralph Cudworth (1731) et Richard Price (1758) ont tous à leur façon développé une théorie éthique défendant l'existence de certains principes moraux non-inférentiels connus par intuition. Cette première phase de l'intuitionnisme a pris fin avec l'arrivée des théories sentimentalistes d'inspiration humienne, puis avec le développement de l'utilitarisme.

Après être tombé dans l'oubli pendant près d'un siècle, l'intuitionnisme a connu une deuxième phase de popularité de la fin du XIX^e siècle jusqu'au premier quart du XX^e siècle. Ce regain d'intérêt est principalement attribuable à Henry Sidgwick et à la publication de *The Methods of Ethics* (1874).⁵⁸ L'approche innovatrice de Sidgwick a été de chercher à réconcilier l'utilitarisme et l'intuitionnisme en soutenant que le principe d'utilité reposait sur une intuition fondamentale. G. E. Moore, un élève de Sidgwick, a poursuivi le relance de l'intuitionnisme avec la publication de *Principia Ethica* (1903), où il défend entre autres l'idée que la propriété d'être « bien » est non-naturelle et non-réductible et ne peut être connue que par intuition. La publication de *The Right and the Good* (1927) par W. D. Ross a ensuite mené à l'apogée de cette seconde phase de l'intuitionnisme. Ross y défend une version déontologique de l'intuitionnisme et développe une forme de pluralisme ancré dans des devoirs *prima facie* connus par intuition. Malgré l'impact

56 Robert Audi (2005, 5) affirme cependant qu'on retrouve des traces de l'intuitionnisme aussi loin que chez Thomas d'Aquin, qui défendait déjà l'idée de principes non-inférentiels connus par intuition.

57 Stephen Darwall (2002, 256) décrit ces deux courants comme étant deux formes de *volontarisme* éthique, la version hobbesienne étant une forme séculaire de volontarisme, et la version du droit naturel étant une forme théologique. Selon Darwall, l'adversaire principal de l'intuitionnisme était donc le volontarisme éthique.

58 Philip Stratton-Lake (2002, 5) présente pour sa part la période sidgwickienne comme étant distincte de la période allant de Moore à Ross, qu'il voit comme une troisième phase de l'intuitionnisme.

important de la publication de Ross, l'intuitionnisme fut une fois de plus discrédité, et parfois même méprisé, suite à la Seconde Guerre mondiale, notamment en raison de l'émergence du non-cognitivism.⁵⁹

Depuis le tournant du nouveau millénaire, on assiste à ce que plusieurs décrivent comme une troisième phase de l'intuitionnisme, qui s'appuie largement sur la version de la théorie défendue par Ross. Bien que la théorie ne puisse être dite dominante aujourd'hui, elle est néanmoins perçue comme une candidate importante dans les débats contemporains en épistémologie morale et en éthique normative.⁶⁰ L'intuitionnisme est aujourd'hui défendu sous différentes formes entre autres par Robert Audi, Jonathan Dancy, John McDowell, David McNaughton, Derek Parfit et David Wiggins.⁶¹ Ce regain de popularité et de crédibilité s'explique par plusieurs facteurs, à commencer par le retour en force du cognitivism dans le dernier tiers du XXe siècle suite aux nombreuses critiques soulevées à l'endroit du non-cognitivism. Comme l'explique Philip Stratton-Lake :

« With the demise of the dominance of non-cognitivism in ethics came a renewed belief that cognitivism and some form of moral realism are defensible after all. And once these views were no longer regarded as crazy, intuitionism could no longer be dismissed simply because it is committed to them » (Stratton-Lake, 2002, 24)

Un autre facteur important que soulignent Stratton-Lake et d'autres est la réhabilitation de l'idée de connaissance *a priori*, voire *synthétique a priori*, après que celle-ci ait été pendant longtemps jugée intenable, principalement suite à la critique autrefois considérée dévastatrice par Quine (1951) de la distinction entre le synthétique et l'analytique. Avec la rédemption du cognitivism et de l'idée de connaissance *a priori*, la porte était ouverte pour un retour en force de la théorie « de la vieille école ».⁶² Robert Audi a contribué plus que tout autre philosophe contemporain à la réhabilitation de l'intuitionnisme, d'abord avec la parution de l'article « Intuitionism, Pluralism, and the Foundations of Ethics » (Audi, 1997), puis avec la parution en

59 D'autres philosophes que l'on associe à cette deuxième phase intuitionniste sont H. Rashdall (1907), H. A. Prichard (1912), C. D. Broad (1930) et A. C. Ewing (1947).

60 En guise d'illustration de cette crédibilité renouvelée, Klemmens Kappel concluait son article « Challenges to Audi's Ethical Intuitionism » de 2002 avec la remarque suivante : « Let me finally emphasize that the issues I have tried to raise obviously do not constitute anything near decisive objections to Audi's intuitionism. I hope that they can be seen as fruitful challenges to the welcome revival of a serious contender in moral epistemology » (Kappel, 2002, 412).

61 La liste provient de Philip Stratton-Lake (2002, 23). Outre Stratton-Lake lui-même, on peut entre autres ajouter à cette liste Brad Hooker, Michael Huemer, Thomas Hurka et Russ Shafer-Landau.

62 Un troisième facteur qui a contribué à la réhabilitation de l'intuitionnisme est le développement de la méthode de l'équilibre réfléchi par Rawls, qui a permis de légitimer l'usage des intuitions en philosophie.

2005 du livre *The Good in the Right : A Theory of Intuition and Intrinsic Value*. Audi a entre autres montré que plusieurs des attaques envers l'intuitionnisme au cours de la deuxième moitié du XXe siècle reposaient sur une certaine incompréhension, voire une caricature de la théorie. Il a également développé une défense sophistiquée de ce qu'il a présenté comme le cœur de la théorie intuitionniste, à savoir l'idée de principes moraux « évidents en eux-mêmes ».

Je n'aborderai dans ce chapitre que les débats intuitionnistes de la deuxième et troisième phases. M'inspirant des distinctions faites par Bernard Williams (1995) et Robert Audi (2005)⁶³ entre deux grands types d'intuitionnisme, j'analyserai dans un premier temps ce qu'on peut appeler l'intuitionnisme *méta-éthique*, qui consiste en l'idée que certains principes moraux sont justifiés de manière *non-inférentielle* (5.1), puis l'intuitionnisme *normatif*, qui soutient qu'il existe une pluralité de principes moraux premiers irréductibles pouvant entrer en conflit (5.2). Après avoir ainsi analysé les principaux types d'intuitionnisme et les enjeux qu'ils soulèvent, je serai en mesure d'aborder la question qui avait conclu le chapitre précédent, à savoir dans quelle mesure l'intuitionnisme et la méthode de l'équilibre réfléchi peuvent être réconciliés (5.3). Je conclurai enfin le chapitre avec un résumé rapide, préparant la conclusion finale du mémoire (5.4).

5.1. Intuitionnisme méta-éthique

La distinction entre l'intuitionnisme méta-éthique et l'intuitionnisme normatif est importante dans la mesure où ces deux théories, bien que souvent placées sous la même étiquette, portent sur des enjeux fort différents et n'ont pas à être défendues conjointement. S'il y a plusieurs exemples d'intuitionnistes qui défendent les deux théories – on peut penser à Clark, Price, Ross, Ewing, Audi et Huemer – et s'il est vrai que pratiquement tous ceux qui endossent la thèse normative endossent également la méta-éthique, l'inverse n'est pas toujours vrai : Sidgwick et Moore, par exemple, défendaient la thèse méta-éthique mais non la normative.

Je débute dans cette section avec l'intuitionnisme méta-éthique. La meilleure façon d'aborder cette thèse est de la comprendre comme une réponse au trilemme de Münchhausen,⁶⁴ réponse qui consiste à dire que l'on peut parvenir à mettre fin à la régression à l'infini en cessant tout

63 Audi reprend essentiellement la même distinction que Williams, mais rebaptise chacune des deux versions de la théorie. Alors que Williams parle d'« intuitionnisme méthodologique » et d'« intuitionnisme épistémologique », Audi parle plutôt d'« intuitionnisme normatif » et d'« intuitionnisme méta-éthique ». Les termes utilisés par Audi me semblent plus clairs, et je me référerai donc à ceux-ci.

64 Voir section 4.2.2 pour une présentation du trilemme.

simplement la chaîne d'inférences une fois que l'on a atteint certains principes fondamentaux ne nécessitant aucune justification au-delà d'eux-mêmes⁶⁵ :

« The easiest way to stop this regress is simply to stop. If a believer can work back to a premise that the believer is justified in believing without being able to infer that premise from anything else, then there is no new premise to justify, so the regress goes no further. That is how moral intuitionists stop the regress, and many moral intuitionists see it as the only way to stop » (Sinnott-Armstrong, 2002, 306).

La façon la plus courante d'interpréter l'idée de croyances justifiées de manière non-inférentielle est en se référant à l'existence de croyances « évidentes en elles-mêmes ». Or, comme le souligne Pekka Väyrynen (2008), il ne s'agit là que de l'une des deux grandes stratégies pouvant être utilisées par les intuitionnistes afin de mettre fin à la chaîne d'inférences. Väyrynen distingue en effet entre l'intuitionnisme (méta-éthique) *a priori* et l'intuitionnisme (méta-éthique) *a posteriori*. Alors que la stratégie *a priori* aborde la justification non-inférentielle sur la base de la nécessité conceptuelle, la stratégie *a posteriori* l'aborde plutôt sur la base d'un modèle perceptuel :

« The standard way to understand *a posteriori* intuitionism is to model at least some of our ethical knowledge on (quasi-)perceptual knowledge (where expressions like 'ethical perception' and 'moral vision' need not be taken literally). If we have any non-inferential *a posteriori* knowledge, some of it is presumably had by perception. Similarly, it might be that we have at least some ethical knowledge by something broadly like perception of ethical facts » (Väyrynen, 2008, 496).

On remarque que Väyrynen parle ici de *connaissances* non-inférentielles plutôt que simplement de justification non-inférentielle.⁶⁶ La distinction entre les deux termes est subtile mais cruciale : une croyance peut être justifiée sans pour autant constituer une connaissance. Cela arrive lorsqu'une croyance justifiée s'avère être fausse. Walter Sinnott-Armstrong donne l'exemple d'un fermier au Moyen-Âge qui croit que sa femme a le devoir moral de lui obéir et qui est justifié dans cette croyance du fait qu'elle est acceptée par tous, à commencer par les autorités religieuses, et qu'il

65 Une erreur fréquente est de soutenir que l'intuitionnisme méta-éthique suppose que l'on ne *puisse* pas inférer les propositions jugées évidentes en soi. Il s'agit d'une erreur dans la mesure où l'intuitionnisme n'affirme pas qu'une telle inférence soit impossible, mais simplement qu'elle n'est pas *nécessaire*. Par exemple, il est possible d'inférer la proposition « mentir est *prima facie* mal », en disant que mentir est mal parce qu'une telle pratique ne peut être universalisée ou qu'elle constitue un manque de respect ou qu'elle enfreint le droit à la vérité. Cependant, ce qu'un intuitionniste soutient est que ce type d'inférence n'est pas nécessaire pour que la proposition « mentir est *prima facie* mal » puisse être justifiée.

66 Merci à Christine Tappolet d'avoir mentionné ce point et de m'avoir suggéré d'aborder plus explicitement la distinction entre justification et connaissance.

n'a connaissance d'aucune raison de douter de cette croyance (Sinnott-Armstrong, 2002, 323). Évidemment, le fait que cet homme soit justifié d'avoir une telle croyance n'implique aucunement qu'il possède une *connaissance* que sa femme a le devoir moral de lui obéir, pour la simple raison qu'une telle croyance, bien que justifiée, n'est pas pour autant *vraie*.⁶⁷ Or, la plupart des intuitionnistes, *a priori* comme *a posteriori*, conçoivent les croyances justifiées de manière non-inférentielle comme étant véritablement des *connaissances*, dans la mesure où elles sont également jugées être vraies.⁶⁸

On remarque également la prudence affichée par Väyrynen lorsqu'il aborde l'idée de « perception » morale. Cela se comprend par le fait qu'un des reproches principaux qui a historiquement été soulevé contre l'intuitionnisme est que la théorie supposerait l'existence d'une faculté perceptuelle morale que plusieurs qualifient d'étrange.⁶⁹ Or, comme le souligne Väyrynen, l'idée de perception n'a pas à être comprise de manière littérale – ce qui n'implique pas non plus qu'elle doive être réduite à une simple analogie ou métaphore. Comme l'explique Christine Tappolet : « On peut [...] soutenir qu'il existe non pas un organe moral ou esthétique, mais des états suffisamment semblables à nos expériences perceptuelles pour qu'il soit légitime de parler d'expériences de valeurs » (Tappolet, 2000, 125). L'exemple que Tappolet a en tête ici est celui des émotions, dont la phénoménologie est à bien des égards similaire à celle de la perception. Tappolet défend une forme d'intuitionnisme méta-éthique *a posteriori*, dans la mesure où elle conçoit les croyances portant sur des valeurs comme étant justifiées de manière *prima facie* par nos émotions, du fait que les émotions sont des perceptions de valeurs.

La stratégie la plus largement défendue par les intuitionnistes est cependant la stratégie *a priori*, et je me concentrerai donc ici sur celle-ci.⁷⁰ Tel que mentionné plus haut, cette approche consiste généralement à soutenir que certains principes moraux sont « évidents en eux-mêmes » et ne nécessitent dès lors aucune justification externe.⁷¹ Cette notion d'évidence en soi était défendue

67 Ces points sont bien sûr controversés. On peut douter de l'idée que le fermier médiéval soit justifié du simple fait de l'absence de désaccord connu et de son ignorance de toute raison de douter de sa croyance. Un nihiliste pourrait également souligner le caractère douteux de la notion même de vérité morale. Je n'aborde cependant pas davantage ces objections possibles, mon but étant simplement d'illustrer l'idée de croyance justifiée mais fautive.

68 Je laisse de côté ici tout le débat initié par Gettier (1963) concernant le caractère non-suffisant des croyances vraies justifiées au statut de connaissances. Aborder ce débat n'est guère essentiel à mon propos.

69 Voir notamment Mackie, 1977.

70 Sans entrer dans les détails, l'un des principaux problèmes auxquels doit répondre l'intuitionniste *a posteriori* est la possibilité que nos perceptions morales soient inconsciemment guidées par nos croyances morales préalables. Väyrynen discute en profondeur dans la section IV de son chapitre de l'impact possible sur l'intuitionnisme *a posteriori* du fait que nos observations soient marquées par nos théories (« that observations are theory-laden ») et analyse également différentes réponses offertes par les défenseurs de l'intuitionnisme *a posteriori*.

71 Il existe au moins une exception: Michael Huemer (2005) défend une forme d'intuitionnisme méta-éthique *a priori* qui ne repose pas sur l'idée d'évidence en soi. Huemer conçoit les intuitions comme des « apparences

par les trois principaux intuitionnistes de la deuxième phase, Sidgwick, Moore et Ross, qui la concevaient cependant de manière différente. Alors que Moore disait simplement qu'une proposition évidente en soi ne fonde sa vérité sur aucune raison autre qu'elle-même – « there is no reason why it is true » (Moore, 1903, §86) –, Sidgwick allait plus loin et offrait quatre critères permettant de déterminer si, et à quel degré, une proposition pouvait être dite évidente en soi :

« Any philosophically respectable conclusions about how we should act must meet the conditions of self-evidence. They should be clear, reflective, consistent, and meet with consensus on the part of the many and the wise, and possess these qualities above certain thresholds. Further, the greater the extent to which they exceed the relevant thresholds, the more self-evident and hence reliable they are » (Sidgwick, 1874, 72).

La définition qui a cependant eu l'impact le plus important et qui est aujourd'hui le point de référence des débats sur la notion d'évidence en soi est celle développée par Ross : une proposition est évidente en elle-même pour Ross si et seulement si elle est connaissable simplement sur la base de sa compréhension (« knowable on the basis of an understanding alone » (Ross, 1927, 20, note 1). Une telle définition peut donner lieu à plusieurs confusions, et Ross s'est assuré d'apporter davantage de précisions :

« [A proposition is evident] not in the sense that it is evident from the beginning of our lives, or as soon as we attend to the proposition for the first time, but in the sense that when we have reached sufficient mental maturity and have given sufficient attention to the proposition it is evident without any need of proof, or of evidence beyond itself. It is self-evident just as a mathematical axiom, or the validity of a form of inference, is evident » (*ibid.*, 29).

Cette définition rossienne de l'évidence en soi a été reprise et raffinée dans le débat contemporain par Robert Audi, qui a entre autres distingué deux grandes étapes dans la définition de Ross : « Self-evident propositions are conceived as truths such that (a) adequately understanding them is sufficient for justification for believing them, [...] and (b) believing them on the basis of adequately understanding them entails knowing them » (Audi, 2008, 503).

De nombreuses objections ont été et continuent à être soulevées à l'endroit de cette idée qu'il

intellectuelles », et il soutient que comme toute autre forme d'apparences (perceptuelles, introspectives, etc.), les intuitions jouissent *de facto* d'une justification *prima facie*. Si sa théorie fait une place importante à l'idée que certaines propositions morales nous apparaissent « manifestes » ou « patentes » (« obvious »), il ne juge pas nécessaire cependant de dire que quelconque proposition soit « évidente en elle-même » (« self-evident »). Voir entre autres pp. 107 et 115 de Huemer, 2005.

existerait des propositions « automatiquement justifiées » du fait de leur évidence. Plusieurs de ces objections reposent cependant sur une mauvaise compréhension du concept d'évidence en soi utilisé par Ross, Audi et autres. Certains reprochent par exemple aux intuitionnistes *a priori* d'accorder un statut épistémique particulier à des propositions morales simplement en vertu du fait qu'elles leur paraissent très intuitives. Or, ces opposants poursuivent, la force intuitive d'une proposition renvoie à la psychologie des individus et ne suppose en elle-même aucun statut épistémique particulier de ces propositions. C'est là une critique que soulève entre autres Allan Gibbard :

« [...] I might, say, find it intuitive that God exists, accepting confidently that he does, on no further grounds. Whether I then have intuitive knowledge, though, that God exists will be controversial. [...] To say that I find it intuitive is to describe my state of mind. To say that I *know* by intuitions is to describe something stronger » (Gibbard, 2002, 225).

Il est erroné cependant de soutenir que les intuitionnistes *a priori* appliquent le concept d'évidence en soi sur la base de la simple « force intuitive » des propositions considérées. Comme le souligne Philip Stratton-Lake, seule une catégorie très limitée de propositions peut potentiellement être dite évidente en elle-même :

« Intuitionists are very careful which of their convictions they claim the status of self-evidence for. It is only those convictions that can plausibly be said to be basic that are even contenders for the status of self-evidence, and of these it is only those that survive careful reflection on their meaning and coherence that are claimed to be self-evident. » (Stratton-Lake, 2002, 22).

Trois autres critiques du concept d'évidence en soi reposent sur une mauvaise compréhension de la nature de ce concept utilisé par les intuitionnistes *a priori*. D'abord, certains font l'erreur de confondre ou de voir comme synonymes les concepts « self-evident » et « obvious » – ou en français, « évident en soi » et « patent » ou « manifeste ». Or, une proposition peut être évidente en soi sans pour autant être patente/manifeste. Par exemple, la proposition « tous cousins immédiats ont la même paire de grand-parents » ou la proposition logique $\neg(p \wedge q) \text{ ssi } (\neg p) \vee (\neg q)$ peuvent ne pas être saisies immédiatement et nécessiter un moment de réflexion sans que cela n'affecte leur caractère évident en soi. À l'inverse, une proposition peut être patente/manifeste sans qu'elle soit évidente en soi. On retrouve des exemples de ce type dans nos expériences empiriques quotidiennes; il est par exemple patent, mais non évident en soi, que si je laisse tomber un objet, sa trajectoire ira vers le bas et non vers le haut.

Une autre critique consiste à soutenir que les principes évidents en soi ne peuvent jamais perdre leur justification et sont ainsi « imbattables » (« indefeasible »). Une telle critique n'est bien sûr pas sans fondement, puisque nombreux sont les intuitionnistes qui soutiennent que certains principes moraux élémentaires sont *nécessairement vrais*, sur la base du modèle kantien de vérités synthétiques *a priori*. C'est le cas notamment de Robert Audi. Or, on peut tout à fait soutenir que certains principes moraux sont nécessairement vrais et soutenir en même temps qu'on peut ne pas être justifié de les croire. Comme le dit Audi : « As we know from studies in logic and mathematics, even what is *a priori* and necessary need not be such that one cannot lose one's justification for believing it. We may, for instance, discover errors in our crucial proof » (Audi, 2002, 49).⁷² Klemmens Kappel apporte une précision éclairante en mentionnant que le caractère faillible des propositions évidentes en soi découle de la faillibilité de notre *capacité* à reconnaître de telles propositions :

« Of course, self-evident propositions are true, perhaps even necessarily true. Hence, it never happens that one accepts a self-evident proposition and yet it is false. Accepting self-evident propositions never leads to falsehoods. Nonetheless, our *capacity* to accept propositions merely on the basis of an understanding of them is fallible » (Kappel, 2002, 395).

Une troisième critique consiste à dire que les propositions évidentes en elles-mêmes doivent obtenir un assentiment universel pour être reconnues comme évidentes. Or, les réponses aux deux critiques précédentes devraient suffire à répondre à celle-ci : puisque les principes évidents en eux-mêmes ne sont pas toujours patents/manifestes et puisque notre capacité à reconnaître de telles propositions est faillible, il n'est guère surprenant qu'il existe des désaccords à leur sujet. Certains insistent cependant pour dire que l'étendue du désaccord moral devrait nous amener à douter du fait qu'il existe réellement quelque proposition évidente en soi, tout spécialement lorsque le désaccord provient de spécialistes en la matière. Berys Gaut (2002) défend un argument de ce type, en prenant l'exemple rossien des devoirs *prima facie* :⁷³

72 Si l'on est inconfortable avec l'idée que la *vérité* (et non la justification) de certains principes moraux élémentaires soit jugée *infaillible*, on peut toujours opter pour un modèle comme celui de Michael Huemer (2005), qui ne comporte aucune référence à l'idée d'évidence en soi. Bien que Huemer ouvre la porte à l'idée de vérités morales infaillibles, il précise qu'une telle thèse n'est pas nécessaire à la forme d'intuitionnisme qu'il défend : « I do not wish to *rule out* [...] the possibility of some intuitions' being infallible; I simply deny that they *must* be infallible » (Huemer, 2005, 106).

73 Walter Sinnott-Armstrong (2002) défend également un argument de ce type, en soulignant que l'existence reconnue de désaccords importants dans le domaine moral en général est suffisant pour fonder le besoin d'inférences pour obtenir une justification morale, *contra* la thèse intuitionniste. Je n'aborde pas plus en détails la critique de Sinnott-Armstrong, essentiellement parce qu'elle ne porte pas directement sur l'existence de principes évidents en eux-

« On a broadly Rossian account of self-evidence, we should say that those who deny that some moral duties are self-evident have failed adequately to understand the propositions about these duties. [...] But such failures of understanding or reflection are scarcely credible as explanations of why many philosophers have denied the self-evidence of these propositions. Moore, for example, was at least as talented a philosopher as Ross. [...] So the self-evidence view does not have a plausible theory of error to explain why the deniers of self-evidence have failed to grasp the purported self-evidence of these propositions » (Gaut, 2002,144).

On peut offrir plusieurs réponses à ce type d'objection. Une première réponse est de souligner que l'existence de désaccords quant à la vérité des devoirs *prima facie* de Ross n'implique aucunement qu'il n'existe *aucune* proposition évidente en elle-même. Même s'il était vrai que les devoirs *prima facie* de Ross n'étaient pas évidents en eux-mêmes, on pourrait encore imaginer d'autres candidats à ce statut. Or, s'il existe *une seule* proposition morale qui soit évidente en soi et donc justifiée de manière non-inférentielle, la thèse intuitionniste est alors validée, comme le souligne Audi dans sa définition minimale de l'intuitionnisme : « I am taking intuitionism in ethical theory to be, in outline and in a minimal version, the view that there is at least one moral principle that is non-inferentially and intuitively knowable » (Audi, 2008, 476).

Ensuite, le fait que certains individus intelligents, spécialistes en philosophie morale rejettent ce que presque tout le monde considère être des principes moraux évidents en eux-mêmes ne nous oblige pas à mettre en doute l'évidence en soi de ces principes. À ce titre, on pourrait aussi bien dire que l'existence de l'univers matériel est incertaine puisque l'évêque Berkeley, un homme intelligent spécialisé sur des enjeux d'épistémologie physique, défendait la thèse de l'immatérialisme.⁷⁴ Plusieurs raisons peuvent en effet expliquer pourquoi un individu intelligent en arrive à rejeter un principe que presque tout le monde juge évident en soi, à commencer par l'idéologie ou l'acceptation de théories d'arrière-fond qui colorent ses intuitions.

Plusieurs contre-arguments sont bien sûr offerts en réponse à cette défense de l'intuitionnisme méta-éthique *a priori*, mais je me contenterai de ce portrait sommaire. Si l'on est convaincu par l'argumentaire intuitionniste et que l'on l'on accepte l'existence de propositions morales évidentes en elles-mêmes, l'étape suivante est de chercher à identifier le type de propositions morales pouvant se voir attribuer un tel statut. Le modèle le plus répandu est celui des devoirs *prima facie* de David Ross, modèle que j'aborde à la section suivante sur l'intuitionnisme

mêmes, mais plutôt sur la légitimité d'une justification non-inférentielle.

74 L'exemple vient de Huemer, 2005, 132.

normatif.

5.2. Intuitionnisme normatif

J'ai souligné au début de la section précédente que la quasi-totalité des intuitionnistes acceptent la thèse *méta-éthique*. À une certaine époque, le consensus allait même au-delà des partisans de la « vieille école »; pour l'essentiel de la deuxième phase de l'intuitionnisme, la thèse intuitionniste méta-éthique était acceptée par pratiquement tous les éthiciens, qu'ils s'identifient ou non comme intuitionnistes. L'idée qu'il existe un ou des principes moraux justifiés de manière non-inférentielle était soit affirmée explicitement, soit acceptée implicitement par la communauté philosophique de l'époque. Le désaccord portait non pas sur l'idée même de principe non-inférentiel, mais sur la nature et le nombre de ces principes. David McNaughton résume bien ce constat :

« The dispute between intuitionists and utilitarians between the 2 wars was about whether or not there was more than one fundamental moral principle. Both parties to this debate were taken to agree that an ethical theory rests on intuition, by which was meant no more than that the most basic ethical principles, since they could not be inferred from more basic ones, must be self-evident. » (McNaughton, 2002, 76)

Alors que la plupart des utilitaristes soutenaient qu'il n'existait qu'un seul principe moral fondamental, à savoir le principe d'utilité,⁷⁵ d'autres défendaient l'existence d'une pluralité de tels principes. Les deux camps acceptaient l'intuitionnisme méta-éthique, mais seul le deuxième acceptait l'intuitionnisme normatif.

Le principal porte-étendard de l'intuitionnisme normatif, W. D. Ross, identifie sept devoirs moraux *prima facie*⁷⁶, à savoir : le devoir de fidélité, de réparation, de gratitude, de bienfaisance, de non malversation, de justice et d'éducation de soi.⁷⁷ Ross ne prétend pas que cette liste soit exhaustive, et d'autres ont proposé des modifications ou ajouts, notamment Robert Audi qui, dans son interprétation kantienne de Ross, a ajouté à cette liste deux autres devoirs (ou principes) moraux

75 En ce qui concerne Sidgwick, la nature de sa position est plus complexe. Bien qu'il ne reconnaisse qu'un seul principe fondamental en ce qui attrait au devoir moral, à savoir le principe d'utilité, il défend par ailleurs d'autres principes fondamentaux sous la forme d'*axiomes moraux*, dont au moins les deux suivants: « whatever action any of us judges to be right for himself, he implicitly judges to be right for all similar persons in similar circumstances » (Sidgwick, 1874, 379 et « the good of any one individual is of no more importance, from the point of view (if I may say so) of the Universe, than the good of any other » (*ibid.*, 382).

76 J'explique le qualificatif « *prima facie* » au paragraphe suivant.

77 Outre ces sept devoirs, Ross identifie également quatre « biens intrinsèques », à savoir la connaissance, le plaisir, la vertu et la justice (qui est à vrai dire la correspondance entre le plaisir et la vertu) (Ross, 1930, 136.)

fondamentaux : la promotion de la liberté et la promotion du respect (Audi, 2008, 476). Indépendamment de la composition finale de la liste de devoirs *prima facie*, l'idée essentielle est que ceux-ci sont valables en eux-mêmes et non en vertu d'un principe plus général, par exemple le principe d'utilité. Cette façon de concevoir une pluralité de principes moraux premiers semble fidèle aux intuitions populaires concernant les enjeux moraux, comme le souligne Philip Stratton-Lake :

« If, for instance, we cannot give up our sense that promises have a special binding force that is independent of the good consequences a commitment to fidelity will bring, then we cannot accept that a theory which requires us to do so could be correct » (Stratton-Lake, 2002, 26)

L'idée de concevoir de tels devoirs comme *prima facie*⁷⁸ était une innovation à l'époque de Ross. La tradition déontologique depuis Kant (1785) avait conçu les devoirs moraux de manière *absolue* : en aucun cas des actions impliquant un mensonge ou un vol, par exemple, ne pouvaient être considérées moralement acceptables. Le principal avantage comparatif de la théorie rossienne fut d'offrir une distinction entre les devoirs *prima facie* et les devoirs *réels* : si le fait de mentir ou de voler ajoute toujours une « valeur négative » à une action, dans la mesure où l'on enfreint par le fait même un devoir de fidélité ou de non-malfaisance, notre devoir *réel* ou *final* peut néanmoins nous enjoindre à commettre un mensonge ou un vol si l'on juge que cela nous permet d'accomplir un devoir encore plus important. On peut par exemple mentir ou voler pour sauver une vie, tout comme on peut parfois briser une promesse pour assurer l'accomplissement de la justice. L'individu doit ainsi comparer le poids relatif des différents devoirs *prima facie* impliqués dans un cas particulier afin de prendre une décision finale sur l'action à prendre. Une action est ainsi rarement, sinon jamais moralement *pure*, puisque l'on doit (presque?) toujours sacrifier un certain devoir *prima facie* pour en réaliser un plus important.

La flexibilité et la nuance permises par la démarche rossienne furent largement encensées et perçues à l'époque comme une nette amélioration par rapport au déontologisme absolutiste de Kant. Comme pour toute autre position philosophique cependant, la théorie de Ross n'est pas sans critiques.⁷⁹ Certains ont par exemple salué le pluralisme de Ross tout en reprochant au philosophe

78 Shelly Kagan (1989) a soutenu que le choix de mot de Ross n'était pas le bon et que ce qu'il avait en tête était des devoirs *pro tanto* et non *prima facie*. La différence essentielle est qu'un principe *prima facie* (c'est à dire « à première vue ») *perd sa valeur* une fois « battu » par un principe plus important, alors qu'un principe *pro tanto* préserve sa valeur même si un principe jugé plus important est adopté.

79 Je laisse de côté ici les critiques conséquentialistes affirmant que les devoirs *prima facie* de Ross peuvent ou doivent être subsumés sous le principe plus fondamental d'utilité. Pour un argumentaire sophistiqué d'un philosophe contemporain qui cherche à réconcilier les intuitions de Ross avec le conséquentialisme de la règle, voir Brad Hooker (1996, 2002).

écossais de ne pas être allé suffisamment loin au niveau de la flexibilité morale. C'est le cas de la plupart des « particularistes », à commencer par Jonathan Dancy (1983, 1993, 2004, 2009), qui juge que l'introduction du caractère *prima facie* des devoirs moraux était un pas dans la bonne direction, mais que Ross aurait dû aller plus loin encore et rejeter l'idée qu'il existe des principes moraux qui sont *invariablement* bons ou mauvais : « Particularists applaud Ross's insistence that there can be many features of the situation each of which makes some difference to how one should act; they merely want to say that the matter is not regular in the way that Ross, as a generalist, supposes » (Dancy, 2009, §5). Pour Dancy, un pluralisme acceptable doit reconnaître que les devoirs *prima facie* n'apportent pas toujours une valeur positive à une situation; selon le contexte, un devoir *prima facie* aura une valeur tantôt positive, tantôt neutre, tantôt négative. Dit autrement, il n'existe aucun principe à polarité fixe et interchangeable.⁸⁰

Une autre critique souvent soulevée est que le pluralisme de Ross offre un amalgame de devoirs n'ayant aucune véritable unité ou cohérence. On peut en effet se demander si une quelconque procédure ou règle sous-tend la liste de sept devoirs *prima facie* de Ross. Bien que Ross admette que sa liste ne se veuille pas exhaustive, il importe au moins de montrer que celle-ci n'est pas arbitraire ou aléatoire. C'est là un problème que soulève notamment D. D. Raphael : « [intuitionism] does not meet the needs of a philosophical theory, which should try to show connections and tie things up into a coherent system » (Raphael, 1981, 55). Une critique de ce type était déjà formulée à l'époque de Ross par H. W. B. Joseph, qui utilisait l'expression « heap of unconnected obligations » (Joseph, 1931, 92) pour décrire l'intuitionnisme défendu par Ross.

Dans un ordre d'idées similaire, Rawls (1971) reprochait à l'intuitionnisme normatif son absence de procédure ou règle permettant de trancher entre les différents principes moraux conflictuels. La thèse essentielle de l'intuitionnisme normatif, soutient Rawls, consiste à dire que le « poids relatif » de chaque principe moral se mesure ultimement *par intuition*, et qu'on ne peut trancher entre différents principes qu'en se référant à notre faculté intuitive.

« Intuitionist theories, then, have two features: first, they consist of a plurality of first principles which may conflict to give contrary directives in particular types of cases; and second, they include no explicit method, no priority rules, for weighing these principles against one another: we are simply to strike a balance by intuition, by what seems to us most nearly right. » (Rawls, 1971, 34)⁸¹

80 Pour une réplique intuitionniste aux arguments particularistes de Dancy et autres, voir Hooker et Little, 2000 et Audi 2006.

81 Rawls n'offre aucun véritable argument dans *A Theory of Justice* afin de montrer que l'intuitionnisme tel qu'il le décrit n'est pas une théorie acceptable. Il va d'ailleurs même jusqu'à affirmer: « Now, there is nothing intrinsically

Différents intuitionnistes ont tenté de répondre à cette critique de l'absence de règles ou procédures permettant soit d'*établir* une liste finale de principes moraux premiers, soit de *trancher* entre ces différents principes lorsqu'ils entrent en conflit. L'une des réponses les plus efficaces est offerte par David McNaughton dans « An Unconnected Heap of Duties? » (2002). Concernant l'absence de règle ou procédure pour identifier les principes premiers, McNaughton souligne que s'il est vrai que Ross ne se sert pas d'une règle ou procédure unitaire pour parvenir à cette identification – comme le font par exemple les kantien avec la règle de l'universalisation des maximes particulières – Ross a néanmoins une procédure, que McNaughton décrit comme holiste.

« The difference between [Ross's] test and the Kantian one is that the latter is atomistic, generating each principle independently of the others, whereas his is holistic, testing each principle by seeing whether it can be derived from the others. But why should a holistic test be less systematic than an atomistic one? The real worry here, I suspect, may be about not the lack of systematic unity in Ross's theory, but the perceived need for a justificatory grounding for each duty. But that is, of course, just to beg the questions against the intuitionist who maintains that these basic duties stand in no need of grounding » (McNaughton, 2002, 85).

Quant à la critique rawlsienne selon laquelle la seule manière d'assigner un « poids relatif » aux différents principes est de se référer à l'intuition individuelle, elle est en partie vraie et en partie erronée. Il est d'abord vrai qu'en dernière instance, toute détermination du poids des principes et toute tentative de trancher entre ces principes ne peut que reposer sur la faculté intuitive. Ce point est admis par tous les intuitionnistes, qui jugent qu'un tel constat s'applique également aux mathématiques, à la logique et à d'autres disciplines sans qu'on y voit un quelconque problème. Rawls lui-même admet d'ailleurs qu'on ne peut éliminer le renvoi aux intuitions pour l'évaluation de nos différents principes (Rawls, 1971, 44) ; son but est simplement de limiter autant que possible le rôle des intuitions à cet effet, et c'est dans cette optique qu'il développe sa procédure impliquant le voile d'ignorance. Or, l'identification de la procédure elle-même, de même que les décisions prises à partir de cette procédure, reposent en grande partie sur l'intuition. La procédure de Rawls n'est en effet pas indépendante des intuitions ; elle fait simplement *canaliser* les intuitions. Le même constat

irrational about this intuitionist doctrine. Indeed, it may be true. We cannot take for granted that there must be a complete derivation of our judgments of social justice from recognizably ethical principles » (Rawls, 1971, 39). Il se contente simplement de dire que c'est un *desiderata* de toute théorie morale de pouvoir offrir une procédure permettant d'établir une priorité entre les différents principes conflictuels, et il conçoit sa propre théorie comme permettant d'atteindre cet objectif.

s'applique pour la procédure kantienne d'universalisation.

S'il est vrai que Ross n'offre pas une procédure comme Rawls ou Kant, il est faux cependant de dire que l'assignation du poids des principes moraux est laissée au bon gré de chaque individu. Ross conçoit en effet une hiérarchie entre les différents principes, et bien que la reconnaissance de cette hiérarchie repose sur l'intuition, elle est conçue par Ross comme une connaissance *a priori* ayant un statut objectif plutôt que subjectif :

« [Ross] holds that the duty of non-maleficence is more important than the duty to promote a maximum of aggregate good (RG 21, 22, FE 75, 130n1), and he suggests that the duties of fidelity, reparation, and gratitude are in general more weighty than the duty to promote the good (RG 19, 30, 41-42; FE 76; OJ 124, 126). [...] Finally, although he does not say it, his view would surely be that the duty of non-maleficence is weightier than the duties of reparation, gratitude, and fidelity : it is (unless much is at stake) wrong to harm others in order to fulfill these duties » (Skelton, 2012, §4.1)

Cette hiérarchie entre les différents principes⁸² nous sert de guide dans la pratique morale. Il faut cependant préciser que la supériorité *in abstracto* d'un principe sur un autre n'implique guère que l'on doive *toujours* agir en fonction du principe supérieur *in abstracto*. Dans certains cas, l'action à prendre peut être celle qui correspond à un principe qui est pourtant inférieur lorsque considéré indépendamment de tout contexte. McNaughton voit dans cette position un juste milieu entre le particularisme, qui soutient que les devoirs sont *entièrement* déterminés par le contexte, et l'absolutisme, qui soutient que les devoirs sont complètement *indépendants* du contexte :

« That midway position is intended to allow us to say something about the ranking of duties in general which falls short of absolutism: namely that some kinds of basic duty might be thought to be, in their intrinsic nature, more weighty than others. This does not mean that the less weighty duty will never win out, only that it starts with an initial handicap which it will have to work hard to overcome. In deciding what to do, it seems, we must take into account not only how weighty an instance of each particular duty we have in the case before us, but also the *general* weight that is to be given to each of these duties » (McNaughton, 2002, 86-87).

Il existe plusieurs autres objections à l'intuitionnisme normatif et contre-objections soulevées par les défenseurs de cette théorie, mais je me contenterai ici de ce portrait sommaire du

82 Ross offre également une hiérarchie entre les différents *biens intrinsèques* au chapitre VI de *The Right and the Good* (1930). Il affirme par exemple : « With regard to pleasure and virtue, it seems to me much more likely to be the truth that *no* amount of pleasure is equal to any amount of virtue, that in fact virtue belongs to a higher order of value, beginning at a point higher on the scale of value than that which pleasure ever reaches » (Ross, 1930, 150).

débat contemporain. J'ai suffisamment présenté les deux formes d'intuitionnisme dans les deux dernières sections afin de pouvoir aborder la dernière question de ce mémoire : l'intuitionnisme et la méthode de l'équilibre réfléchi sont-ils réconciliables?

5.3. Intuitionnisme et méthode de l'équilibre réfléchi

Le problème de la compatibilité entre l'intuitionnisme éthique et la méthode de l'équilibre réfléchi ne saurait obtenir une réponse directe et généralisée. Cela s'explique bien sûr par le fait qu'il existe différentes formes d'intuitionnisme et différentes façons de concevoir la méthode de l'équilibre réfléchi. On ne peut dès lors aborder la question de la compatibilité sans avoir d'abord précisé de quel type d'intuitionnisme et d'équilibre réfléchi on parle. Ce constat nous permet également de comprendre pourquoi certains philosophes (DePaul, 1986; Ebertz 1993; Tappolet 2000) soutiennent qu'intuitionnisme et équilibre réfléchi sont tout à fait compatibles, tandis que d'autres (Rawls, 1971; Daniels, 1980b) les voient au contraire comme mutuellement exclusifs : leur désaccord s'explique en partie par le fait qu'ils ont en tête des catégories différentes.

J'ai présenté dans les deux sections précédentes les deux grandes formes d'intuitionnisme éthique : l'intuitionnisme méta-éthique et l'intuitionnisme normatif. Quant aux différentes formes de l'équilibre réfléchi, on peut bien sûr souligner la différence entre les équilibres réfléchis *large* et *étroit*. Cependant, ces deux versions ne diffèrent guère en ce qui attrait à leur compatibilité avec l'intuitionnisme, puisque l'une n'est qu'une version plus sophistiquée de l'autre. Je souhaite ici proposer une différente typologie de la méthode de l'équilibre réfléchi, typologie que j'ai indirectement introduite au chapitre 4. On peut essentiellement distinguer deux grandes formes ou usages possibles de la méthode de l'équilibre réfléchi : on peut l'aborder soit comme une manière de *construire* des théories, soit comme une manière de *justifier* des théories ou principes. Une fois que l'on a devant nous les deux formes de l'intuitionnisme et les deux façons de concevoir la méthode de l'équilibre réfléchi, on est plus à même de déterminer si celles-ci sont compatibles ou non.

J'apporte d'abord quelques précisions sur ces deux façons de concevoir l'équilibre réfléchi. La distinction entre ces deux formes de la méthode est implicite dans le débat sur les interprétations cohérentistes et fondationnalistes de l'équilibre réfléchi. Tel que souligné à la section 4.2.2, Michael DePaul soutient qu'une interprétation fondationnaliste de l'équilibre réfléchi est tout à fait concevable dans la mesure où l'équilibre réfléchi n'est ultimement qu'une méthode permettant de *construire* des théories : « Reflective equilibrium is a method of theory construction [while

foundationalism is] a type of account of the epistemic status of our beliefs » (DePaul, 1986, 68). Or, si rien n'empêche que l'on aborde l'équilibre réfléchi comme une simple démarche de construction théorique qui demeure agnostique sur les enjeux épistémologiques, il est cependant clair que ce n'est pas ainsi que Rawls et Daniels concevaient leur méthode. Les deux auteurs soulignent en effet à plusieurs reprises qu'ils conçoivent l'équilibre réfléchi comme un modèle de justification cohérentiste de théories et de principes moraux.⁸³ Or, ce serait à mon sens une erreur de chercher à voir qui a raison entre DePaul et Rawls/Daniels, puisqu'on peut tout à fait accepter l'existence d'une pluralité de conceptions : on peut tout à fait se servir de la méthode sans aucune prétention épistémique, à des simples fins de construction théorique, mais on peut également s'en servir à des fins de justification. Les deux façons de concevoir la méthode sont acceptables.

En concevant la méthode de l'équilibre réfléchi de la première façon, c'est à dire comme une manière de construire des théories sans leur accorder un statut épistémique particulier, on constate que celle-ci est parfaitement compatible avec chacune des deux versions de l'intuitionnisme. En effet, comme l'explique DePaul, aborder l'équilibre réfléchi de cette façon revient simplement à prendre en considération l'ensemble de nos croyances, de juger de leur plausibilité respective, et de rechercher une cohérence entre elles. Or, une telle démarche est compatible avec l'intuitionnisme normatif puisque l'on peut arriver à la fin de ce processus réflexif et conclure qu'il existe une pluralité de principes moraux irréductibles. Si l'on inclut la thèse, centrale à la critique rawlsienne, de l'impossibilité de trancher entre ces principes moraux sinon que par intuition, on ne se retrouve pas non plus confrontés à une forme d'incompatibilité, puisqu'il s'agit encore là d'une conclusion à laquelle on pourrait arriver suite à un long processus d'équilibrage entre nos jugements, principes et théories d'arrière-fond.

Cette façon de concevoir la méthode est également compatible avec l'intuitionnisme méta-éthique. En effet, rien n'empêche que l'on accepte, à l'intérieur d'un tel processus réflexif, que certaines croyances sont justifiées de manière non-inférentielle. Sans doute pourrait-on répliquer que l'intuitionnisme méta-éthique, en *s'arrêtant* à des principes jugés évidents en eux-mêmes, commet une forme d'irrationalité en ne poussant pas davantage la réflexion ou en accordant un statut épistémique privilégié à certains principes, ce qui serait, pourrait-on soutenir, incompatible avec la méthode de l'équilibre réfléchi. Cette critique repose sur la proposition de DePaul selon laquelle la méthode de l'équilibre réfléchi est la seule démarche rationnelle pouvant être adoptée (voir 4.3), dans la mesure où la méthode enjoint à n'ignorer aucune croyance et à pousser à sa limite

83 Voir par exemple Rawls, 1971, 579 et Daniels, 1979, 264-5.

le raisonnement sur un enjeu (DePaul, 1998, 304) Or, on pourrait soutenir que l'intuitionnisme méta-éthique ne respecte pas cette démarche et commet une forme d'irrationalité en s'arrêtant à des principes jugés évidents en eux-mêmes. Il est cependant clair qu'une telle position n'est pas tenable, puisqu'il n'est pas irrationnel, au sens décrit par DePaul, de s'arrêter à des principes auxquels on accorde un statut particulier. En effet, les intuitionnistes n'affirment pas *dogmatiquement* leur épistémologie et sont au contraire en dialogue constant avec d'autres philosophes dans le but de tester cette épistémologie. Dit autrement, ce n'est pas que les intuitionnistes *refusent* de creuser davantage une fois les principes évidents en eux-mêmes identifiés, mais c'est simplement qu'ils soutiennent qu'il n'est *pas nécessaire* d'aller plus loin une fois ces principes identifiés. Il n'y a donc rien d'irrationnel dans la démarche de l'intuitionnisme méta-éthique, et la théorie est ainsi tout à fait compatible avec la première façon de concevoir la méthode de l'équilibre réfléchi.

Reste donc la deuxième forme de l'équilibre réfléchi, à savoir celle qui conçoit la méthode comme une démarche de justification cohérentiste. C'est clairement cette façon particulière de concevoir la méthode de l'équilibre réfléchi qui semble entrer en conflit avec l'intuitionnisme, et plus précisément avec l'intuitionnisme méta-éthique. Il n'y a en effet aucune incompatibilité avec l'intuitionnisme normatif, dans la mesure où l'on peut tout à fait justifier sur une base cohérentiste l'idée qu'il existe une pluralité de principes moraux premiers pouvant entrer en conflit. La seule incompatibilité possible se situe donc entre la deuxième forme de l'équilibre réfléchi et l'intuitionnisme méta-éthique.

Il est facile de voir en quoi ces deux versions particulières peuvent être dites mutuellement exclues. Pour Rawls et Daniels, c'est uniquement l'atteinte d'un équilibre et d'une cohérence entre nos différents principes et jugements (et théories d'arrière-fond pour Daniels) qui permet de conférer une justification, alors que pour l'intuitionniste méta-éthique, certains principes élémentaires sont *ipso facto* justifiés de par leur nature même – soit parce qu'ils sont évidents en eux-mêmes (intuitionnisme *a priori*), soit parce qu'ils sont des objets de perception (intuitionnisme *a posteriori*). Or, dans la mesure où les deux approches impliquent des modèles de justification antagonistes, elles paraissent effectivement irréconciliables.

On doit cependant noter que les intuitionnistes méta-éthiques ont tendu une branche d'olivier aux cohérentistes de l'équilibre réfléchi. La plupart des intuitionnistes reconnaissent en effet que la cohérence est une condition *sine qua non* de la justification; Sidgwick fait de la cohérence l'un de ses quatre critères de l'évidence (e.g., Sidgwick, 1874, 72), Ross souligne à maintes reprises la nécessité d'atteindre une harmonie dans nos opinions (e.g., Ross, 1939, 3) et Audi affirme que

même nos principes évidents en eux-mêmes peuvent perdre leur justification dans la recherche d'un équilibre réfléchi (Audi, 2002, 49). Tout ce que les intuitionnistes contestent est l'idée que *seule* la cohérence confère de la justification. Le cohérentiste de l'équilibre réfléchi peut-il reconnaître qu'autre chose que la cohérence puisse apporter une valeur justificatrice? La thèse de l'*exclusivité* justificatrice de la cohérence est-elle réellement intouchable pour lui? Évidemment, si l'on définit le cohérentisme comme la thèse selon laquelle la cohérence est une condition nécessaire et *suffisante* pour la justification, il est alors faux *par définition* d'affirmer que le cohérentisme peut reconnaître d'autres sources de justification.

Il y a cependant de bonnes raisons de croire que la cohérence ne constitue pas l'alpha et l'oméga de la justification. Après tout, même lorsqu'un état d'équilibre réfléchi est atteint, les croyances en équilibre n'ont pas toutes le même niveau de plausibilité ou de crédibilité, ce qui suggère que la cohérence ne suffit pas à expliquer ce qui confère une valeur épistémique à nos croyances. Bien évidemment, cet argument à lui seul ne risque guère de convertir le cohérentiste au fondationnalisme, et considérant l'ampleur de ce débat en épistémologie, même un mémoire entier ne saurait rendre justice aux nombreux arguments pouvant être déployés par le cohérentiste. Je me contenterai donc de conclure que le mariage *complet* entre intuitionnisme et équilibre réfléchi ne saurait avoir lieu, puisque les cohérentistes s'opposent farouchement à l'union de leur méthode justificatrice avec l'intuitionnisme méta-éthique.

5.4. Conclusion

J'ai présenté dans ce chapitre deux grandes versions de l'intuitionnisme éthique. J'ai débuté par l'intuitionnisme méta-éthique, qui affirme que certains principes moraux sont justifiés de manière non-inférentielle. J'ai distingué deux versions différentes de cette approche: une version *a posteriori*, qui situe la justification non-inférentielle dans le fait de percevoir certains faits moraux, et une version *a priori*, qui conçoit plutôt la justification non-inférentielle comme le fait que certaines propositions soient évidentes en elles-mêmes. J'ai abordé certaines objections soulevées envers cette dernière approche, en montrant dans quelles mesures ces objections reposaient en grande partie sur une mauvaise compréhension de l'idée d'évidence en soi utilisée par les intuitionnistes *a priori*. J'ai ensuite abordé la deuxième forme d'intuitionnisme, à savoir l'intuitionnisme normatif, qui soutient qu'il existe une pluralité de principes moraux premiers pouvant entrer en conflit. J'ai centré mon attention sur la version de la théorie défendue de W. D.

Ross, et j'ai souligné les critiques les plus importantes, notamment la critique particulariste qui reproche à Ross de concevoir les devoirs *prima facie* comme ayant une polarité fixe.

Après avoir présenté les principales thèses et les principaux arguments associés à chaque version de l'intuitionnisme, j'ai pu retourner à la question qui avait conclu le chapitre 4, à savoir dans quelle mesure l'intuitionnisme et la méthode de l'équilibre réfléchi sont compatibles. Après avoir distingué entre deux grandes façons de concevoir la méthode de l'équilibre réfléchi, j'ai conclu que la seule incompatibilité envisageable était entre l'intuitionnisme méta-éthique et la conception de l'équilibre réfléchi comme modèle de justification cohérentiste.

La compatibilité entre intuitionnisme et équilibre réfléchi constituait la dernière grande question de mon mémoire. Je me tourne sans plus tarder vers la conclusion de mon mémoire.

6. CONCLUSION

J'ai cherché dans ce mémoire à répondre à trois grandes questions concernant les intuitions utilisées en philosophie morale : « que sont ces intuitions? », « ces intuitions sont-elles fiables? » et « quel est leur rôle en philosophie morale contemporaine »?

La première question était l'objet du chapitre 2. Après avoir souligné que le concept d'intuition était fort difficile à cerner, point que j'ai illustré notamment en présentant une dizaine de caractérisations différentes du concept, j'ai cherché à délimiter autant que possible les contours du concept tel qu'utilisé en philosophie. Cela m'a amené d'abord à explorer l'usage populaire du concept (2.1), en remarquant notamment qu'en dépit d'une certaine compréhension « ésotérique » de l'intuition, il existait dans le langage populaire une compréhension de l'intuition qui se rapproche de l'usage philosophique, à savoir l'intuition comme une forme d'impression spontanée ou de croyance à première vue. J'ai ensuite offert une analyse étymologique du concept (2.2), ce qui m'a entre autres amené à constater que celui-ci avait des origines augustinienne et renvoyait à une forme de perception de l'esprit. À la section suivante (2.3), j'ai cherché à démystifier l'intuition en montrant que, contrairement à ce que laissait entendre Mackie (1977), les intuitions et la faculté intuitive n'ont rien d'étranges. J'ai cherché à renforcer ce point à la section suivante (2.4) en montrant que l'on peut « naturaliser » l'intuition en l'identifiant simplement au « système 1 » dans la théorie à processus duaux développée en sciences cognitives. Avant de conclure le chapitre, j'ai apporté quelques distinctions entre l'intuition, l'émotion et le raisonnement, soulignant que les trois notions, bien que reliées, renvoient à des phénomènes différents et ne doivent pas être confondues.

La deuxième question était l'objet du chapitre 3. Plutôt que de présenter des raisons pouvant nous amener à croire en la fiabilité des intuitions, ma stratégie dans ce chapitre a consisté à considérer différentes raisons de *douter* de la fiabilité des intuitions et à offrir une réponse à chacune de ces raisons. J'ai distingué trois types de raisons pouvant nous faire douter de la fiabilité des intuitions. Dans un premier temps, j'ai considéré les principales recherches scientifiques pouvant nous amener à mettre en doute la fiabilité des intuitions (3.1). J'ai d'abord considéré différentes recherches en économie comportementale (3.1.1) qui illustrent de nombreuses formes d'évaluations irrationnelles des agents. J'ai conclu que ces études doivent nous inciter à la prudence dans l'utilisation de notre faculté intuitive, mais qu'elles ne viennent aucunement délégitimer l'usage de cette faculté – pas plus que des études sur les illusions d'optique, par exemple, pourraient délégitimer l'usage de notre faculté visuelle. J'ai ensuite considéré un deuxième type d'études

scientifiques, qui portaient cette fois-ci sur les jugements de catégorisation et sur la nature des concepts (3.1.2). J'ai conclu que ces études ne faisaient que confirmer un constat déjà largement accepté en philosophie, à savoir qu'il est illusoire de rechercher des critères nécessaires et suffisants derrière nos concepts évaluatifs. Un tel constat, j'ai souligné, est tout à fait compatible avec l'usage contemporain des intuitions en philosophie.

Dans un deuxième temps, j'ai considéré une autre famille de raisons pouvant nous amener à douter de la fiabilité des intuitions, à savoir les nombreuses critiques soulevées et les recherches effectuées par les partisans du mouvement de philosophie expérimentale (3.2). Le mouvement X-Phi soutient essentiellement que les intuitions soulevées par les philosophes lors d'expériences de pensée sont loin d'obtenir l'assentiment universel que ceux-ci supposent, et plusieurs recherches effectuées par les philosophes expérimentaux ont cherché à démontrer ce point. Ma réponse a consisté d'une part à souligner l'existence de failles importantes dans la démarche des expérimentaux, et d'autre part à montrer que certaines de leurs études tendaient en fait à démontrer qu'il existait bel et bien un noyau intuitif universel, et qu'un tel noyau était suffisant pour justifier un certain usage des intuitions en philosophie. J'ai conclu que le mouvement X-Phi, malgré ses failles méthodologiques, apportait des contributions importantes au débat philosophique contemporain qui pouvaient aider à améliorer la méthode philosophique traditionnelle.

Dans un troisième temps, j'ai exploré le problème de l'origine des intuitions comme raison de douter de leur fiabilité (3.3). J'ai distingué trois sources possibles des intuitions : la socialisation (3.3.1), les théories tacites (3.3.2) et les théories explicites (3.3.3). J'ai reconnu que si nos intuitions provenaient entièrement de la socialisation et résultaient de l'intériorisation de normes sociales *arbitraires*, cela pouvait mettre en doute leur fiabilité, dans la mesure où le renvoi aux intuitions à des fins de justification reviendrait alors à une forme de chauvinisme. J'ai également reconnu que les intuitions résultant de l'intériorisation de théories explicites n'étaient guère plus fiables que celles résultant de la socialisation, et j'ai ajouté que les intuitions de certains experts étaient souvent moins fiables que les intuitions pré-théoriques de non-experts compétents. J'en suis ainsi arrivé à l'idée que la seule source non-problématique de nos intuitions était les théories tacites, mentionnant entre autres que les normes sociales reposant sur de telles théories tacites n'avaient rien d'arbitraires. Les intuitions résultant de théories tacites peuvent être comprises comme le produit de certains mécanismes cognitifs universels; des théories comme l'intuitionnisme social de Jonathan Haidt et la théorie de la grammaire morale universelle offrent une manière de concevoir certaines intuitions comme étant véritablement constitutives de la nature humaine. C'est là à mon sens la meilleure

façon de concevoir la source de plusieurs de nos intuitions, tout spécialement nos intuitions morales fondamentales. J'ai souligné rapidement que le modèle réaliste offrait une autre façon de fonder la fiabilité de nos intuitions, mais que je préférais l'explication renvoyant à des théories tacites, essentiellement pour des raisons métaphysiques.

Après avoir offert une réponse aux trois principales sources de doute concernant la fiabilité des intuitions, j'ai conclu qu'il était faux d'affirmer que les intuitions n'étaient tout simplement pas fiables; s'il est vrai que certaines études empiriques montrent que notre faculté intuitive n'est pas toujours efficace, ce constat doit nous amener non pas à cesser d'utiliser cette faculté – ce qui serait de toute façon impossible en philosophie – mais plutôt à cultiver une certaine prudence dans son usage. Or, l'utilisation des intuitions en philosophie morale contemporaine témoigne d'un tel souci de prudence, notamment par l'importance accordée à la cohérence et à la recherche d'un « équilibre réfléchi » entre nos différentes croyances.

Les deux derniers chapitres m'ont amené à répondre à la troisième grande question de mon mémoire en explorant les différents usages faits des intuitions en philosophie morale contemporaine. J'ai distingué trois rôles principaux des intuitions : les intuitions sont utilisées (1) dans l'élaboration d'expériences de pensée, (2) dans l'application de la méthode de l'équilibre réfléchi et (3) dans le développement de la théorie de l'intuitionnisme éthique. J'ai abordé conjointement les deux premiers usages, puisque ceux-ci sont généralement combinés en philosophie morale contemporaine.

Le chapitre 4 a ainsi porté sur ce que j'ai appelé la « méthodologie standard » en philosophie morale contemporaine, qui consiste en l'identification de nos intuitions sur différents enjeux par le recours à des expériences de pensée, puis en la tentative de systématiser ces intuitions par la méthode de l'équilibre réfléchi. J'ai commencé par présenter quelques exemples classiques d'expériences de pensée développées en philosophie (4.1.1), puis je me suis interrogé sur la nature et la valeur épistémique de ce type d'exercice intellectuel (4.1.2). J'ai abordé ces enjeux en me concentrant sur le débat contemporain opposant l'interprétation empiriste de John Norton et l'interprétation rationaliste de James Brown. J'ai conclu que le débat sur ces enjeux se réduisait essentiellement à un débat entre réalisme et anti-réalisme, et que la plupart des questions soulevées à l'endroit des expériences de pensée étaient des questions concernant la fiabilité des intuitions – sujet largement discuté au chapitre 3.

J'ai ensuite abordé la méthode de l'équilibre réfléchi, véritable pilier en philosophie morale contemporaine (4.2). J'ai présenté la version standard de la méthode telle que développée par John

Rawls dans *A Theory of Justice* (1971), et j'ai souligné les différentes critiques qui lui étaient souvent associées, notamment le fait que la méthode permette uniquement de décrire et non de justifier nos jugements et principes moraux (4.2.1). J'ai ensuite présenté la méthode de l'équilibre réfléchi *large* développée par Norman Daniels comme une tentative de répondre aux faiblesses associées à la méthode rawlsienne (4.2.2). La méthode de Daniels consiste à inclure dans l'équilibre réfléchi les « théories d'arrière-fond » qui sous-tendent nos jugements et principes. L'inclusion de ces théories d'arrière-fond, soutient Daniels, offre une nouvelle manière de concevoir la justification conférée par la méthode développée par Rawls. J'ai ensuite abordé la possibilité de réconcilier la méthode décrite par Daniels avec un modèle de justification fondationnaliste. Enfin, j'ai considéré différents autres modèles ayant été développés à partir de la méthode standard de Rawls, notamment un modèle visant à limiter l'équilibre réfléchi légitime à certains experts ou individus compétents, et un autre visant à étendre l'équilibre réfléchi à des collectivités plutôt qu'à des individus (4.2.3). Aucun de ces modèles, j'ai conclu, n'était préférable au modèle développé par Daniels, qui pouvait même être considéré comme un modèle *par défaut* en philosophie.

Le chapitre 5 a porté sur une manière alternative d'aborder les intuitions, soit en les concevant à la manière de l'intuitionnisme éthique. J'ai distingué deux grandes formes d'intuitionnisme éthique, à savoir l'intuitionnisme méta-éthique (5.1) et l'intuitionnisme normatif (5.2). L'intuitionnisme méta-éthique soutient que certaines croyances sont justifiées de manière non-inférentielle, c'est à dire qu'elles n'ont besoin de rien d'autre qu'elles-mêmes pour être justifiées. J'ai distingué entre une version *a posteriori* de cette théorie, concevant les intuitions auto-justifiées comme étant des perceptions de faits moraux, et une version *a priori*, concevant plutôt ces intuitions comme étant justifiées sur une base conceptuelle. Cette dernière version, qui est la plus répandue, repose sur l'idée que certains principes moraux sont « évidents en eux-mêmes ». J'ai présenté différentes manières de concevoir ce concept d'évidence en soi, et j'ai abordé les différentes critiques soulevées envers cette notion, soulignant que la plupart de ces objections reposaient sur une mauvaise compréhension du concept utilisé.

J'ai ensuite abordé l'intuitionnisme normatif, qui se résume à la thèse qu'il existe une pluralité de principes moraux irréductibles pouvant entrer en conflit. J'ai présenté la version classique de cette théorie, à savoir celle de W. D. Ross, qui défend l'idée d'une pluralité de devoirs *prima facie*. J'ai souligné les principales critiques soulevées envers la théorie rossienne, notamment la critique particulariste reprochant à Ross de concevoir l'existence de principes moraux à polarité fixe et la critique rawlsienne concernant l'absence de règle ou procédure permettant de trancher

entre les différents principes conflictuels. J'ai cherché à défendre la théorie de Ross contre ces différentes critiques, en affirmant notamment que Ross proposait une hiérarchie des différents principes et que sa démarche constituait un juste milieu entre le particularisme, qui soutient que les valeurs sont *entièrement* déterminées par le contexte, et l'absolutisme, qui soutient que les valeurs sont complètement *indépendantes* du contexte. J'ai finalement conclu le chapitre en me demandant dans quelle mesure l'intuitionnisme éthique et la méthode de l'équilibre réfléchi pouvaient être dits compatibles. Après avoir distingué deux façons de concevoir la méthode de l'équilibre réfléchi, j'ai conclu qu'une seule de ces approches, à savoir celle qui conçoit la méthode comme une forme de justification cohérentiste, pouvait entrer en conflit avec l'intuitionnisme méta-éthique.

Pour finir, s'il y a un constat général que je peux tirer de ce mémoire, c'est que les intuitions sont là pour rester en philosophie morale. En effet, en prenant conscience de l'impossibilité d'étudier le phénomène moral strictement sur une base empirique – aucune recherche scientifique ne parviendra jamais à découvrir « le juste » ou « le bien » – on arrive rapidement à la conclusion que l'intuition morale est tout simplement incontournable. Même l'un des plus farouches opposants à l'utilisation des intuitions en philosophie, Robert Cummins, admet ne pas savoir ce qui pourrait bien remplacer les intuitions en philosophie morale (Cummins, 1998, 126). Je soupçonne par ailleurs que plusieurs des « éliminativistes de l'intuition » ne saisissent pas pleinement la nature et le rôle de l'intuition, sans quoi ils admettraient son inévitabilité en philosophie morale. Pour illustrer ce point, je conclus comme j'ai débuté ce mémoire, soit avec une anecdote.

Il y a environ un an, je participais au 79^e congrès de l'ACFAS à Sherbrooke. Ma présentation avait pour titre « Intuitionnisme Social et Relativisme Culturel ». Une heure à peine avant ma présentation, je discutais avec un autre participant à la conférence qui était farouchement opposé à l'utilisation des intuitions en philosophie morale. Il affirmait notamment que les mathématiques avaient commencé à progresser le jour où elles s'étaient débarrassées des intuitions, et que la philosophie morale devrait en faire de même. Je lui ai alors demandé à quoi pourrait bien ressembler une philosophie morale dépourvue d'intuitions, et il m'a donné un exemple du type d'exercice intellectuel que la philosophie morale devrait selon lui développer. Son exemple, qui cherchait à fonder un principe de justice, allait comme suit : « Imagine deux jumeaux parfaitement identiques à tous les niveaux. Un gâteau tombe du ciel et ils doivent le partager en deux. Or, si les propriétés morales surviennent sur les propriétés non-morales, alors la seule manière acceptable de partager le gâteau est de le diviser en deux parts égales. On a là un principe de justice qui ne dépend d'absolument aucune intuition ».

Si mon exposé au cours des cent dernières pages était le moins efficace, le lecteur devrait constater que l'exemple offert par mon interlocuteur repose en fait entièrement sur l'intuition : c'est uniquement sur la base de notre intuition que l'on parvient à reconnaître le principe de justice que l'exemple – qui est en fait une expérience de pensée – cherche à illustrer. Si c'est à cela que ressemblerait une philosophie morale dépourvue d'intuitions, force est de constater qu'elle aurait des airs bien familiers.

7. BIBLIOGRAPHIE

- ALEXANDER, J. et J. M. Weinberg (2007) « Analytic Epistemology and Experimental Philosophy », *Philosophy Compass* Vol. 2 (1), 56-80.
- ARISTOTE (1997) *Topics : Books I and VII*, New York : Oxford University Press.
- ARTEMIS, A. (2011) *Navigating by Intuition : How to Follow the Signs*, Synchronicity Publishing.
- AUDI, R. (1997) *Moral Knowledge and Ethical Character*, New York : Oxford University Press.
- AUDI, R. (2002) « Prospects for a Value-Based Intuitionism » dans Stratton-Lake (dir.) *Ethical Intuitionism : Re-Evaluations*, New York : Oxford University Press, 29-58.
- AUDI, R. (2005) *The Good in the Right : A Theory of Intuition and Intrinsic Value*, Princeton : Princeton University Press.
- AUDI, R. (2006) *Practical Reasoning and Ethical Decision*, Londres: Routledge.
- AUDI, R. (2008) « Intuition, Inference, and Rational Disagreement in Ethics », *Ethical Theory and Moral Practice*, Vol. 11 (5), 475-492.
- BALGUY, J (1728) *The Foundation of Moral Goodness*, New York: Garland (1978).
- BEALER, G. (1998) « Intuition and the Autonomy of Philosophy » dans DePaul et Ramsey (dir.) *Rethinking Intuition : The Psychology of Intuition and its Role in Philosophical Inquiry*, Rowman Littlefield Publishers, 201-221.
- BEDKE, M. S. (2008) « Ethical Intuitions : What They Are, What They Are Not, and How They Justify », *American Philosophical Quarterly*, Vol. 45 (3), 253-269.
- BEN-ZE'EV, A. (2009) « The Thing Called Emotion » dans Goldie (dir.) *The Oxford Handbook of Philosophy of Emotion*, Oxford : Oxford University Press, 41-62.
- BIARD (2011) « Intuitive and Abstractive Cognition » dans Laguarda (dir.) *Encyclopedia of Medieval Philosophy*, Springer, 568-571.
- BRANDT, R. (1979) *A Theory of the Good and the Right*, Clarendon Press.
- BRENDEL, E. (2004) « Intuition Pumps and the Proper Use of Thought Experiments », *Dialectica*, Vol. 58, 88-108.
- BROAD (1930) *Five Types of Ethical Theory*, New York : Harcourt, Brace et Co.
- BROWN, J. (1993) *The Laboratory of the Mind : Thought Experiments in the Natural Sciences*, New York : Routledge.

BROWN, J. et Y. Feige (2011) « Thought Experiments », *The Stanford Encyclopedia of Philosophy* [En ligne] : <http://plato.stanford.edu/entries/thought-experiment>

BUZZONI, M. (2008) *Esperimento Ed Esperimento Mentale*, Milan : FrancoAngeli.

CHOMSKY, N. (1965) *Aspects of the Theory of Syntax*, Cambridge : MIT Press.

CHOMSKY, N. (1971) *Studies on Semantics in Generative Grammar*, Mouton de Grutyer.

CLARKE, S. (1728) « A Discourse of Natural Religion » dans Raphael (dir.), *British Moralists 1650-1800*, Hackett (1990).

CONEE, E. et R. Feldman, (1983) « Stich and Nisbett on Justifying Inference Rules », *Philosophy of Science*, Vol. 50 (2), 326-331.

CRISP, R. (2002) « Sidgwick and the Boundaries of Intuitionism » dans Stratton-Lake (dir.) *Ethical Intuitionism : Re-Evaluations*, New York : Oxford University Press, 56-75

CUDWORTH, R. (1731) *A Treatise Concerning Eternal and Immutable Morality*, Hutto (éd.. 1996), Cambridge University Press.

CULLEN, S (2010) « Survey-Driven Romanticism », *Review of Philosophy and Psychology*, Vol. 1 (2), 275-296.

CUMBERLAND, R. (1672) *De Legibus Naturae Disquisitio Philosophica*, Kessinger Publishing (2009).

CUMMINS, R. (1998) « Reflection on Reflective Equilibrium » dans DePaul et Ramsey (dir.) *Rethinking Intuition : The Psychology of Intuition and its Role in Philosophical Inquiry*, Rowman Littlfield Publishers, 112-124.

DANCY, J. (1983) « Ethical Particularism and Morally Relevant Properties », *Mind*, Vol. 92 (368), 530-547.

DANCY, J. (1993) *Moral Reasons*, Oxford : Blackwell Publishing.

DANCY, J. (2004) *Ethics Without Principles*, Oxford : Oxford University Press.

DANCY, J. (2009) « Moral Particularism », *Stanford Encyclopedia of Philosophy*, [En ligne] : <http://plato.stanford.edu/entries/moral-particularism>.

DANIELS, N. (1979) « Wide Reflective Equilibrium and Theory Acceptance in Ethics », *The Journal of Philosophy*, Vol 76 (5), 256-282.

DANIELS, N. (1980a) « On Some Methods of Ethics and Linguistics », *Philosophical Studies : An International Journal for Philosophy in the Analytic Tradition*, Vol. 37 (1), 21-36.

DANIELS, N. (1980b) « Reflective Equilibrium and Archimedian Points », *Canadian Journal of*

Philosophy, Vol. 10 (1), 83-103.

DARWALL, S. (2002) « Ethical Intuitionism and the Motivation Problem » dans Stratton-Lake (dir.) *Ethical Intuitionism : Re-Evaluations*, New York : Oxford University Press, 248-270.

DEPAUL, R. M. (1986) « Reflective Equilibrium and Foundationalism », *American Philosophical Quarterly*, Vol. 23 (1), 59-69.

DEPAUL, M. (1998) « Why Bother With Reflective Equilibrium ? », dans DePaul et Ramsey (dir.) *Rethinking Intuition : The Psychology of Intuition and its Role in Philosophical Inquiry*, Rowman Littlfield Publishers, 293-310.

DEPAUL, R. M. et Ramsey, (1998) « Preface » dans DePaul et Ramsey (dir.) *Rethinking Intuition : The Psychology of Intuition and its Role in Philosophical Inquiry*, Rowman Littlfield Publishers.

DESOUSA, R. (2009) « The Mind's Bermuda Triangle : Philosophy of Emotions and Empirical Science » dans Goldie (dir.) *The Oxford Handbook of Philosophy of Emotion*, Oxford : Oxford University Press, 95-119.

DE WAAL, F. (2009) *The Age of Empathy : Nature's Lessons for a Kinder Society*, New York : Crown Publishing.

DILLARD, S. (2008) *Discover Your Psychic Type : Developing and Using Your Natural Intuition*, Llewellyn Publications.

DUHEM, P. (1913) *La théorie physique: son objet et sa structure*, Paris: Chevalier & Rivière.

DWYER, S. (2008) « How Not to Argue that Morality Isn't Innate: Comments on Prinz », dans Sinnott-Armstrong (dir.) *Moral Psychology, Vol. 2.*, Boston : MIT Press, 407-418.

EBERTZ, R. (1993) « Is Reflective Equilibrium a Coherentist Model ? », *Canadian Journal of Philosophy*, Vol. 23 (2), 193-214.

EVANS, J. (2008) « Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition », *Annual Review of Psychology*, Vol. 59, 255-278.

EWING, A. C. (1947) *The Definition of Good*, Londres : Routledge.

FEHIGE, J. (2009) « Thought Experimenting with God: Revisiting the Ontological Argument », *Neue Zeitschrift für Systematische Theologie und Religionsphilosophie*, Vol. 51, 249-267.

FOLEY, R. (1998) « Rationality and Intellectual Self-Trust », dans DePaul et Ramsey (dir.) *Rethinking Intuition : The Psychology of Intuition and its Role in Philosophical Inquiry*, Rowman Littlfield Publishers, 241-257.

FOOT, P. (1978) *Virtues and Vices and Other Essays in Moral Philosophy*. Berkeley et Los Angeles : University of California Press.

- GAUT, B. (2002) « Justifying Moral Pluralism », dans Stratton-Lake (dir.) *Ethical Intuitionism : Re-Evaluations*, New York : Oxford University Press, 137-160.
- GENDLER, T. (1998) « Galileo and the Indispensability of Scientific Thought Experiment », *The British Journal for the Philosophy of Science*, Vol. 49, 397-424.
- GETTIER, E. L. (1963) « Is Justified True Belief Knowledge ? », *Analysis*, Vol. 23, 121-123.
- GIBBARD, A. (2002) « Knowing What to Do, Seeing What to Do », dans Stratton-Lake (dir.) *Ethical Intuitionism : Re-Evaluations*, New York : Oxford University Press, 212-228.
- GIROUX, J. (2011) « The Origin of Moral Norms : A Moderate Nativist Account », *Dialogue : Canadian Philosophical Review*, Vol. 50 (2), 381-406.
- GOLDMAN, A. et J. Pust (1998) « Philosophical Theory and Intuitional Evidence » dans DePaul et Ramsey (dir.) *Rethinking Intuition : The Psychology of Intuition and its Role in Philosophical Inquiry*, Rowman Littlfield Publishers, 179-200.
- GOODMAN, N. (1955) *Fact, Fiction, and Forecast*, Cambridge : Harvard University Press.
- GOPNICK, A. et E. Schwitzgebel (1998) « Whose Concepts Are They, Anyway ? The Role of Philosophical Intuition and Cognitive Mechanism » dans DePaul et Ramsey (dir.) *Rethinking Intuition : The Psychology of Intuition and its Role in Philosophical Inquiry*, Rowman Littlfield Publishers, 75-94.
- GRAHAM, G. et T. Horgan (1998), « Southern Fundamentalism and the End of Philosophy » dans DePaul et Ramsey (dir.) *Rethinking Intuition : The Psychology of Intuition and its Role in Philosophical Inquiry*, Rowman Littlfield Publishers, 271-292.
- GREENE, J. (2007) « Why Are VMPFC Patients Utilitarian ? : A Dual-Process Theory of Moral Judgment Explains », *Trends in Cognitive Sciences*, Vol. 11 (8), 322-323.
- GREENE, J., R. Sommerville, R. B. Nystrom, L.E., Darely, J. M et J. D. Cohen (2001) « An fMRI Investigaton of Emotional Engagement in Moral Judgment », *Science*, Vol. 293, 2105-2108.
- GUTTING, G. (1998) « 'Rethinking Intuition' : A Historical and Metaphilosophical Introduction », dans DePaul et Ramsey (dir.) *Rethinking Intuition : The Psychology of Intuition and its Role in Philosophical Inquiry*, Rowman Littlfield Publishers, 3-16.
- HAIDT, J. (2001) « The Emotional Dog and its Rational Tail : A Social Intuitionist Approach to Moral Judgment », *Psychological Review*, Vol 108, 814-834.
- HAIDT, J. et C. Joseph (2004) « Intuitive Ethics: How Innately Prepared Intuitions Generate Culturally Variable Virtues », *Daedalus: On Human Nature*, Vol, 133, 55-66.
- HAIDT, J. et F. Bjorklund (2008) « Social Intuitionists Answer Six Questions about Moral Psychology », dans Sinnott-Armstrong (dir.) *Moral Psychology, Vol. 2.*, Boston : MIT Press, 181-217.

- HARSANYI, J. (1953) « Cardinal Utility in Welfare Economics and in the Theory of Risk-Taking », *Journal of Political Economy*, Vol. 61 (5), 434-435.
- HAUSER, M. D. (2006) *Moral Minds: How Nature Designed our Universal Sense of Right and Wrong*, New York : Ecco Press.
- HAUSER, M. D., F. Cushman, L. Young, R. K-X. Jin et J. Mikhail (2007) « A Dissociation Between Moral Judgments and Justifications », *Mind & Language*, Vol. 22 (1), 1-21.
- HAUSER, M. D., L. Young et F. Cushman (2008) « Reviving Rawls's Linguistic Analogy: Operative Principles and the Causal Structure of Moral Actions », dans Sinnott-Armstrong (dir.) *Moral Psychology*, Vol. 2., Boston : MIT Press, 107-143.
- HOBBS, T. (1651) *Leviathan or The Matter, Forme and Power of a Common Wealth Ecclesiasticall and Civil*, Andrew Crooke et William Crooke.
- HOLMGREN, M. (1989) « The Wide and Narrow of Reflective Equilibrium », *Canadian Journal of Philosophy*, Vol. 19 (1), 43-60.
- HOOKER, B. (1996) « Ross-Style Pluralism Versus Rule-Consequentialism », *Mind*, Vol. 106, 531-532.
- HOOKER, B. (2002) « Intuitions and Moral Theorizing », dans Stratton-Lake (dir.) *Ethical Intuitionism : Re-Evaluations*, New York : Oxford University Press, 161-183.
- HOOKER, B. et M. Little (dir.) (2000) *Moral Particularism*, Oxford : Oxford University Press.
- HUEMER, M. (2005) *Ethical Intuitionism*, Palgrave MacMillan.
- HUME, D. (1739) *A Treatise of Human Nature*, Thoemmes Continuum (2001).
- JACOBSON, D. (2008) « Does Social Intuitionism Flatter Morality or Challenge It ? » dans Sinnott-Armstrong (dir.) *Moral Psychology*, Vol. 2., Boston : MIT Press, 218-224.
- JOSEPH, H W. B. (1931) *Some Problems in Ethics*, Oxford : Clarendon Press.
- KACHELMEIER, S. J. et S. Shehata (1992) « Examining Risk Preferences under High Monetary Incentives : Experimental Evidence from the People's Republic of China », *The American Economic Review*, Vol. 82 (5), 1120-1141.
- KAGAN, S. (1989) *The Limits of Morality*, Oxford : Clarendon Press.
- KAHNEMAN, D. et A. Tversky (1979) « Prospect Theory : An Analysis of Decisions under Risk », *Econometrica*, Vol. 47, 313-327.
- KANT, I. (1785) *Groundwork of the Metaphysics of Morals*, Cambridge : Cambridge University Press (1998).

- KAPPEL, K. (2002) « Challenges to Audi's Ethical Intuitionism », *Ethical Theory and Moral Practice*, Vol. 5 (4), 391-413.
- KORNBLITH, H. (1998) « The Role of Intuition in Philosophical Inquiry : An Account with No Unnatural Ingredients » dans DePaul et Ramsey (dir.) *Rethinking Intuition : The Psychology of Intuition and its Role in Philosophical Inquiry*, Rowman Littlfield Publishers, 129-142.
- KUHN, T. (1964) « A Fucnction for Thought Experiments », reproduit dans Hacking (dir.) *Scientific Revolutions* (1981), Oxford : Oxford University Press, 6-27.
- LAURENCE, S. et E. Margolis (1999) « Concepts and Cognitive Science », dans Margolis et Laurence (dirs.) *Concepts : Core Readings*, Cambridge : MIT Press, 3-81.
- LAURENCE, S. et E. Margolis (2011) « Concepts », *Stanford Encyclopedia of Philosophy*, [En ligne] : <http://plato.stanford.edu/entries/concepts>.
- LEHRER, K. (1990) *Theory of Knowledge*, Londres : Routledge.
- LIAO, S. M. (2008) « A Defense of Intuitions », *Philosophical Studies*, Vol. 140 (2), 247-262.
- LICHTENSTEIN, S. et P. Slovic (1973) « Response-induced Reversals of Preferences in Gambling : An Extended Replication in Las Vegas » *Journal of Experimental Psychology*, Vol. 101, 16-20.
- LIEBERMAN, M. (2003) « Reflexive and Reflective Judgment Processes : A Social Cognitive Neuroscience Approach », dans Forgas et al. (dirs) *Social Judgments : Implicit and Explicit Processes*, New York : Cambridge University Press, 44-67.
- LOCKE, J. (1689) *Two Treatises of Government*, CreateSpace (2010).
- MACH, E. (1897) « Über Gedankenexperimente », *Zeitschrift für den physikalischen und chemischen Unterricht*, Vol. 10, 1-5.
- MACKIE, J. (1977) *Ethics : Inventing Right and Wrong*, Londres : Penguin.
- MCNAUGHTON, D. (2002) « An Unconnected Heap of Duties ? » dans Stratton-Lake (dir.) *Ethical Intuitionism : Re-Evaluations*, New York : Oxford University Press, 76-91.
- MIKHAIL, J. (2008) « The Poverty of the Moral Stimulus », dans Sinnott-Armstrong (dir.) *Moral Psychology*, Vol. 1., Boston : MIT Press, 353-359.
- MIKHAIL, J. (2011) *Elemets of Moral Cognition : Rawls' Linguistic Analogy and the Cognitive Science of Moral and Legal Judgment*, Cambridge : Cambridge University Press.
- MILLER, D. T. et C. McFarland (1986) « Counterfactual Thinking and Victim Compensation : A Test of Norm Theory », *Personality and Social Psychology Bulletin*, Vol. 13, 228-238.

- MOORE, G. E. (1903) *Principia Ethica*, Cambridge : Cambridge University Press (1993).
- NAHIMAS, E., S. Morris, T. Nadelhoffer et J. Turner (2005) « Surveying Freedom : Folk Intuitions about Free Will and Moral Responsibility », *Philosophical Psychology*, Vol. 18 (5), 561-584.
- NARVAEZ, D. (2008) « The Social Intuitionist Model : Some Counter-Intuitions » dans Sinnott-Armstrong (dir.) *Moral Psychology, Vol. 1.*, Boston : MIT Press, 233-240.
- NICHOLS, S. (2004) *Sentimental Rules: On the Natural Foundations of Moral Judgment*, New York: Oxford University Press.
- NIELSEN, K. (1993) « Relativism and Wide Reflective Equilibrium », *The Monist*, Vol. 76 (3), 316-332.
- NORTON, J. D. (1996) « Are Thought Experiments Just What You Thought ? » *Canadian Journal of Philosophy*, Vol. 26 (3), 333-366.
- NORTON, J. D. (2004) « On Thought Experiments : Is There More to the Argument ? », *Proceedings of the 2002 Biennial Meeting of the Philosophy of Science Association*, Vol. 71, 1139-1151.
- PEIRCE, P. (2009) *The Intuitive Way : The Definitive Guide to Increasing Your Awareness*, Hillsboro : Atria Books.
- POPPER, K. (1959) « On the use and misuse of imaginary experiments, especially in Quantum Theory », dans Popper (dir.) *The Logic of Scientific Discovery*, Londres: Hutchinson, 442-456 .
- PRICE, R. (1758) *A Review of the Principal Questions in Morals*, dans Raphael (dir.), *British Moralists, 1650-1800*, Hackett (1990).
- PRICHARD, H. A. (1912) « Does Moral Philosophy Rest on a Mistake ? », *Mind*, Vol. 21 (81), 21-37.
- PRINZ, J. J. (2007) *The Emotional Construction of Morals*, Oxford: Oxford University Press.
- PRINZ, J. J. (2008) « Is Morality Innate? », dans Sinnott-Armstrong (dir.) *Moral Psychology, Vol. 2.*, Boston : MIT Press, 367-406.
- PUFENDORF (1660) *Elementorum jurisprudentiae universalis libri II*, Edwin Mellen (1990).
- PUTNAM, H. (1975) *Mind, Language, and Reality. Philosophical Papers Vol. 2*, Cambridge : Cambridge University Press.
- QUINE, W. V. O. (1951) « Two Dogmas of Empiricism », *Philosophical Review*, Vol. 60, 20-43.
- RAMSEY, W. (1998) « Prototypes and Conceptual Analysis », dans DePaul et Ramsey (dir.) *Rethinking Intuition : The Psychology of Intuition and its Role in Philosophical Inquiry*, Rowman Littlefield Publishers, 161-178.

- RAPHAEL, D. D. (1981) *Moral Philosophy*, Oxford : Oxford University Press.
- RASHDALL, H. (1907) *The Theory of Good and Evil*, Londres : Oxford University Press.
- RAWLS, J. (1951) « Outline of a Decision Procedure for Ethics », *Philosophical Review*, Vol. 60 (2), 177–97.
- RAWLS, J. (1971) *A Theory of Justice*, Cambridge : Harvard University Press.
- RAWLS, J. (1993) *Political Liberalism*, New York : Columbia University Press.
- RAWLS, J. (2001) *Justice as Fairness : A Restatement*, Cambridge : Harvard University Press.
- ROSCH, E. (1978) « Principles of Categorization », dans Rosch et Lloyd (dirs.), *Cognition and Categorization*, Hillsdale : Lawrence Erlbaum Associates, 27–48.
- ROSCH, E. et C. Mervis (1975) « Family Resemblances: Studies in the Internal Structure of Categories », *Cognitive Psychology*, Vol. 7, 573–605.
- ROSS, W. D. (1927) *The Right and the Good*, Oxford: Oxford University Press.
- ROSS, W. D. (1939) *Foundations of Ethics*, Oxford : Oxford University Press.
- SELLARS, W. (1963) *Science, Perception and Reality*, New York : The Humanities Press.
- SHAFIR, E. (1993) « Choosing Versus Rejecting : Why Some Options Are Both Better than Others », *Memory & Cognition*, Vol. 21 (4), 546-556.
- SHAFIR, E. (1998) « Philosophical Intuitions and Cognitive Mechanisms », dans DePaul et Ramsey (dir.) *Rethinking Intuition : The Psychology of Intuition and its Role in Philosophical Inquiry*, Rowman Littlefield Publishers, 59-74.
- SHER, G. (2009) *Who Knew ? Responsibility Without Awareness*, New York : Oxford University Press.
- SIDGWICK, H. (1874) *The Methods of Ethics*, Londres : Macmillan.
- SINGER, P. (1972) « Famine, Affluence, and Morality », *Philosophy and Public Affairs*, Vol. 1, 229–43.
- SINGER, P. (1974) « Sidgwick and Reflective Equilibrium », *The Monist*, Vol. 58 (3), 490-517.
- SINGER, P. (2005) « Ethics and Intuitions », *The Journal of Ethics*, Vol. 9 (¾), 331-352.
- SINNOTT-ARMSTRONG, W. (2002) « Moral Relativity and Intuitionism », *Philosophical Issues*, Vol. 12, 305-328.

- SKELTON (2012) « William David Ross », *The Stanford Encyclopedia of Philosophy* [En ligne] : <http://plato.stanford.edu/entries/william-david-ross>.
- SLOVIC, P. et S. Lichtenstein (1983) « Preference Reversal : A Broader Perspective », *American Economic Review*, Vol. 72, 923-955.
- SMITH, E. et D. Medin (1981) *Categories and Concepts*, Cambridge : Harvard University Press.
- SORENSEN, R. A. (1992) *Thought Experiments*, Oxford : Oxford University Press.
- SOSA, E. (1980) « The Raft and the Pyramid : Coherence Versus Foundation in the Theory of Knowledge », *Midwest Studies in Philosophy*, Vol. 5, 3-25.
- SOSA, E. (1998) « Minimal Intuition », dans DePaul et Ramsey (dir.) *Rethinking Intuition : The Psychology of Intuition and its Role in Philosophical Inquiry*, Rowman Littlfield Publishers, 257-270.
- SOSA, E. (2007) *A Virtue Epistemology : Apt Belief and Reflective Knowledge, Vol. 1*, Oxford : Oxford University Press.
- STICH, S. (1998) « Reflective Equilibrium, Analytical Epistemology and the Problem of Cognitive Diversity », dans DePaul et Ramsey (dir.) *Rethinking Intuition : The Psychology of Intuition and its Role in Philosophical Inquiry*, Rowman Littlfield Publishers, 95-112.
- STICH, S. et R. Nisbett (1980) « Justification and the Psychology of Human Reasoning », *Philosophy of Science*, Vol. 47 (2), 188-202.
- STRATTON-LAKE, P. (2002) « Introduction » dans Stratton-Lake (dir.) *Ethical Intuitionism : Re-Evaluations*, New York : Oxford University Press, 1-28.
- SWAIN, S., J. Alexander et J. Weinberg (2008) « The Instability of Philosophical Intuitions : Running Hot and Cold on Truetemp », *Philosophy and Phenomenal Research*, Vol. 76 (1), 138-155.
- TAPPOLET, C. (2000) *Émotions et Valeurs*, Paris : Presses Universitaires de France.
- THOMSON, J. J. (1971) « A Defense of Abortion », *Philosophy and Public Affairs*, Vol. 1 (1), 47-66.
- TVERSKY, A., P. Slovic, et D. Kahneman (1990) « The Causes of Preference Reversal », *American Economic Review*, Vol. 80, 204-217.
- TVERSKY, A. et R. Thaler (1990) « Preference Reversals », *Journal of Economic Perspectives*, Vol. 4 (2), 193-205.
- VÄYRYNEN, P. (2008) « Some Good and Bad News for Ethical Intuitionism », *The Philosophical Quaterly*, Vol. 58 (232), 489-511.
- WEINBERG, J. M., S. Nichols et S. Stich (2001) « Normativity and Epistemic Intuitions »,

Philosophical Topics, Vol. 29 (1-2), 429-460.

WIGGINS, D. (1987) « A Sensible Subjectivism? », dans *Needs, Values, Truth: Essays in the Philosophy of Value*, Oxford: Blackwell, 185-214.

WILLIAMS, B. (1995) *Making Sense of Humanity*, Cambridge : Cambridge University Press.

WISNIEWSKI, E. J. (1998) « The Psychology of Intuition », dans DePaul et Ramsey (dir.) *Rethinking Intuition : The Psychology of Intuition and its Role in Philosophical Inquiry*, Rowman Littlefield Publishers, 45-58.

WITTGENSTEIN, L. (1953) *Philosophical Investigations*, Wiley-Blackwell (2009).

