

Table des matières

Sommaire

Résumé	iii
Table des matières	v
Remerciements	vii
Introduction	1
1. Le modèle déductif-nomologique de l'explication scientifique tel que défendu par Carl Hempel et le modèle unificationniste soutenu par Michael Friedman et Philip Kitcher	5
1.1 Introduction	5
1.2 Modèle déductif-nomologique	6
1.2.1 Présentation générale du modèle	6
1.2.2 Problèmes et difficultés liés au modèle	8
1.2.3 Conclusion : synthèse critique	11
1.3 Modèle unificationniste	12
1.3.1 Friedman et les fondements du modèle unificationniste	12
1.3.2 Développement du modèle unificationniste par Kitcher	14
1.3.3 Solutions aux difficultés du modèle DN	16
1.3.4 Difficultés du modèle unificationniste	18
1.3.5 Synthèse critique	21
2. Le modèle causal mécanique de l'explication scientifique défendu par Wesley Salmon	23
2.1 Introduction	23
2.2 Présentation générale du modèle causal mécanique	24
2.3 Processus causaux et interactions causales	25
2.4 L'explication scientifique selon le modèle causal mécanique	28
2.5 Difficultés du modèle causal mécanique	32
2.5.1 Le problème de la pertinence causale	32
2.5.2 Le modèle causal mécanique, les systèmes complexes et la physique quantique	34
2.6 Synthèse critique	37
3. L'empirisme constructif de Bas van Fraassen	39
3.1 Introduction	39
3.2 Remarques préalables	40

3.3 Définition et caractérisation de l'empirisme constructif.....	41
3.3.1 Thèse du réalisme scientifique	41
3.3.2 Thèse de l'empirisme constructif et la notion d'adéquation empirique.....	42
3.3.3 Observabilité.....	47
3.4 Objection de Ladyman: observabilité et réalisme modal	50
3.5 Synthèse critique.....	53
4. La théorie pragmatique de l'explication de Bas van Fraassen	55
4.1 Introduction	55
4.2 Considérations générales à propos des théories pragmatiques de l'explication	56
4.3 La théorie pragmatique de l'explication de van Fraassen: une théorie des questions-pourquoi	58
4.3.1 Mise en contexte : l'explication et les théories scientifiques	58
4.3.2 Importance du contexte dans les relations causales et les questions-pourquoi	62
4.3.3 L'explication selon van Fraassen: une théorie des questions-pourquoi.....	65
4.4 La théorie de l'explication de van Fraassen au sein de l'empirisme constructif	68
Conclusion.....	75
Bibliographie.....	81

Remerciements

La rédaction de ce mémoire a été possible grâce à l'aide et au soutien de quelques personnes à qui, par la présente, j'aimerais exprimer ma très sincère reconnaissance.

D'abord, j'aimerais remercier chaleureusement ma directrice de mémoire Madame Renée Bilodeau, qui a été d'une aide sans pareil dans toutes les étapes de ma rédaction. Avant de bâtir mon projet de maîtrise, je ne savais trop sur quel sujet travailler, même si je savais que l'épistémologie m'intéressait tout particulièrement. C'est Madame Bilodeau qui m'a orienté vers un auteur que je n'avais jamais étudié auparavant. Elle a visé on ne peut plus juste : j'ai tout de suite été enchanté par mes lectures et j'ai pu déterminer très rapidement le sujet précis de mon mémoire. Par ailleurs, Madame Bilodeau a toujours lu et révisé mes chapitres avec la plus minutieuse attention pour que je sois en mesure de produire un travail au meilleur de mes capacités. Merci pour tout votre temps.

En outre, je tiens à remercier Messieurs Victor Thibaudeau et Pierre-Olivier Méthot d'avoir accepté d'examiner mon mémoire. Leurs commentaires m'ont permis d'approfondir plusieurs questions abordées dans mon mémoire et je leur en suis très reconnaissant.

Je dois également exprimer ma gratitude envers le Conseil de recherches en sciences humaines du Canada ainsi que la Faculté de philosophie de l'Université Laval pour leur soutien financier durant mes études de deuxième cycle.

J'aimerais aussi remercier chaudement mes parents, Sophie et Raïf, pour leur support inébranlable depuis le début de mes études jusqu'à aujourd'hui. Ils m'ont toujours appuyé de manière inconditionnelle, peu importe les choix liés à mon cheminement scolaire. Sans ce support moral et matériel qu'ils m'offrent, il m'aurait été impossible de produire ce mémoire. Je suis conscient de la chance que j'ai.

Finalement, j'aimerais aussi témoigner de ma gratitude envers ma grand-mère Constance, qui non seulement m'inspire chaque jour par sa sagesse et son amour de la réflexion, mais m'encourage sans relâche à poursuivre ce que j'aime.

Introduction

Si Aristote est souvent considéré comme le fondateur de la science, on peut certainement affirmer du même souffle qu'il est aussi le fondateur de l'épistémologie ou, pour prendre un terme plus au goût du jour, de la philosophie de la science. Jusqu'à récemment dans l'histoire, science et philosophie n'étaient pas réellement distinguées; d'ailleurs, le terme « philosophie de la nature » était bien souvent employé pour faire référence à ce que l'on nomme aujourd'hui « science ». Depuis que ces deux disciplines intellectuelles sont considérées comme séparées l'une de l'autre, et depuis la division de la science en plusieurs branches plus spécifiques (physique, biologie, chimie, etc.), le terme « philosophie des sciences » est plus fréquemment utilisé. Les philosophes des sciences s'intéressent à de nombreuses questions à propos des fondements, des implications logiques et de la définition de plusieurs concepts qui forment le cœur de la pensée scientifique. Par exemple, « Qu'est-ce qu'une loi de la nature? », « Qu'est-ce qui distingue une théorie scientifique d'une théorie non scientifique? » ou « Est-ce que la science nous dit la vérité sur le monde? » sont des questions récurrentes en philosophie des sciences.

La notion d'explication est très importante au sein de la science. En effet, il semble que les scientifiques se servent des théories qu'ils avancent ou qu'ils adoptent dans le but, entre autres, d'expliquer les phénomènes de ce monde. La notion d'explication scientifique est justement l'une des notions les plus étudiées en philosophie des sciences. C'est précisément sur cette notion que portera le présent mémoire, notion qui est à la source de nombreux questionnements d'ordre philosophique. Les discussions contemporaines sur l'explication scientifique, qui ont été lancées par les travaux de Carl G. Hempel sur le modèle déductif-nomologique de l'explication, semblent pour la plupart présupposer, comme le remarque James Woodward, que la tâche principale des philosophes qui s'intéressent à cette notion est de mettre de l'avant un modèle ou une théorie de l'explication scientifique qui puisse rendre compte de la structure ou de la forme de cette explication (Woodward, 2014). Or, cette manière de concevoir les choses implique qu'il existe une forme unique de l'explication qui est celle de l'explication scientifique. D'abord, de cette idée découle qu'il existe un contraste entre les explications qui sont scientifiques et les explications qui sont d'un autre type. En outre, d'affirmer qu'il existe une structure particulière de l'explication scientifique implique que l'explication scientifique doit être distinguée d'autres types de relations, par exemple la description ou la classification.

L'une des importantes figures de la philosophie des sciences contemporaine est sans contredit Bas van Fraassen, philosophe canadien d'origine néerlandaise, surtout connu pour avoir développé la

théorie de l'empirisme constructif, laquelle est une théorie antiréaliste de la science. Dans son ouvrage majeur *The Scientific Image*, van Fraassen défend cette version de l'antiréalisme scientifique en abordant une multitude de questions à propos de la nature des théories scientifiques, des difficultés conceptuelles rattachées au réalisme scientifique, de la méthode scientifique, etc. Par ailleurs, un chapitre entier de cet ouvrage est consacré à la notion d'explication scientifique. Effectivement, van Fraassen intègre, dans sa défense de l'empirisme constructif, une théorie pragmatique de l'explication scientifique, qu'il oppose aux théories de l'explication dites traditionnelles qui ont été le centre d'attention des philosophes des sciences au cours des dernières années. Je propose donc, dans le présent travail, de présenter cette théorie pragmatique de l'explication en l'opposant aux théories traditionnelles de l'explication les plus influentes afin de mieux comprendre comment elle s'insère dans le cadre général défini par l'empirisme constructif.

Dans le premier chapitre, j'exposerai les grandes lignes du modèle de l'explication déductif-nomologique mis de l'avant par Hempel. Comme il s'agit sans doute du modèle de l'explication scientifique le plus influent dans la deuxième moitié du vingtième siècle, il s'agira d'un point de départ adéquat pour entamer une réflexion philosophique sur l'explication scientifique. Nous verrons comment Hempel définit et caractérise cette notion ainsi que certaines difficultés conceptuelles qui se rattachent à son modèle. La présentation de ces difficultés nous permettra, par la suite, d'aborder le modèle unificationniste de l'explication, le plus souvent associé à Michael Friedman et Philip Kitcher. Similairement, nous verrons de quelle manière est définie l'explication scientifique selon ce modèle. De plus, nous tenterons de comprendre pourquoi les défenseurs du modèle unificationniste pensent qu'il permet de résoudre certains des problèmes auxquels faisait face le modèle déductif-nomologique. Nous constaterons cependant que ce modèle unificationniste, lui aussi, n'est pas sans impliquer quelques difficultés conceptuelles.

Le deuxième chapitre sera entièrement consacré au modèle causal mécanique de l'explication tel que développé par Wesley Salmon, figure importante de la philosophie des sciences des dernières décennies. Ardent défenseur du réalisme scientifique, il est aussi connu pour ses travaux portant sur la notion de causalité, qui est d'ailleurs tout à fait centrale dans son modèle causal mécanique de l'explication scientifique. Non seulement ce modèle permettra de mieux comprendre le questionnement général sur l'explication scientifique, en ce qu'il incarne une approche assez différente des deux modèles abordés dans le premier chapitre, mais il est d'autant plus intéressant de l'étudier, dans le cadre de ce mémoire, en ce qu'il permettra de bien cerner le caractère propre de la théorie pragmatique de l'explication de van Fraassen. Ce modèle causal mécanique sera donc

exposé de manière plus détaillée, et encore une fois, nous nous pencherons sur quelques problématiques conceptuelles qui s'y rattachent.

Dans le troisième chapitre, nous nous intéresserons à la théorie de l'empirisme constructif. D'abord, nous verrons comment la science est caractérisée selon cette théorie, par opposition au réalisme scientifique. Cela nous amènera ensuite à nous pencher sur les notions centrales de l'empirisme constructif, notamment sur celles d'adéquation empirique et d'observation, afin de bien cerner le volet épistémique de cette théorie antiréaliste de la science. La dernière portion du chapitre sera consacrée à une objection qui a été dirigée à l'endroit de van Fraassen concernant justement la notion d'observation. La réponse que propose van Fraassen pour neutraliser cette objection va permettre de raffiner notre compréhension de la notion d'observation et, par le fait même, de celle d'adéquation empirique. Une présentation brève de l'empirisme constructif, telle que proposée dans ce chapitre, est importante, car elle permettra de comprendre le cadre général dans lequel s'insère la théorie pragmatique de l'explication que défend van Fraassen.

Le quatrième et dernier chapitre incorporera des éléments qui ont été abordés dans les trois chapitres précédents. En effet, ce chapitre sera dédié à la compréhension de la théorie pragmatique de l'explication mise de l'avant par van Fraassen. Nous verrons en premier lieu en quoi ce type de théorie de l'explication s'oppose aux théories plus traditionnelles telles que présentées dans les deux premiers chapitres, ce qui permettra de mieux en saisir l'essence. Par la suite, nous explorerons les détails plus techniques et formels qui définissent et caractérisent cette théorie de l'explication. Finalement, je consacrerai quelques pages pour montrer en quoi je pense que cette théorie pragmatique de l'explication, nonobstant le fait qu'elle puisse être étudiée en elle-même et adéquatement comprise sans faire référence aux notions centrales de l'empirisme constructif, se fonde particulièrement bien dans le cadre théorique défini par ce dernier, étant donné qu'elle rappelle de manière assez probante la modestie épistémique dont fait preuve l'empirisme constructif.

1. Le modèle déductif-nomologique de l'explication scientifique tel que défendu par Carl Hempel et le modèle unificationniste soutenu par Michael Friedman et Philip Kitcher

1.1 Introduction

Ce chapitre portera sur deux des modèles de l'explication scientifique les plus influents des dernières décennies, soit le modèle déductif-nomologique et le modèle unificationniste. Le modèle déductif-nomologique de l'explication scientifique a non seulement été le modèle le plus communément accepté au milieu du vingtième siècle (années 1950-1960), mais c'est surtout à partir de la critique de ce modèle que beaucoup d'autres théories de l'explication scientifique ont été développées. Ainsi, il est pratiquement indispensable qu'une présentation de ce modèle soit faite dans un travail portant sur l'explication scientifique, puisque la bonne compréhension de celui-ci fournit une assise solide pour l'étude des théories concurrentes de l'explication scientifique. C'est d'ailleurs sur la base d'une critique du modèle déductif-nomologique que le modèle unificationniste s'est développé. Il me semble ainsi nécessaire d'aborder ce modèle dans le présent travail étant donné qu'il est non seulement l'un des premiers modèles étayés à concurrencer le modèle déductif-nomologique, mais aussi parce qu'il est issu du même paradigme épistémique que ce dernier, selon lequel une explication prend la forme d'un argument logique valide. Il sera donc plus aisé de comprendre le modèle qui sera exploré dans le deuxième chapitre, qui est fondé dans un tout autre paradigme, après avoir étudié deux modèles certes différents, mais partageant tout de même une conception semblable de l'explication. Pour chacun des deux modèles, il sera d'abord question de les présenter de manière générale et de montrer en quoi consiste une bonne explication scientifique. Par après, dans le cas du modèle déductif-nomologique, nous nous pencherons sur quelques problèmes qui l'affectent, pour ensuite conclure avec une brève synthèse critique. En ce qui concerne le modèle unificationniste, il sera aussi question de quelques difficultés qui lui sont liées, mais il sera d'abord question de montrer en quoi ses défenseurs prétendent offrir une solution aux problèmes présentés par le modèle déductif-nomologique, pour finalement conclure, encore une fois, avec une synthèse critique.

1.2 Modèle déductif-nomologique

1.2.1 Présentation générale du modèle

Carl G. Hempel (1905-1997), éminent philosophe des sciences du vingtième siècle, est notamment connu pour avoir articulé et défendu un modèle de type déductif-nomologique de l'explication scientifique, conjointement avec, entre autres, Oppenheim et Popper. Avant de plonger dans l'exploration de ce modèle, notons que Hempel, en tant que critique du positivisme logique, a largement contribué au développement du courant plus nuancé qu'est celui de l'empirisme logique et que ses travaux ont eu une influence monumentale qui se fait sentir encore aujourd'hui à travers les développements en philosophie des sciences.

L'une des idées les plus fondamentales du modèle déductif-nomologique est qu'une explication scientifique a la forme d'un argument logique, plus précisément d'une déduction. Ainsi, selon ce modèle, une explication scientifique consiste en un *explanandum* (description linguistique du phénomène qui demande à être expliqué) déduit logiquement à partir d'un *explanans*, qui est constitué par les prémisses de l'argument déductif qui fournissent l'explication (Hempel, 1965, p.247). Il est toutefois pertinent de noter que le terme « *explanandum* » peut parfois désigner non pas l'expression désignant le phénomène qu'on cherche à expliquer, mais ledit phénomène lui-même (Fetzer, 2001, p.70). Évidemment, pour que l'explication scientifique soit acceptable, la déduction doit être valide, et les prémisses formant l'*explanans* doivent être vraies (Hempel, 1965, p.248), ce qui implique la vérité de la conclusion (l'*explanandum*). Ce premier aspect du modèle déductif-nomologique, on l'aura deviné, comble la partie déductive du modèle.

En ce qui concerne la composante nomologique du modèle, elle est soutenue par l'idée qu'au moins une loi de la nature doit faire partie des prémisses de la déduction de manière essentielle, de sorte que la validité de la déduction de l'*explanandum* à partir de l'*explanans* nécessite que la ou les lois soient contenues dans l'*explanans* (Woodward, 2011). Il est important de préciser qu'il est ici question de lois de la nature générales $L1, L2, \dots, Ln$ qui ne laissent place à aucune exception, c'est-à-dire de lois strictement universelles. Elles prennent la forme, en langage symbolique, $(x)(Ax \rightarrow Bx)$.

Les autres prémisses faisant partie de l'*explanans* seront constituées par l'énonciation de circonstances empiriques particulières $C1, C2, \dots, Cn$. Ainsi, ces circonstances empiriques particulières, en conjonction avec certaines lois universelles de la nature, impliqueront l'occurrence de l'*explanandum* E . Une explication scientifique, selon le modèle déductif-nomologique, sera donc un argument déductif de la forme suivante :

$L1, L2, \dots, Ln$

$C1, C2, \dots, Cn$

E

C'est pourquoi le modèle déductif-nomologique fait partie de la catégorie plus générale des « covering-law models », en ce sens que l'*explanandum* est déductivement subsumé (Fetzer, 2001, p.70) sous les lois générales $L1, L2, \dots, Ln$ en vertu des conditions empiriques particulières $C1, C2, \dots, Cn$. En d'autres mots, selon le modèle déductif-nomologique de l'explication scientifique, le phénomène qui demande à être expliqué devient une instantiation spatio-temporelle particulière d'une loi universelle de la nature, d'où l'idée que la loi recouvre le phénomène.

Cela nous mène à un autre aspect de ce modèle, à savoir la manière dont il est lié à la compréhension scientifique des phénomènes. En effet, les rapports entre l'explication scientifique et la compréhension scientifique des phénomènes est un sujet récurrent en philosophie des sciences¹. Ces deux concepts sont intimement liés, tellement que Michael Friedman affirmera que « We can find out what scientific understanding consists in only by finding out what scientific explanation is and vice versa » (Friedman, 1974, p.6). Or, dans un passage souvent cité de son grand ouvrage *Aspects of Scientific Explanation*, Hempel montre clairement de quelle manière il comprend la relation entre ces deux notions :

Thus a D-N Explanation answers the question 'Why did the explanandum-phenomenon occur?' by showing that the phenomenon resulted from certain particular circumstances, specified in $C1, C2, \dots, Ck$, in accordance with the laws $L1, L2, \dots, Lr$. By pointing this out, the argument shows that, given the particular circumstances and the laws in question, the occurrence of the phenomenon *was to be expected*; and it is in this sense that the explanation enables us to *understand why* the phenomenon occurred (Hempel, 1965, p.337)².

Hempel voit donc un lien très intime entre la notion de compréhension d'un phénomène et l'idée que ce phénomène était anticipé. De par la nature même du modèle déductif-nomologique de l'explication, l'*explanandum* est justement anticipé à cause de la conjonction de la loi universelle et

¹ La notion de compréhension scientifique mériterait certainement d'être explorée. À ce sujet, voir, entre autres, de Regt, Leonelli et Eigner (2009).

² Texte souligné par l'auteur. D'ailleurs, dans tout le reste du présent mémoire, les mises en évidence dans les citations seront toujours faites par les auteurs et auteures.

des circonstances particulières qui composent l'*explanans* (il s'agit d'une « nomic expectability » (Salmon, 1989, p.57), ce qui montre en quoi l'explication scientifique aboutit à la compréhension du phénomène en question). Autrement dit, la compréhension d'un phénomène quelconque découle de l'anticipation de ce phénomène à partir d'une loi universelle et générale de la nature appliquée dans un contexte particulier.

1.2.2 Problèmes et difficultés liés au modèle

En premier lieu, le modèle déductif-nomologique de l'explication scientifique, bien qu'il semble assez intuitif à première vue, n'est pas sans poser de difficultés qui ont fait l'objet de nombreuses discussions parmi les philosophes des sciences. L'une d'entre elles est intimement liée à l'une des notions constitutives du modèle, c'est-à-dire celle de loi universelle et générale de la nature. Il ne s'agit pas vraiment ici de remettre en question le statut ontologique des lois de la nature, malgré que cela puisse certainement être pertinent pour l'étude détaillée du modèle déductif-nomologique. Il s'agit plutôt de soulever la question épistémologique, voire sémantique, de la distinction entre une loi générale et universelle de la nature et un simple énoncé général vrai. Une distinction claire entre ces deux types d'énoncés semble nécessaire pour la crédibilité du modèle déductif-nomologique. Selon Hempel, on peut distinguer, parmi les énoncés généraux vrais, les véritables lois et les énoncés qui sont vrais de façon accidentelle (Woodward, 2011). Par exemple, l'énoncé « Tous les joueurs de hockey de la ligue nationale mesurent au moins 1,70m » est un énoncé général accidentellement vrai, alors que l'énoncé « Tout métal, lorsque chauffé, se dilate, toute chose étant égale par ailleurs » est une loi physique de la nature. C'est à l'aide de ce genre d'exemple³ qu'Hempel tente d'appuyer son intuition. Ainsi, le deuxième énoncé, en tant que loi, pourrait être utilisé pour expliquer pourquoi tel ou tel morceau de métal augmente de volume s'il est chauffé, alors que le premier énoncé ne pourrait pas être utilisé pour expliquer pourquoi le joueur de hockey Daniel Brière mesure au moins 1,70m.

Malgré que ce genre d'exemple nous fasse percevoir assez intuitivement une distinction entre une loi et un énoncé accidentellement vrai, on est en droit de revendiquer une distinction conceptuelle plus rigoureuse entre les deux types d'énoncés. Or, Hempel reconnaissait la légitimité de la requête, mais n'a jamais été en mesure de fournir une distinction claire et précise entre une loi de la nature et une généralisation accidentellement vraie, affirmant que la notion de loi est « hautement récalcitrante » (Hempel, 1965, p.338). À ce sujet, une remarque semble s'imposer : il faut noter que le problème de la définition conceptuelle de la notion de loi, encore aujourd'hui, n'a pas reçu de

³ Cf. Hempel, 1965, p.339

solution généralement acceptée. Il serait donc ingrat d'accuser Hempel d'avoir esquivé ce problème et d'ainsi rejeter son modèle déductif-nomologique. Peut-être que le fait pour certains énoncés d'être quasi unanimement reconnus comme étant des lois de la nature est suffisant pour l'utilisation adéquate de la notion de loi impliquée dans ce modèle. Cependant, il ne faudrait tout de même pas minimiser le problème; en l'absence d'une définition rigoureuse de la notion de loi, l'une des idées au cœur du modèle déductif-nomologique reste sans fondement solide, et cela est certainement une raison qui devrait pousser les tenants de ce modèle, bien que peu nombreux de nos jours, à fournir une clarification du concept de loi de la nature.

En deuxième lieu, examinons une autre difficulté soulevée par le modèle déductif-nomologique, cette fois-ci d'ordre plus général : le problème des explications non pertinentes. C'est entre autres Salmon qui a popularisé cette critique, à l'aide d'un exemple souvent utilisé dans la littérature (Salmon, 1971, p.34) :

(L) Aucun homme qui ingère régulièrement des pilules anticonceptionnelles ne tombe enceinte

(C) X est un homme qui ingère régulièrement des pilules anticonceptionnelles

(E) X ne tombe pas enceinte

Dans cette déduction, la première prémisse pourrait probablement être considérée comme étant une loi générale et universelle⁴; du moins, cela est une possibilité envisageable. En outre, l'*explanandum* découle d'une manière nécessaire de la conjonction de (L) et (C) qui constitue l'*explanans*. Il semble donc que cette déduction satisfasse les conditions requises d'une explication valide dans le modèle déductif-nomologique.

Toutefois, il est clair que bien peu de gens considéreraient que (L) et (C) expliquent véritablement l'occurrence de (E). Pourquoi? Parce que l'énoncé (L), bien que vrai, ne semble aucunement pertinent pour expliquer qu'un homme quelconque ne tombe pas enceinte, étant donné que ledit homme, évidemment, ne tomberait pas enceinte même s'il n'ingérait pas régulièrement de pilules anticonceptionnelles. C'est le manque de pertinence de l'*explanans* par rapport à l'*explanandum* qui nous pousse intuitivement à rejeter ce genre d'explication.

⁴ La question de la nature des lois qui constituent en partie l'*explanans* des explications est traitée dans les paragraphes précédents.

Pour Salmon, ce genre de situation dans laquelle un manque de pertinence invalide l'explication met en relief l'idée qu'une explication adéquate devrait traquer la cause de l'*explanandum*, ce qui n'est justement pas le cas dans l'exemple des pilules anticonceptionnelles. C'est l'absence d'un lien causal entre les phénomènes décrits par l'*explanans* et l'*explanandum* qui expliquerait ici le manque de pertinence de l'explication proposée. Nous reviendrons plus tard sur la théorie de l'explication scientifique que propose Salmon, mais retenons pour l'instant que ce genre de difficulté semble indiquer que l'anticipation d'un phénomène à partir d'une loi et de circonstances particulières ne garantit en rien que cette loi exprime un lien causal entre ce qui est décrit dans l'*explanans* et le phénomène qu'on cherche à expliquer.

Cependant, il n'est pas sans intérêt de souligner qu'Hempel et de nombreux autres défenseurs du modèle déductif-nomologique avaient une conception qu'on pourrait qualifier d'empiriste ou humienne de la causalité, selon laquelle la relation de causalité n'était pas première et n'était rien d'autre qu'une relation de régularité nomologique : « Generally, the assertion of a causal connection between individual events seems to me unintelligible unless it is taken to make, at least implicitly, a nomological claim to the effect that there are laws which provide the basis for the causal connection asserted » (Fetzer, 2001, p.91). Ainsi, pour Hempel, il ne serait pas vraiment sensé d'affirmer qu'une explication mettant en valeur une relation nomologique entre deux événements n'est valable que dans la mesure où elle identifie une relation causale véritable, car les relations causales ne sont justement rien d'autre que des relations nomologiques explicitant des régularités entre l'occurrence de la cause et l'occurrence de l'effet. Autrement dit, pour Hempel, les explications causales impliquent l'existence de lois et ne sont donc qu'un sous-ensemble des explications admises par le modèle déductif-nomologique.

Une troisième difficulté rattachée au modèle déductif-nomologique, souvent relevée dans la littérature, est celle des explications asymétriques. Dans plusieurs cas, alors que ce modèle semble adéquat pour déduire un *explanandum* E à partir d'une loi générale L et de circonstances particulières C, et donc pour expliquer l'occurrence de E, il semble contre-intuitif, voire absurde, de procéder à la déduction inverse qui consisterait à déduire C à partir de L et de E. Or, selon les critères du modèle, la déduction inverse est explicative au même titre que l'est la déduction originale. Plusieurs exemples paradigmatiques ont été donnés pour illustrer cette problématique. L'un d'entre eux est celui du baromètre : il semble tout à fait approprié de vouloir expliquer la chute de l'indicateur d'un baromètre (E) en vertu de lois générales liant la pression atmosphérique, l'approche d'une tempête et la position de l'indicateur du baromètre (L) et de l'approche d'une tempête (C). Toutefois, il semblerait quelque peu ridicule de vouloir expliquer l'occurrence de la

tempête à partir de l'indication du baromètre et de l'ensemble de lois (L). Il y a donc ici une asymétrie, en ce sens que l'approche de la tempête semble pertinente pour expliquer la chute de l'indicateur du baromètre, alors que la chute de l'indicateur du baromètre ne semble pas pertinente pour expliquer la présence de la tempête. Or, rien ne permet, dans le modèle déductif-nomologique, de distinguer ces deux déductions, dans la mesure où elles sont toutes deux logiquement valides et basées sur des prémisses vraies, dont l'une est constituée par des lois générales de la nature. Autrement dit, ce modèle reste insensible à l'aspect asymétrique qui semble être inhérent à certains types d'explications (Woodward, 2011).

1.2.3 Conclusion : synthèse critique

À la lumière de la présentation générale du modèle déductif-nomologique et de quelques difficultés qui s'y rattachent, je pense qu'il nous est maintenant possible d'élaborer une synthèse critique de ce modèle. Le modèle de l'explication que propose Hempel est une conception claire, précise et relativement simple de l'explication scientifique. Comme le remarque Hempel lui-même, il est bien important de comprendre que son modèle a pour but de rendre compte des aspects théoriques de l'explication, sans intégrer une compréhension de ses aspects pragmatiques : « This ideal suggests the problem of constructing a nonpragmatic conception of scientific explanation [...]. And this is this nonpragmatic conception of explanation with which the two covering-law models are concerned » (Fetzer, 2001, p.82). Friedman reconnaît que cette caractéristique du modèle déductif-nomologique est en quelque sorte un avantage, dans la mesure où elle permet de comprendre l'explication de manière objective, en ce sens que la validité de l'explication, dans ce modèle, ne dépend pas des contextes particuliers dans lesquels elle est donnée (préférences biaisées des scientifiques, etc.) (Friedman, 1974, p.9). D'ailleurs, selon Hempel, plusieurs des objections qui ont été faites contre son modèle concernent justement l'aspect pragmatique des explications, et c'est pourquoi, selon lui, elles manquent leur cible (Fetzer, 2001, p.83).

Néanmoins, il reste que certaines objections ayant été dirigées contre le modèle déductif-nomologique n'ont jamais pu être démantelées ni par Hempel, ni par les partisans de son modèle. Il s'agit notamment de deux des difficultés explorées ci-haut, c'est-à-dire le problème de la non-pertinence et le problème des explications asymétriques. C'est donc dire que ce modèle, sans nécessairement être totalement erroné (ce que soutiendra toutefois Salmon), est incapable, à lui seul, de répondre aux exigences conceptuelles relatives au problème de la nature de l'explication scientifique. Les problèmes de la non-pertinence et de l'asymétrie font voir que le modèle déductif-nomologique établit, dans le meilleur des cas, des conditions *nécessaires* pour rendre compte adéquatement de l'explication scientifique, mais insuffisantes à elles seules. D'ailleurs, c'est un peu

sur cette dernière idée que se fonde le modèle unificationniste de l'explication scientifique (Woodward, 2011; Weber & al., 2013, p.12), élaboré principalement par Kitcher, dont il sera question dans la section suivante.

1.3 Modèle unificationniste

1.3.1 Friedman et les fondements du modèle unificationniste

C'est Michael Friedman (1974) qui a donné le coup d'envoi au modèle unificationniste de l'explication scientifique. Dans la section précédente, il a été mentionné que Friedman considérait comme intimement liés les concepts d'explication scientifique et de compréhension scientifique. D'ailleurs, la question de déterminer quelle relation peut permettre de considérer qu'un phénomène quelconque peut servir d'explication pour un autre phénomène, et surtout en quoi cette même relation pouvait permettre de comprendre ce phénomène, était tout à fait capitale pour Friedman (1974, p.6). C'est donc en ayant en tête d'éclairer les liens entre l'explication et la compréhension que Friedman en est venu à élaborer l'une des premières versions du modèle unificationniste de l'explication scientifique.

En effet, Friedman n'était pas satisfait par la manière dont les modèles de l'explication scientifique en vogue, notamment le modèle déductif-nomologique d'Hempel, rendaient compte de la compréhension scientifique des phénomènes. Nous avons déjà vu qu'Hempel rattachait la notion de compréhension à celle d'anticipation, en ce sens qu'un phénomène est compris scientifiquement s'il est anticipé à partir d'une loi universelle de la nature mise en conjonction avec des circonstances particulières. Or, cette idée, selon Friedman, est problématique : les notions d'anticipation et de compréhension sont fort distinctes, et il ne semble pas acceptable de réduire cette dernière à la première : « To have grounds for rationally expecting some phenomenon is not the same as to understand it » (*Ibid.*, p.8). Friedman donne des exemples pour appuyer cette objection : le fait d'être en mesure de prédire l'occurrence d'une tornade à l'aide d'un baromètre et de certaines règles ou lois générales de météorologie appliquées dans certaines conditions initiales ne permet en rien de comprendre le phénomène de la tornade. En outre, Friedman remarque que seuls les phénomènes particuliers, c'est-à-dire spatiotemporellement localisés, peuvent être anticipés. Ainsi, il ne peut pas être question d'anticiper les phénomènes généraux ou régularités, ceux-ci n'étant pas spatiotemporellement localisés; pourtant, il semble qu'une théorie adéquate de l'explication scientifique devrait permettre de rendre compte de la compréhension scientifique de ces régularités. C'est pourquoi Friedman en vient à la conclusion que le modèle d'explication déductif-nomologique n'est pas en mesure de rendre compte adéquatement de la compréhension scientifique, bien qu'il ne soit

pas exclu que l'anticipation d'un phénomène puisse faire partie de la compréhension de ce dernier (Ibid., p.9).

Ainsi, après avoir critiqué quelques modèles de l'explication scientifique, Friedman en vient à identifier trois caractéristiques qui, selon lui, devraient être respectées par une théorie adéquate de l'explication scientifique (Ibid., pp.13-14). Premièrement, une telle théorie devrait être suffisamment générale de sorte qu'elle devrait être compatible avec les présentes théories scientifiques reconnues généralement comme étant explicatives; Friedman tient donc à ce qu'une théorie de l'explication scientifique puisse s'accorder en bonne partie avec la science telle que pratiquée par les scientifiques. Deuxièmement, un modèle de l'explication scientifique devrait être objectif, c'est-à-dire qu'il devrait permettre de comprendre et d'identifier une explication scientifique adéquate sans tenir compte des préférences des scientifiques ou encore des circonstances contingentes d'ordre historique. Troisièmement – et c'est sur ceci qu'insiste Friedman – une théorie de l'explication scientifique devrait être en mesure de lier clairement les notions d'explication et de compréhension en permettant de mettre en évidence de quelle manière une explication scientifique permet de comprendre les phénomènes expliqués.

Quelle est la conclusion de Friedman? C'est que l'essence d'une théorie de l'explication qui respecte ces trois exigences réside dans le concept d'unification, en ce sens qu'une bonne explication scientifique diminue le nombre de phénomènes indépendants :

I claim that this [unification] is the crucial property of scientific theories we are looking for; this is the essence of scientific explanation – science increases our understanding of the world by reducing the total number of independent phenomena that we have to accept as ultimate or given. A world with fewer independent phenomena is, other things equal, more comprehensible than one with more (Ibid., p.15).

Friedman précise qu'il est ici question d'une notion globale, par opposition à locale, de compréhension scientifique. En effet, Friedman ne s'intéresse pas vraiment à la compréhension d'un phénomène en particulier ou individuel; il s'intéresse plutôt à l'idée d'une compréhension scientifique générale du monde, en ce sens que les bonnes explications scientifiques simplifient le portrait qu'on se fait de la nature en réduisant le nombre de phénomènes indépendants les uns des autres. Or, c'est justement ce qu'il reproche aux autres modèles de l'explication d'avoir ignoré (Ibid., p.18). Cela est manifeste dans le modèle déductif-nomologique : les phénomènes particuliers sont compris dans le sens qu'ils sont anticipés, mais le modèle ne fournit aucune base pour rendre compte de la compréhension scientifique du monde en général. Ainsi, en identifiant l'unification comme étant la caractéristique essentielle de l'explication scientifique, Friedman a en quelque sorte

jeté les fondements du modèle unificationniste, ultérieurement développé considérablement par Philip Kitcher.

1.3.2 Développement du modèle unificationniste par Kitcher

Kitcher, bien qu'il partage avec Friedman l'idée générale que l'explication scientifique est une affaire d'unification, a critiqué la manière dont ce dernier rendait compte théoriquement de l'unification. C'est d'ailleurs Kitcher qui a développé le plus abondamment le modèle unificationniste de l'explication scientifique. Il s'agira donc, dans cette sous-section, de présenter les points essentiels des travaux de Kitcher. La première chose qu'il importe de noter est que tout comme ce que proposait Hempel avec le modèle déductif-nomologique, une explication, pour Kitcher, est conçue comme un argument logique duquel découle une conclusion qui décrit le phénomène à expliquer : « However, the systematization approach retains the Hempelian idea that to explain a phenomenon is to produce an argument whose conclusion describes the phenomenon [...] » (Kitcher, 1989, p.431; 1981, p.509). Ainsi, pour bien comprendre la manière dont Kitcher propose de rendre compte du concept d'unification, il est très important de nous pencher en premier lieu sur les termes techniques qui sont au cœur de sa notion d'argument⁵.

Premièrement, une *phrase schématique* est une expression dans laquelle certains des termes non logiques constituant une phrase sont remplacés par des variables. Pour reprendre l'exemple de Kitcher, à partir de la phrase « Les organismes homozygotes pour l'allèle de l'anémie falciforme développent l'anémie falciforme », il est possible de construire plusieurs phrases schématiques telles que « Les organismes homozygotes pour A développent B » ou « Pour tout X, si X est A et B, alors X est C ». Les *instructions de remplacement* fournissent les directions à suivre pour remplacer les variables d'une phrase schématique. Une séquence de phrases schématiques constituera un *argument schématique*, alors qu'une *classification* est un ensemble d'expressions qui décrivent les caractéristiques logiques ou inférentielles d'un argument schématique, c'est-à-dire qu'une classification permet d'exposer quels termes de l'argument doivent être compris comme des prémisses, quelles sont les règles d'inférence utilisées, etc. Finalement, un *patron argumentatif général* est le tout formé par un argument schématique, un ensemble de groupes d'instructions de remplacement (un groupe pour chacune des phrases schématiques) et une classification.

Pour bien saisir le modèle unificationniste que propose Kitcher, il faut aussi examiner la manière dont il pose le problème de l'explication scientifique. D'abord, Kitcher ne considère pas que les

⁵ Kitcher définit les termes techniques dont il sera question dans le prochain paragraphe dans Kitcher (1981, p.516 et 1989, p.432).

sciences offrent une multitude d'arguments isolés les uns des autres, qui ne peuvent ainsi servir que pour des explications individuelles. Il pense plutôt que la science est en quelque sorte un réservoir d'arguments explicatifs (*explanatory store*). Ainsi, pour Kitcher, une théorie de l'explication scientifique devrait avoir pour but de rendre compte des conditions qui devraient être remplies par un bon réservoir explicatif (1981, p.512; 1989, p.430). Plus formellement, considérons K, c'est-à-dire l'ensemble des énoncés acceptés par la communauté scientifique à un certain moment de l'histoire. La tâche que se donne Kitcher « is that of specifying E(K), the explanatory store over K, which is the set of arguments acceptable as the basis for acts of explanation by those whose beliefs are exactly the members of K » (1981, p.512).

On s'en doute bien, la réponse de Kitcher à cette question sera de dire que E(K) est l'ensemble d'arguments qui réalise la meilleure unification de K (1981, p.519; 1989, p.431). C'est ici que devient fondamentale la notion de patron argumentatif général. En effet, c'est en dérivant des descriptions et des explications de différents phénomènes *à partir du minimum de patrons argumentatifs généraux* qu'il est possible d'unifier la connaissance scientifique et d'ainsi parvenir à une meilleure compréhension du monde : « Science advances our understanding of nature by showing us how to derive descriptions of many phenomena, using the same patterns of derivation again and again, and, in demonstrating this, it teaches us how to reduce the number of types of facts we have to accept as ultimate (or brute) » (1989, p.432). Il est intéressant de noter qu'on retrouve l'intuition centrale de Friedman, à savoir que la compréhension du monde est améliorée en diminuant le plus possible le nombre de phénomènes ou de faits que l'on doit accepter comme ultimes ou irréductibles.

C'est donc parce qu'un patron argumentatif général peut être utilisé pour expliquer et décrire une multitude de phénomènes différents qu'il est possible de parler d'unification de K. Cependant, Kitcher amène une précision cruciale : ce n'est que par le biais de patrons argumentatifs *restrictifs* qu'une réelle unification est possible. Plus un patron argumentatif est restrictif et plus il permet d'être appliqué pour décrire une grande variété de phénomènes, alors plus les explications de ces phénomènes sont unifiées. Différents arguments peuvent être soit similaires quant à leur structure logique, soit quant au vocabulaire non logique qu'ils contiennent (1981, p.518; 1989, p.433). Or, la notion de patron argumentatif général met justement en évidence que différents arguments peuvent partager certaines caractéristiques (logiques ou non logiques). La restrictivité d'un patron argumentatif est justement déterminée par ces caractéristiques : plus les conditions logiques ou non logiques d'instanciation d'un argument qui respecte le patron argumentatif sont difficiles à satisfaire, plus le patron argumentatif est restrictif (1989, p.433). Ainsi, cela revient à dire que c'est

la classification d'un patron argumentatif (qui identifie sa structure logique) et les instructions de remplacement qui lui sont rattachées (qui spécifient le vocabulaire non logique) qui vont déterminer son niveau de restrictivité (1981, p.518)⁶. Par exemple, un patron argumentatif qui n'est que très peu restrictif, en ce sens qu'une multitude d'arguments différents pourraient facilement le respecter, n'est pas ce qui permet une bonne unification, justement parce qu'il y a unification lorsque des arguments assez semblables peuvent être utilisés pour expliquer des phénomènes différents. Cette citation résume bien le modèle que propose Kitcher : « The intuitive idea behind unification is the generation of as many conclusions as possible using as few patterns. It is also important that the instantiations of the patterns should genuinely be similar, that is, that the patterns in question should be stringent » (1989, p.434).

1.3.3 Solutions aux difficultés du modèle DN

Le modèle unificationniste de l'explication que propose Kitcher, en plus d'être intuitivement assez intéressant, est en mesure de contrer deux difficultés majeures qui affligeaient le modèle déductif-nomologique. En premier lieu, considérons le problème des explications sans pertinence, illustré par l'exemple de la pilule anticonceptionnelle et de l'homme qui ne tombe pas enceinte (voir section 1.1.2). Un autre exemple est assez commun dans la littérature : celui du sel ensorcelé. Imaginons un magicien qui ensorcelle un échantillon de sel de table en passant sa main au-dessus et en exprimant une formule magique. Il est évidemment vrai que tous les échantillons de sel ayant été ensorcelés se dissoudront dans l'eau. Il serait donc possible d'établir une déduction à partir de la « loi » générale selon laquelle tout échantillon de sel ensorcelé se dissout si placé dans de l'eau et des circonstances particulières entourant cette dissolution pour expliquer pourquoi cet échantillon de sel ensorcelé s'est effectivement dissout. Tout comme le cas de l'homme qui ne tombe pas enceinte et de la pilule anticonceptionnelle, le modèle déductif-nomologique ne réussit pas, de manière convaincante, à rendre compte du caractère non explicatif de ces arguments.

Or, justement, le modèle unificationniste de Kitcher permet d'exclure ces arguments du réservoir d'arguments explicatifs. Pour ce faire, il faudra montrer que l'ensemble d'arguments contenant l'argument que nous désirons exclure ne réalise pas la meilleure unification possible de nos croyances (l'ensemble K). Si nous considérons l'argument qui déduit du fait d'avoir ensorcelé le sel le fait qu'il se dissout dans de l'eau, il faudra utiliser un autre argument pour expliquer la dissolution du sel non ensorcelé. De même, dans le cas de l'homme et de la pilule

⁶ Kitcher n'élabore pas de théorie complète de la restrictivité des patrons argumentatifs. Il suffit pour son argumentation d'expliquer cette notion de manière qualitative.

anticonceptionnelle, un autre argument sera nécessaire pour expliquer qu'un homme qui n'utilise pas la pilule anticonceptionnelle ne tombe pas enceinte. Ainsi, contrairement à la personne qui utilise un seul argument (d'ordre chimique) pour expliquer *tous* les cas de dissolution de sel dans de l'eau ou encore un seul argument (d'ordre biologique) pour expliquer qu'*aucun* homme ne puisse tomber enceinte, la personne qui utilise l'argument fondé sur l'ensorcellement du sel pour expliquer sa dissolution ou encore sur l'usage de la pilule anticonceptionnelle pour expliquer que l'homme ne tombe pas enceinte utilisera ainsi deux patrons d'arguments différents. Dans ces deux cas d'explication non pertinence, l'argument explicatif est construit sur la base d'une propriété accidentelle du phénomène à expliquer. Au contraire, l'argument qui explique toute dissolution de sel et le fait qu'aucun homme ne tombe enceinte peut être appliqué généralement : il explique une plus grande diversité de phénomènes, ce pour quoi il permet une meilleure unification. Cela permet donc de rejeter les arguments de l'ensorcellement et de la pilule anticonceptionnelle hors du réservoir d'arguments explicatifs.

Kitcher prétend que son modèle unificationniste peut aussi offrir une solution au problème des asymétries qui affectait le modèle déductif-nomologique. Pour bien comprendre cette solution, considérons l'exemple paradigmatique du mât de drapeau, lequel illustre le fait qu'il est intuitivement correct d'expliquer la grandeur de l'ombre du mât à partir de la longueur de ce dernier, alors qu'il ne semble pas adéquat de vouloir expliquer la longueur réelle dudit mât à partir de la grandeur de son ombre. Or, pour les raisons évoquées plus haut (avec l'exemple du baromètre et de la tempête), le modèle que propose Hempel ne permet pas de distinguer clairement ces deux explications et de rejeter la seconde. La solution de Kitcher sera la suivante : l'ensemble des procédés explicatifs ou des arguments *O*, c'est-à-dire la systématisation des connaissances, selon lequel la longueur réelle du mât est expliquée à partir de la grandeur de son ombre réalise une moins bonne unification que celui selon lequel la longueur réelle du mât est expliquée à partir d'un seul patron d'explication, qu'il nomme le patron de l'origine et du développement (*origin and development [pattern of explanation]*)⁷ (Kitcher, 1989, p.485).

En effet, l'ajout de *O* au patron de l'origine et du développement, alors qu'il augmente la quantité de patrons explicatifs dans le réservoir explicatif, ne permet de parvenir à aucune nouvelle conclusion, car la dimension de n'importe quel objet peut être déterminée par le patron de l'origine et du développement. Toutefois, si l'on éliminait le patron de l'origine et du développement de notre réservoir explicatif pour le remplacer uniquement par *O*, la taille du réservoir explicatif serait

⁷ Selon le patron de l'origine et du développement, la dimension de n'importe quel objet peut être déterminée à partir des conditions sous lesquelles il a été créé et des modifications qui l'ont affecté.

similaire, mais elle ne permettrait pas de dériver autant de conclusions, car de nombreux objets ne possèdent pas d'ombre (par exemple, s'ils sont à la noirceur), ou alors leurs ombres ne sont pas assez nettes pour en dériver leur longueur exacte (Woodward, 2011). Bref, O ne fait pas partie du réservoir explicatif, ce pourquoi la prétendue explication de la longueur réelle du mât à partir de la longueur de son ombre ne peut être considérée comme valable. Ce même type de solution peut être donné au problème du baromètre : le patron argumentatif selon lequel la chute de l'indicateur du baromètre explique la présence de la tempête ne fait pas partie du réservoir explicatif, car il faudrait ajouter un autre patron pour expliquer la présence d'une tempête dans une situation dans laquelle nous n'avons pas accès à un baromètre analogique, alors qu'une explication fondée sur des données climatologiques précises (pression atmosphérique, déplacement des masses d'air, courants marins, etc.) permet d'expliquer la présence d'une tempête dans tous les cas. Ainsi, comme le patron argumentatif en question ne fait pas partie du réservoir explicatif, il ne peut être utilisé pour expliquer la présence de la tempête, ce qui résout cette instanciation du problème de l'asymétrie.

1.3.4 Difficultés du modèle unificationniste

La première difficulté, et sans doute la plus sérieuse pour le modèle unificationniste de Kitcher, provient d'une conséquence logique liée à la notion de patron argumentatif. Il s'agit de dire, brièvement, que le modèle de Kitcher pourrait admettre que n'importe quel fait F puisse être expliqué par une dérivation effectuée à partir de F lui-même; ainsi, l'auto-explication serait un type d'explication universellement valide dans le modèle unificationniste (Gijsbers, p.13). Pourquoi? Considérons l'argument déductif suivant, valide de façon triviale :

p
 _____,
 p

selon lequel p est une conséquence logique de p . Suivant ce patron argumentatif que nous pourrions qualifier d'auto-déductif, toute connaissance scientifique pourrait être expliquée (par elle-même), et ce, à partir de ce seul patron; il semblerait donc que ce patron soit on ne peut plus unificateur. Toutefois, il est assez évident que ce patron a une valeur explicative la plus minime qui soit s'il en est, étant donné que pour reprendre le vocabulaire qu'utilise Kitcher, il n'est absolument pas restrictif; il ne contient aucun vocabulaire logique ou non logique pour limiter son utilisation.

Mais la difficulté provient du fait qu'il semble possible de rendre ce même patron argumentatif suffisamment restrictif pour qu'il puisse être acceptable dans le modèle unificationniste. Considérons un patron argumentatif A que nous présupposons accepté par la science contemporaine de par ses caractères unificateur et restrictif. De ce patron A découle un ensemble de conclusions $C(A)$. La difficulté est précisément la suivante : il est possible de constituer un patron argumentatif A' tel qu'il soit aussi unificateur et restrictif que A , qui prend la forme de l'argument auto-déductif ci-haut. En effet, pour ce faire, il suffit de remplacer p par n'importe quel énoncé faisant partie de $C(A)$. Autrement dit, il faut compléter le patron de l'argument auto-déductif avec les instructions de remplacement I suivantes : « remplacer p par n'importe quelle conclusion dérivable à partir d'un patron argumentatif acceptable ». Ainsi, de cette manière, A' mènerait exactement aux mêmes conclusions que A , c'est-à-dire que $C(A)$ serait identique à $C(A')$. C'est pourquoi A' serait non seulement aussi unificateur que A , mais aussi restrictif, car toutes et seulement les conclusions acceptables selon A peuvent en découler. Ainsi, ces deux patrons argumentatifs étant aussi unificateurs l'un que l'autre, il semble qu'ils soient aussi explicatifs, bien que cela soit manifestement étrange; le patron auto-déductif A' ne semble pas posséder de grande valeur explicative.

Il faut toutefois noter que Kitcher est bien conscient de ce problème, qu'il nomme celui de l'unification fallacieuse (*spurious unification*) (1981, pp.526-529). D'ailleurs, il tente d'y apporter une solution en ajoutant une exigence qui devrait être respectée par un patron argumentatif valable : « If the filling instructions associated with a pattern P could be replaced by different filling instructions, allowing for the substitution of a class of expressions of the same syntactic category, to yield a pattern P' and if P' would allow the derivation of *any* sentence, then the unification achieved by P is spurious » (1981, pp.527-528). L'idée de Kitcher est donc de dire que bien que le patron A' , complété par I , soit aussi restrictif que A , l'unification qu'il procure n'est qu'illusoire, car il serait possible, en lui joignant des instructions de remplacement autres que I , de conclure n'importe quoi étant donné sa nature auto-déductive. Cependant, on est en droit de se questionner sur le caractère *ad hoc* de cette exigence que propose Kitcher; elle ne semble jouer aucun rôle théorique important dans sa théorie unificationniste, mis à part celui de contrer cette objection de l'unification fallacieuse. Par ailleurs, Gijbers remet fortement en question la légitimité de cette exigence, car « [it] is both too strong and not strong enough: it banishes some patterns that we need to keep, but does not bar all forms of spurious unification » (p.14). L'argument de Gijbers pour soutenir cette idée étant assez complexe, il ne sera pas présenté dans le présent travail; cela

demanderait un trop grand espace. Il suffit de noter que la solution de Kitcher au sérieux problème de l'unification fallacieuse est bien loin de faire l'unanimité.

Une deuxième objection que l'on pourrait adresser à Kitcher se fonde dans l'intuition que certaines explications sont meilleures que d'autres, c'est-à-dire qu'il existe un continuum entre diverses explications, certaines étant plus ou moins bonnes, d'autres étant excellentes. Certaines explications présumées peuvent être superficielles ou médiocres, mais tout de même minimalement explicatives. À première vue, le modèle unificationniste paraît s'accorder très bien avec une telle intuition, dans la mesure où l'unification semble être une affaire de degrés, tout comme la notion d'explication; certaines systématisations de nos connaissances sont moyennement unifiées, d'autres le sont énormément, etc. Toutefois, si l'on s'attarde quelque peu aux solutions que Kitcher a proposées aux problèmes des explications non pertinentes et des asymétries, il semble qu'il soit au contraire contraint d'adopter une position qui pourrait être dite du « tout ou rien », c'est-à-dire une position qui n'admet aucune gradation entre les différentes possibles explications d'un phénomène quelconque.

Pour bien comprendre ce point, reprenons l'exemple du sel ensorcelé qui se dissout dans l'eau. Nous avons vu que Kitcher rejetait l'argument de l'ensorcellement hors du réservoir explicatif, car il ne permettait pas une *aussi bonne* unification que celui basé sur les propriétés chimiques du sel et de l'eau, ce dernier permettant d'expliquer la dissolution de n'importe quel échantillon de sel dans l'eau. Mais cela ne revient pas à dire que l'argument du sel ensorcelé, conjointement à l'argument basé sur les propriétés chimiques, ne permet *aucune* unification du phénomène observé, soit la dissolution du sel. Mais Kitcher le rejette totalement, c'est-à-dire qu'il le qualifie de *non explicatif*, sous prétexte qu'il n'offre pas une aussi bonne unification que l'explication fondée uniquement sur les propriétés chimiques. Autrement dit, Kitcher ne conclut pas que l'explication par l'ensorcellement est *moins* explicative que celle par les propriétés chimiques, mais qu'elle n'est *pas* explicative, et ce, même s'il est ici question de dire que l'explication par l'ensorcellement offre une *moins bonne* unification du phénomène de la dissolution.

Il en est de même si l'on se penche sur la solution de Kitcher au problème des asymétries. En effet, nous avons vu qu'il excluait l'argument qui permet d'expliquer la longueur réelle du mât à partir de la longueur de son ombre du réservoir explicatif parce qu'il n'offrait pas une aussi bonne unification que le patron général d'argument de l'origine et du développement. Mais il serait bien difficile de soutenir que l'argument de l'ombre, pris conjointement à celui de l'origine et du développement, n'offre *aucune* unification de la mesure de la dimension des objets; pourtant,

Kitcher semble dire de cet argument qu'il est absolument non explicatif. Autrement dit, pour Kitcher, il semble que la seule explication valable pour un phénomène quelconque soit celle fondée sur le patron argumentatif qui offre la meilleure unification, les autres étant automatiquement qualifiées de non explicatives. Par ailleurs, cela n'est pas qu'un détail, car les solutions de Kitcher aux problèmes des explications dépourvues de pertinence et des asymétries reposent justement sur cette idée du tout ou rien, dans la mesure où il réussit à nier le caractère explicatif des arguments de la pilule contraceptive et de l'ombre du mât en soutenant qu'ils n'offrent pas la *meilleure* unification possible des phénomènes à expliquer. Autrement, Kitcher serait forcé de conclure que l'explication fondée sur l'usage de la pilule contraceptive est véritablement explicative (mais dans une moindre mesure que celle fondée sur la physiologie masculine), conclusion qu'il veut manifestement éviter. Bref, il semble que le modèle unificationniste soit incapable de rendre compte de l'idée qu'il puisse y avoir des explications plus ou moins bonnes et d'autres explications présumées qui ne sont, finalement, absolument pas explicatives; seule la théorie la plus unificatrice est explicative, comme le remarque Michael Strevens (2009, p.4).

1.3.5 Synthèse critique

Kitcher, avec son modèle unificationniste, prétend être en mesure de rendre compte de l'explication scientifique en évitant les problèmes qui s'imposaient au modèle déductif-nomologique, notamment celui des explications non pertinentes et celui des asymétries. Il me semble toutefois pertinent de relever une similarité théorique qui me paraît importante entre ces deux modèles : tous deux veulent rendre compte de l'explication sans être fondés sur la notion de causalité. Non seulement Hempel et Kitcher ne pensent pas que la causalité doit être au fondement de la théorie de l'explication, mais encore il semble qu'ils considèrent ces deux notions comme quasiment indépendantes. D'ailleurs, Kitcher écrit explicitement que le « parce que » propre à la causalité est toujours une certaine dérivation du « parce que » propre aux explications (1989, p.477); dans cette optique, il n'y a donc pas de causalité ontologiquement indépendante de nos explications. Cette manière de voir les choses repose certainement sur l'idée humienne ou empiriste qu'il n'est pas possible de cerner clairement, de manière théorique, ce qu'est une véritable relation causale (Woodward, 2011). Ainsi, d'une certaine façon, de ne pas être aux prises avec les difficultés théoriques inhérentes au concept de causalité représente un avantage non négligeable pour Kitcher.

Cependant, d'évacuer ce concept hors d'une théorie de l'explication comme l'unificationnisme de Kitcher ne va pas sans poser de difficulté. En effet, en plus des problèmes pour le modèle unificationniste explorés dans la section précédente, d'autres difficultés s'imposent à ce modèle, et ce, peut-être précisément parce qu'il n'inclut pas la notion de causalité. Par exemple, considérons

un système physique newtonien⁸ selon lequel la position de planètes quelconques dans un temps futur $T1$ est calculée à partir de la position de ces mêmes planètes au temps présent $T0$. Il s'agirait ici d'une dérivation prédictive, en ce sens que des informations sur le système dans un temps futur sont inférées à partir d'informations sur le système dans un temps présent. Mais, théoriquement, une dérivation rétrodictive est tout autant possible : dans un tel système newtonien, il est possible de déduire la position présente des planètes au temps $T0$ à partir des données sur la position de ces planètes au temps $T1$. Selon le modèle unificationniste, ces deux dérivations sont aussi explicatives l'une que l'autre, car elles font appel exactement aux mêmes patrons explicatifs newtoniens. Toutefois, intuitivement, seule la dérivation prédictive semble explicative; la dérivation rétrodictive, bien que théoriquement acceptable, ne semble pas posséder de valeur explicative. Pourquoi? Car seule la dérivation prédictive semble expliquer l'effet (la position au temps $T1$) à partir de la cause (la position au temps $T0$)⁹, ce qui paraît nécessaire pour qu'une dérivation soit considérée comme réellement explicative. Le modèle unificationniste semble donc incapable de rendre compte de telles asymétries fondées sur les relations entre la cause et l'effet. Bref, on est en droit de douter de la thèse de Kitcher selon laquelle les jugements de causalité sont toujours dérivés à partir d'explications.

Pour conclure ce chapitre, même si les modèles déductif-nomologique, développé notamment par Hempel, et unificationniste, développé principalement par Kitcher, offrent des manières théoriquement assez convaincantes de comprendre l'explication scientifique, il n'en reste pas moins qu'ils soulèvent des difficultés quelque peu gênantes, qui découlent peut-être du fait qu'ils sont tous deux indépendants de la notion de causalité. C'est pourquoi, dans le chapitre suivant, nous nous pencherons sur le modèle de l'explication scientifique développé par Salmon, qui, contrairement aux deux autres modèles tout juste abordés, est justement centré autour des relations causales.

⁸ Dans lequel la temporalité est symétrique, c'est-à-dire qu'il est possible d'inférer, théoriquement, des conclusions à propos du système dans un temps futur à partir de données d'un temps passé ou présent, et vice-versa.

⁹ Alors qu'inversement, la dérivation rétrodictive semble expliquer la cause à partir de l'effet, ce qui est tout à fait contre-intuitif.

2. Le modèle causal mécanique de l'explication scientifique défendu par Wesley Salmon

2.1 Introduction

Wesley Salmon (1925-2001) est certainement l'un des philosophes des sciences les plus influents du vingtième siècle. Il a orienté plusieurs de ses travaux sur la critique et la clarification des modèles de l'explication scientifique, ce pour quoi il me semble nécessaire d'accorder quelques pages à cet auteur dans le présent travail, surtout que quelques-uns de ses textes ont précisément porté sur la conception pragmatique de l'explication développée par van Fraassen; d'ailleurs, leurs désaccords ont donné vie à une série d'échanges fort intéressants. Ainsi, non seulement les travaux de Salmon sont tout à fait pertinents en eux-mêmes, mais ils permettront de mettre en évidence les caractéristiques du modèle pragmatique de van Fraassen qui fera l'objet du dernier chapitre de ce travail.

Salmon est particulièrement connu pour avoir accordé une importance capitale à la notion de causalité dans son traitement de l'explication scientifique. D'ailleurs, Salmon a écrit abondamment sur le concept de causalité; cela n'a rien d'étonnant si l'on considère les difficultés qui lui sont rattachées, mises en évidence entre autres par David Hume. Ainsi, Salmon n'avait guère le choix de tenter de démystifier cette notion pour la mettre au cœur de son traitement de l'explication scientifique. Il sera donc nécessaire de s'arrêter quelque temps aux travaux de Salmon sur la causalité pour être en mesure de cerner clairement sa manière de comprendre l'explication scientifique. Il est à la source de deux modèles de l'explication scientifique auxquels la communauté philosophique a accordé une grande importance : le modèle de pertinence statistique (*statistical-relevance model*) et le modèle causal mécanique (*causal mechanical model*). Ce chapitre sera consacré à ce dernier, étant donné qu'il s'agit du dernier modèle de l'explication scientifique proposé par Salmon, modèle qui se veut plus achevé et convaincant que le modèle de pertinence statistique. Il ne sera toutefois pas inutile d'aborder à quelques reprises le modèle de pertinence statistique étant donné que ces deux modèles sont similaires sur plusieurs points. D'abord, une présentation très globale du modèle causal mécanique sera faite, pour ensuite nous pencher quelque temps sur des notions plus techniques qui sont essentielles pour bien le comprendre. Nous verrons par la suite comment l'explication scientifique, pour Salmon, s'articule autour de ces notions. Par après, nous porterons notre attention sur certaines difficultés qui peuvent se poser pour le modèle causal mécanique, pour finalement conclure, encore une fois, avec une synthèse critique.

2.2 Présentation générale du modèle causal mécanique

Dans plusieurs de ses travaux¹⁰, Salmon soutient une distinction entre trois conceptions générales de l'explication scientifique : la conception épistémique, la conception modale et la conception ontique. La conception épistémique est parfaitement exemplifiée par le modèle déductif-nomologique de Hempel étudié dans le premier chapitre; il s'agit d'une conception selon laquelle une explication scientifique est un argument qui montre en quoi le phénomène qu'on cherche à expliquer pouvait être anticipé. La conception modale, quant à elle, implique qu'une explication scientifique doit montrer en quoi le phénomène à expliquer devait se produire, c'est-à-dire pourquoi ce qui s'est produit n'aurait pas pu être différent. Finalement, en ce qui concerne la conception ontique, il s'agit de montrer comment le phénomène à expliquer s'insère dans les lois ou les régularités de la nature – habituellement comprises comme étant de nature causale. Il s'agit donc d'une conception ontique en ce sens qu'il n'est pas question de considérer l'explication comme un argument ou une inférence; il s'agit plutôt de dévoiler la constitution du monde naturel lui-même. C'est une telle conception ontique de l'explication scientifique que Salmon cherche à défendre avec son modèle causal mécanique. Bref, Salmon voit un lien extrêmement étroit et fondamental entre la notion d'explication et celle de causalité (1984, p.296) : une explication scientifique doit décrire la toile (*nexus*) ou le réseau causal qui produit et qui entoure l'*explanandum* (Weber & al., 2013, p.12).

Qu'est-ce à dire? Une toile causale est un réseau physique constitué par ce que Salmon nomme des processus causaux et des interactions causales (ces deux notions, capitales dans les travaux de Salmon, seront présentées dans la sous-section suivante). Ainsi, le rôle de l'explication est en quelque sorte de situer le phénomène à expliquer dans ce nœud causal, de fournir une « carte causale », pour employer l'expression de Christopher R. Hitchcock (1995, p.308). Plus précisément, Salmon distingue deux aspects de l'explication : l'aspect étiologique et l'aspect constitutif. L'aspect étiologique met en lumière les interactions et les processus causaux qui précèdent l'*explanandum* afin d'expliquer pourquoi il s'est produit : « [it] involves the placing of the explanandum in a causal network consisting of relevant causal interactions that occurred previously and suitable causal processes that connect them to the fact-to-be-explained » (Salmon, 1984, p.269). Autrement dit, l'aspect étiologique de l'explication concerne la position de l'*explanandum* dans la toile causale qui l'entoure. L'aspect constitutif, quant à lui, concerne la nature du phénomène lui-même : « If we want to show why E manifests certain characteristics, we place inside the volume occupied by E the internal causal mechanisms that account for E's nature. This is the constitutive aspect of our

¹⁰ Notamment, Salmon (1984; 1998, pp.53-54)

explanation; it lays bare the causal structure of E » (*Ibid.*, p.275). Le volume en question représente l'intérieur du phénomène E tel qu'illustré dans la *Figure 1*¹¹, si l'on présuppose que E occupe un volume fini dans l'espace-temps. Bref, l'aspect constitutif de l'explication se rapporte aux mécanismes causaux internes à l'*explanandum*.

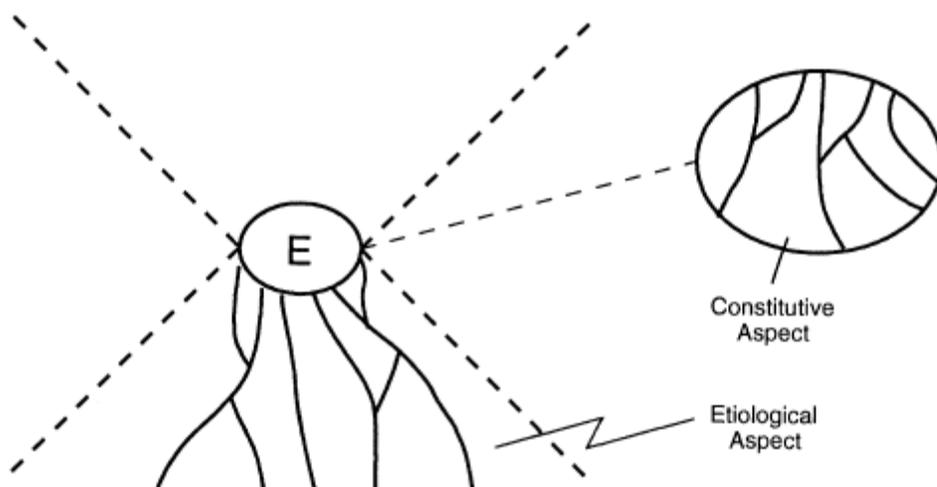


Figure 1. Aspects étiologique et constitutif d'une explication selon Salmon

2.3 Processus causaux et interactions causales

La causalité est une notion extrêmement problématique depuis les réflexions de Hume au début du dix-huitième siècle. Hume notait qu'il était incapable de mettre le doigt sur une connexion nécessaire entre la cause et l'effet, ou encore sur un pouvoir mystérieux quelconque permettant à la cause de créer l'effet. Il est d'ailleurs fort probable que le refus d'auteurs comme Hempel ou Kitcher d'inclure la notion de causalité dans leurs théories de l'explication soit dû à ce genre de difficultés inhérentes à cette notion. Salmon, sans évidemment avoir la prétention de neutraliser les critiques humiennes, pense toutefois être en mesure de traiter de la notion de causalité d'une nouvelle manière que Hume n'aurait peut-être pas envisagée (Salmon, 1998, p.59). En effet, au lieu de centrer son attention sur les causes et les effets en tant qu'événements, Salmon propose d'accorder une importance prépondérante à la notion de processus : « I have argued that causal processes are precisely the kinds of causal connections Hume sought but was unable to find » (*Ibid.*, p.71). Un processus, pour Salmon, est une trajectoire ou une collection de points dans l'espace-temps qui fait preuve d'une certaine continuité dans ses propriétés (Hitchcock, 1995, p.306); il peut

¹¹ Tirée de Hitchcock, 1995, p.309



s'agir d'un avion volant dans le ciel ou encore de l'ombre de ce dernier qui se déplace sur le sol. Il faut aussi distinguer un processus d'un événement, un événement étant habituellement défini comme étant l'instanciation d'une propriété dans un point ou un intervalle précis de l'espace-temps¹², alors que l'idée directrice derrière la notion de processus est la *continuité* dans l'espace-temps. Salmon définit d'ailleurs le processus d'une autre manière en disant qu'un processus est toute chose qui fait preuve de constance ou de persistance dans sa structure à travers le temps (1984, p.144). L'idée de persistance et de continuité est donc ce qui semble distinguer le processus de l'événement, et c'est pourquoi les objets matériels sont inclus parmi les processus pour Salmon, même si cela peut sembler contre-intuitif. Un exemple pourrait sans doute permettre de mieux comprendre cette importante distinction: alors qu'un pare-brise de voiture et un caillou sont deux processus en tant que tous deux persistent dans le temps, l'éclatement du pare-brise à l'instant de son contact avec le cailloux est un événement (car cet éclatement a une occurrence spatiotemporelle ponctuelle). Si Salmon accorde une importance capitale à la notion de processus dans son ontologie, c'est justement parce qu'il croit que c'est cet aspect de constance à travers le temps (au lieu de parler d'événements spatiotemporellement séparés) qui peut permettre de rendre compte adéquatement de la causalité¹³. C'est notamment ce qu'illustre la notion de transmission de marque dont il sera question dans le paragraphe suivant.

Par ailleurs, Salmon distingue deux types fondamentaux de processus : les pseudo-processus et les processus causaux. C'est la capacité à transmettre une marque, ou encore à transmettre de l'information, qui permet de différencier ces deux types de processus (Salmon, 1984, p.297). Qu'est-ce que cela veut dire, que de transmettre une marque? C'est comme suit que répond Salmon :

Let P be a process that, in the absence of interactions with other processes, would remain uniform with respect to characteristic Q, which it would manifest consistently over an interval that includes both of the space-time points A and B ($A \neq B$). Then a mark (consisting of a modification of Q into Q'), which has been introduced into process P by means of a single local interaction at point A, is transmitted to point B if P manifests the modification Q' at B and at all stages of the process between A and B without additional interventions (Salmon, 1984, p.148).

¹² La notion d'événement est abondamment discutée et il n'existe pas de définition universellement acceptée de ce terme. Il suffit de comprendre qu'un événement est lié à un temps t et à un lieu l précis.

¹³ Une analyse détaillée de la théorie de la causalité que propose Salmon dépasserait largement le cadre de ce mémoire, car elle est très riche et assez complexe. Plus précisément, pour une meilleure compréhension des rapports entre les notions de processus et d'événement en lien avec la causalité, cf. Salmon (1984, pp. 178-179) et Chakravarty (2005).

On s'en doute bien, ce ne sont que les processus causaux qui peuvent transmettre des marques, alors que les pseudo-processus n'en ont pas la capacité. L'un des exemples les plus connus que donne Salmon (1984, pp.143-144) pour illustrer cette distinction est celle de la voiture qui roule et de son ombre qui la suit. Si, pour une raison quelconque, la voiture entre en collision avec une roche, la marque causée par cette collision (l'égratignure ou l'éclat dans le pare-brise, par exemple) va suivre la voiture tout au long de son trajet, de manière spatiotemporellement continue. Si, au contraire, l'ombre de la voiture percute un gros rocher, elle sera momentanément déformée, mais dès l'instant où elle aura passé le rocher, elle reprendra sa forme initiale. Pour que la modification de la forme de l'ombre persiste, il faudrait continuellement intervenir tout au long de son parcours, car la modification n'est pas transmise à la structure de l'ombre elle-même, contrairement à un cas de réel processus causal comme l'est la voiture en mouvement (Woodward, 2011).

En outre, il est important de noter que cette notion de transmission de marque est une notion contrefactuelle, et ce, sous deux aspects (Hitchcock, 1995, p.307). En effet, c'est d'abord la notion elle-même de transmission de marque qui est contrefactuelle : pour savoir si un processus quelconque transmet une marque ou non, il faut savoir comment ce processus se comporterait s'il n'avait pas été marqué. En outre, un processus causal est un processus qui a la possibilité ou la capacité de transmettre une marque, qu'il transmette une marque ou non dans les faits. Pour Salmon, ce caractère contrefactuel n'est en rien problématique, car il est possible de le tester empiriquement (*Ibid.*); par exemple, on sait qu'une balle de baseball, de manière générale, est un processus causal, car si une balle particulière était fortement frappée par un bâton de baseball, elle se déformerait et garderait toujours sa difformité.

Une autre notion capitale dans le modèle de la causalité que propose Salmon est celle d'interaction causale. En quelques mots, une interaction causale se produit lorsque deux processus causaux se rencontrent dans l'espace-temps et que ces deux processus causaux se retrouvent structurellement modifiés par cette rencontre, c'est-à-dire qu'ils possèdent des caractéristiques qu'ils ne possèderaient pas n'eût été cette rencontre : « When two or more causal processes intersect in spacetime, they may or may not produce lasting modifications in one another. If they do, the intersection constitutes a causal interaction » (Salmon, 1984, p.297). Ainsi, de manière plus formelle, soit $P1$ et $P2$ deux processus causaux, soit Q une caractéristique de $P1$, soit R une caractéristique de $P2$ et soit S un point d'intersection dans l'espace-temps entre $P1$ et $P2$: « Then, the intersection of $P1$ and $P2$ at S constitutes a causal interaction if: 1) $P1$ exhibits the characteristic Q before S , but it exhibits a modified characteristic Q' throughout an interval immediately following S ; and 2) $P2$ exhibits the characteristic R before S , but it exhibits a modified characteristic R'

throughout an interval immediately following S » (Salmon, 1984, p.174). L'exemple paradigmatique d'une interaction causale est celui d'une collision entre deux voitures: les deux voitures sont des processus causaux qui se rencontrent à un point particulier de l'espace-temps, chacune des deux se trouvant modifiée par la collision. Cependant, toute intersection entre deux processus causaux ne résulte pas nécessairement en une interaction causale. Si deux rayons lumineux (qui sont deux processus causaux) se croisent, il y aura distorsion au point d'intersection, mais chaque rayon lumineux reprendra ses caractéristiques initiales après le point d'intersection (Hitchcock, 1995, p.308). Il n'en reste pas moins qu'un rayon lumineux est bel et bien un processus causal, car il est possible pour lui de participer à des interactions causales (un rayon lumineux puissant peut, par exemple, faire fondre un objet matériel quelconque).

Une brève clarification s'impose ici. En effet, ce qui vient d'être mentionné dans le paragraphe précédent pourrait faire croire qu'il existe, selon Salmon, deux critères différents d'identification des processus causaux, soit la capacité de transmettre des marques (telle qu'illustrée par l'exemple de la voiture) et la capacité de participer à des interactions causales (telle que mentionnée en ce qui concerne les rayons lumineux). Or, ces deux critères, en réalité, sont absolument identiques: une interaction entre deux processus est une interaction causale, nous l'avons vu, justement quand les deux processus se retrouvent structurellement modifiés après cette rencontre, c'est-à-dire lorsqu'ils se transmettent mutuellement une marque définitive. Autrement dit, un processus quelconque est un processus causal s'il peut participer à une interaction causale, car une interaction causale n'est pas autre chose qu'une interaction entre deux ou plusieurs processus qui se sont transmis des marques.

2.4 L'explication scientifique selon le modèle causal mécanique

Il a été mentionné plus haut que Salmon abandonnait l'idée, défendue entre autres par Hempel, qu'une explication scientifique était un argument. D'ailleurs, Salmon avait déjà abandonné cette idée avec son modèle de pertinence statistique élaboré dans les années 1970, avec lequel il tentait de cerner la pertinence explicative uniquement à l'aide de la notion de pertinence statistique (Woodward, 2011). En effet, l'intuition fondamentale du modèle de pertinence statistique est que seules les propriétés pertinentes du point de vue statistique sont pertinentes du point de vue explicatif. Un attribut quelconque C est statistiquement pertinent par rapport à un autre attribut B , étant donné une classe quelconque A , si et seulement si la probabilité de B conditionnellement à A et à C est différente de la probabilité de B conditionnellement à A uniquement, c'est-à-dire si $P(B|A.C) \neq P(B|A)$. L'exemple de l'homme qui ne tombe pas enceinte et de la pilule anticonceptionnelle permet de bien comprendre l'idée centrale de ce modèle. Considérons A comme la classe de tous les

hommes, B comme étant le fait de tomber enceinte et C comme étant le fait d'ingérer des pilules anticonceptionnelles. Dans une telle situation, il est clair que $P(B|A.C) = P(B|A)$, c'est-à-dire que le fait pour un homme de consommer des pilules anticonceptionnelles n'a absolument aucune pertinence statistique concernant le fait qu'il ne tombe pas enceinte. C'est pourquoi le modèle de pertinence statistique permet, dans un tel cas, de rejeter la pertinence explicative de la consommation de pilules anticonceptionnelles vis-à-vis le fait que tel ou tel homme ne tombe pas enceinte, ce qui était pourtant un problème non négligeable pour le modèle déductif-nomologique.

Ce court passage sur le modèle de pertinence statistique n'est pas sans intérêt pour bien comprendre le modèle causal mécanique, développé environ une décennie plus tard par Salmon. Effectivement, Salmon considère que l'explication scientifique est une affaire de deux étapes, comme le note Woodward (1989, p.358). La première consiste en une collecte d'informations concernant les relations statistiques qui formaient le cœur du modèle de pertinence statistique. Il s'agit ensuite, pour fournir une explication complète, de rendre compte de manière causale, c'est-à-dire en utilisant les notions d'interaction et de processus causaux, de ces relations de pertinence statistique. Ainsi, le modèle causal mécanique que propose Salmon ne rompt pas avec le modèle de pertinence statistique qu'il proposait quelques années plutôt; il s'agit plutôt de le compléter et de lui joindre une théorie de la causalité. Autrement dit, Salmon ne considère pas que les relations statistiques soient par elles-mêmes explicatives, en ce sens que c'est seulement une fois comprises en termes causaux qu'elles le deviennent (*Ibid.*).

C'est notamment grâce à la notion de cause commune, empruntée entre autres à Hans Reichenbach (1956, §19) et à Bertrand Russell (1948, chap.9), que Salmon va traduire en langage causal les relations purement statistiques telles que trouvées dans son modèle de pertinence statistique. Plus précisément, Salmon utilisera l'expression « fourche conjonctive » (*conjunctive fork*) pour analyser les corrélations entre des effets spatiotemporellement séparés, provenant de processus causaux différents, mais découlant par contre d'une cause commune. Le principe de la cause commune est clairement résumé par Salmon : « When apparently unconnected events occur in conjunction more frequently than would be expected if they were independent, then assume that there is a common cause » (Salmon, 1998, p.110). Qu'est-ce que deux événements¹⁴ indépendants du point de vue statistique? Soit deux événements, A et B . Si la probabilité de l'occurrence commune de ces

¹⁴ Salmon utilise ici la notion d'événement plutôt que celle de processus parce qu'il est question de statistiques. Or, seuls des événements, qui sont spatio-temporellement distincts l'un de l'autre, peuvent être pertinents pour évaluer des relations statistiques, car celles-ci portent sur des occurrences spatiotemporelles précises.

événements, $P(A.B)$, est égale au produit de la probabilité de l'occurrence individuelle de chacun de ces deux événements, c'est-à-dire si $P(A.B) = P(A) \times P(B)$, alors A et B sont deux événements statistiquement indépendants. Au contraire, si la probabilité de l'occurrence jointe de deux événements est supérieure ou inférieure à celle du produit de leur occurrence individuelle, alors ces deux événements sont statistiquement significatifs l'un envers l'autre (*Ibid.*).

Évidemment, deux événements statistiquement indépendants n'ont aucune valeur explicative l'un pour l'autre, notamment parce qu'ils sont causalement non pertinents l'un pour l'autre (*Ibid.*). Au contraire, si deux événements ne sont pas statistiquement indépendants, alors il est possible que l'occurrence de l'un de ces deux événements soit directement pertinente pour expliquer l'occurrence du second. Si je remarque, par exemple, une forte corrélation entre la sonnerie de mon téléphone résidentiel et la perte de la connexion Internet sur mon ordinateur (autrement dit, si je remarque que ces deux événements ne sont pas statistiquement indépendants), je conclurai sans doute que l'activation du téléphone fait partie de l'explication de la perte de la connexion Internet, peut-être à cause d'une interférence entre les différentes lignes de communication numériques. Toutefois, deux événements statistiquement pertinents l'un pour l'autre ne font pas nécessairement partie d'une relation explicative de cette sorte. Salmon donne l'exemple dont il a été question au chapitre précédent, à savoir celui de la tempête et du baromètre (*Ibid.*, p.111). En effet, la chute de l'aiguille du baromètre ne peut pas elle-même expliquer la tempête à venir, ni l'inverse. Ainsi, même si ces deux événements sont en forte corrélation statistique, aucun des deux ne peut directement faire partie de l'explication de l'autre événement¹⁵, car aucun des deux n'est la cause de l'autre. C'est ici que la notion de cause commune devient importante.

En effet, Salmon tient *mordicus* à l'un des principes à la base de la théorie de l'explication de Reichenbach, selon lequel toute relation de pertinence statistique doit être expliquée par une relation de pertinence causale (*Ibid.*). Autrement dit, dès qu'il y a une corrélation statistique entre des événements, il doit y avoir un certain lien causal entre eux, sans que ce ne soit nécessairement une connexion directe de cause à effet. Effectivement, ces deux événements peuvent avoir une cause commune, c'est-à-dire qu'il peut y avoir un autre événement C , plus ou moins éloigné dans la chaîne causale menant à chacun des deux événements A et B , qui est commun à chacune de ces deux chaînes causales, de sorte que C est à la fois une cause de A et une cause de B : « If A and B

¹⁵ Pourtant, au chapitre précédent, il a été dit que l'approche de la tempête semblait pertinente pour expliquer la chute de l'aiguille du baromètre. Mais cela n'est juste que dans la mesure où l'approche de la tempête, dans ce cas, n'est pas distinguée des conditions météorologiques particulières (pression atmosphérique, mouvement des masses d'air chaud, etc.) qui sont justement la cause commune de la tempête et de la chute de l'aiguille du baromètre.

are two events that are statistically relevant to each other, but neither is causally relevant to the other, then there must be a common cause C in the region of overlap of the past light cones of A and B » (*Ibid.*, p.121). Nous avons vu plus tôt que deux événements A et B sont statistiquement indépendants si $P(A.B) = P(A) \times P(B)$. Cette définition a pour corolaire que A est statistiquement pertinent pour B si $P(B|A)$, c'est-à-dire la probabilité de B étant donné A , n'est pas identique à $P(B)$. Si la relation de pertinence statistique est positive entre les événements A et B , on obtiendra $P(B|A) > P(B)$ et $P(A|B) > P(A)$. On peut donc conclure, dans un tel cas, que $P(A.B) > P(A) \times P(B)$.

L'inclusion de la cause commune C dans l'équation statistique aura pour conséquence d'annuler la relation de dépendance statistique entre A et B . Dit autrement, A et B deviennent statistiquement indépendants étant donnée la présence de C dans l'équation : $P(A.B|C) = P(A|C) \times P(B|C)$. Cela veut dire que la relation de dépendance statistique entre A et B est masquée par les relations de dépendance causale entre A et C et B et C , exprimées par $P(B|C) > P(B)$ et $P(A|C) > P(A)$ (Sober, 1984, p.217). En résumé, la relation de dépendance statistique entre A et B peut être expliquée par les relations causales qui lient A et B à une cause commune C , d'où l'idée que les relations de pertinence statistique peuvent toujours être analysées en termes causaux : « The claim that statistical-relevance relations can always be explained in terms of causal-relevance relations therefore means that causal-relevance relations screen off other kinds of statistical-relevance relations » (Salmon, 1998, p.121). C'est justement ce qui est illustré par l'exemple de la tempête et du baromètre. La chute de l'aiguille du baromètre ainsi que la présence de la tempête sont certainement liées de manière statistique, mais nous avons vu qu'aucun de ces deux événements n'est la cause de l'autre. Il faut donc chercher une cause commune à ces deux événements, à savoir un certain type de conditions météorologiques, qui permet d'expliquer la dépendance statistique de ces événements d'un point de vue causal. En résumé, une fois les corrélations statistiques entre les différents phénomènes relevées, il reste à les expliquer soit en montrant qu'ils sont la cause l'un de l'autre, ou encore en montrant qu'ils découlent d'une cause commune.

Ainsi, l'idée de Salmon, nous l'avons vu plus haut, est somme toute assez simple : expliquer un phénomène, c'est montrer en quoi ce phénomène s'insère dans un réseau causal formé par des processus et des interactions causales : « It seems to me that the nature of scientific explanation can be summed up as follows : to give scientific explanations is to show how events and statistical regularities fit into the causal network of the world » (Salmon, 1998, p.104). Considérons l'exemple¹⁶ d'une boule de billard blanche, elle-même frappée par un bâton de billard, qui frappe

¹⁶ Tiré de Woodward, 2011.

une autre boule de billard au repos; alors que cette dernière se met en mouvement, la boule blanche, après l'impact, change de direction. D'abord, il s'agit de noter que ces trois entités (le bâton et les deux boules) sont des processus causaux; par exemple, on pourrait considérer que le bâton, dont le bout est taché de craie bleue, transfère cette marque de craie à la boule blanche lors du premier impact, alors que cette dernière la transmet à l'autre boule lors du second impact. C'est donc en citant les faits entourant ces différentes interactions causales (les collisions) qu'il est possible d'expliquer, pour Salmon, les différents mouvements des boules de billard à la suite des collisions qu'elles subissent.

Jusqu'à présent, il a surtout été question de l'aspect causal du modèle de l'explication développé par Salmon. Qu'en est-il de l'aspect mécanique? Je crois qu'il provient essentiellement de la notion d'interaction causale dont il a été question dans la section précédente. En effet, Salmon insiste sur l'aspect local des interactions causales – c'est-à-dire sur l'idée qu'elles se produisent toujours entre des processus causaux spatiotemporellement adjacents. Dit plus simplement, les interactions causales sont toujours, pour Salmon, des rencontres mécaniques de processus causaux : les interactions causales « à distance » sont tout simplement impossibles. Comme Woodward le remarque, cela va d'ailleurs dans le sens de la théorie de la relativité restreinte, qui a pour conséquence que les transferts d'énergie doivent toujours être conservés localement, de manière spatiotemporellement continue (Woodward, 1989, p.359). Bref, Salmon affirme que « Spatiotemporal continuity [...] plays a vital role in theoretical explanation » (Salmon, 1998, p.115). Le terme « mécanique » est donc en quelque sorte, pour Salmon, indissociable de la notion de causalité, en ce sens que les processus causaux qui constituent le monde ne peuvent entrer en interaction que s'ils sont en contact direct.

2.5 Difficultés du modèle causal mécanique

2.5.1 Le problème de la pertinence causale

Nous avons vu, dans le premier chapitre, que le modèle de l'explication déductif-nomologique était aux prises avec le problème des explications non pertinentes. Salmon est d'ailleurs l'un de ceux ayant développé cette critique du modèle Hempélien, sur la base, nous l'avons vu, d'exemples tels que celui de l'homme qui ne tombe pas enceinte et de la pilule anticonceptionnelle. Or, aussi ironique que cela puisse sembler, il s'avère que certains auteurs, notamment Hitchcock, perçoivent dans le modèle causal mécanique défendu par Salmon une difficulté peut-être pas identique, mais à tout le moins semblable, à savoir le problème de la non-pertinence causale. Dans les pages précédentes, il a été question de montrer en quoi les notions de processus causal et d'interaction

causale permettent de situer un phénomène quelconque dans une toile causale afin de l'expliquer. Reprenons l'exemple des boules de billard : celles-ci (ainsi que le bâton) sont des processus causaux qui entrent dans des interactions causales, et l'explication scientifique unanimement acceptée (en simple, bien sûr) pour rendre compte du mouvement des différentes boules concerne entre autres leur énergie cinétique et la loi de la conservation d'énergie.

Toutefois, un problème se présente : bien qu'il ne soit pas remis en question que les boules soient bel et bien des processus causaux (et non des pseudo-processus), il est difficile de voir en quoi le modèle de Salmon permet de sélectionner l'énergie cinétique linéaire des boules et non, par exemple, leur couleur ou la tache de craie bleue qu'elles arborent, pour expliquer leur mouvement. Ainsi, il semble que la notion de processus causal n'est pas à elle seule suffisante pour identifier les *propriétés* des processus qui ont une pertinence du point de vue causal vis-à-vis du phénomène à expliquer. Bref, la notion de processus causal (ainsi que celle de transmission de marque) semble insuffisante pour déterminer quelles sont les propriétés pertinentes du point de vue causal pour l'explication d'un phénomène quelconque et les distinguer des propriétés non pertinentes (comme la tache de craie bleue sur la boule) (Woodward, 2011). Comme l'explique clairement Hitchcock, « For example, in stating that a process is a causal process rather than a pseudoprocess, it is asserted only that some later properties of the process are dependent upon some earlier properties; no information is given about what these properties are. Our demand that explanations provide relevant information requires something stronger - that we be told which earlier properties the properties specified in the explanandum depend upon » (Hitchcock, 1995, p.311). La notion de processus causal ne permet pas d'effectuer une telle distinction, car n'importe quelle marque transmise (la quantité d'énergie cinétique linéaire ou encore la tache de craie bleue) est aussi bonne qu'une autre pour identifier qu'un processus quelconque est causal; c'est en regard de l'explication que ces deux marques sont très inégales, l'une étant pertinente, l'autre non : « From the perspective of explanation, however, these two marks are not on a par: the linear momentum of the cue ball was relevant to the eight ball's landing in the corner pocket; the blue spot on the cue ball was irrelevant to this outcome » (*Ibid.*, p.312).

En ce qui concerne l'exemple de l'homme et de la pilule anticonceptionnelle, le modèle de Salmon n'est pas sans poser de difficultés. En effet, lorsqu'un homme particulier ingère une pilule anticonceptionnelle, une multitude de processus causaux ont lieu : la pilule se dissout dans son estomac, ses particules chimiques actives traversent son intestin pour entrer dans le flux sanguin, elles sont ensuite métabolisées par diverses cellules du corps, etc. Toutefois, évidemment, aucun de ces processus causaux n'est pertinent pour expliquer l'incapacité de l'homme à tomber enceinte,

alors qu'ils le sont si c'était une femme qui consommait la même pilule. Cela semble peut-être banal, mais il reste que le modèle de Salmon ne permet pas clairement de distinguer ces processus causaux du point de vue de la pertinence causale si l'on compare le cas de l'homme et celui de la femme. Bref, Hitchcock résume bien cette première difficulté qui affecte le modèle causal mécanique de Salmon : « The concepts of causal process and interaction are physically legitimate and philosophically interesting regardless of whether either of Salmon's theories provides a definitive characterization of them. But these concepts do not provide the resources to explicate the notion of explanatory relevance that Hempel's models failed to capture » (*Ibid.*, p.316).

2.5.2 Le modèle causal mécanique, les systèmes complexes et la physique quantique

Alors que la difficulté explorée dans la sous-section précédente en est une inhérente à la manière dont Salmon articule diverses notions constitutives de son modèle, celles qui seront abordées dans les prochaines lignes concernent davantage l'application ou l'utilisation du modèle causal mécanique par rapport à des situations particulières. Alors que ce modèle n'est pas trop problématique lorsqu'il est question de systèmes physiques relativement simples comme c'est le cas des collisions de boules de billard, cela est autrement lorsque l'on considère, par exemple, des systèmes beaucoup plus complexes étudiés par des sciences comme la physique quantique, la biologie ou la psychologie, qui ne font pas nécessairement appel explicitement à des processus causaux liés de manière spatiotemporelle (Woodward, 2011). Un exemple qui me semble particulièrement éclairant est celui d'un gaz quelconque, renfermé dans un contenant de volume x , auquel on permet de se dilater dans un contenant ayant un volume supérieur à x ¹⁷. Il est possible d'expliquer le comportement du gaz en faisant appel à quelques lois scientifiques générales, comme celle des gaz parfaits.

Cependant, il est important de noter que ces lois considèrent le gaz comme un tout même s'il est évidemment composé par une immense quantité de molécules individuelles; elles font appel, par exemple, à des considérations générales sur la distribution des molécules de gaz dans l'espace, sur la fréquence des collisions entre ces molécules, etc. Or, Salmon considère que seules les molécules individuelles sont des processus causaux, et non le gaz pris comme un tout (Woodward, 2011). Ainsi, la seule manière, selon Salmon, d'offrir une véritable explication du comportement du gaz lors de sa dilatation serait d'analyser les trajectoires et les interactions de chacune des molécules composant le gaz. Cela est absolument impossible, voire farfelu; non seulement il n'est pas humainement possible de réaliser une telle tâche, mais même si ça l'était, l'explication qui en

¹⁷ Cet exemple, ainsi que les considérations s'y rattachant, est tiré de Woodward (1989, p.363).

découlerait ne serait probablement pas celle à préférer. En effet, peu importe la position et la vitesse initiales des molécules de gaz, le comportement macroscopique final du gaz après la dilatation serait le même; autrement dit, les trajectoires particulières des molécules ne sont pas pertinentes pour expliquer le comportement général de la masse gazeuse qu'elles forment.

Si, toutefois, l'explication du comportement du gaz passait par l'analyse de ces trajectoires particulières, il serait impossible de comprendre la multitude de conditions initiales et de trajectoires des molécules différentes qui auraient pu mener au même résultat, ce qui est pourtant scientifiquement très pertinent. Il semble donc qu'en ne permettant qu'une explication à partir des processus causaux individuels, le modèle causal mécanique ne puisse pas permettre, du moins à première vue, d'expliquer simplement et adéquatement les phénomènes qui forment des systèmes complexes dans lesquels c'est l'ensemble plutôt que les multiples parties individuelles qui devraient mériter le plus clair de notre attention, comme le conclut Woodward : « In such cases it is often hopeless to try to understand the behavior of the whole system by tracing each individual process. Instead one needs to find a way of representing what the system does on the whole or on the average, which abstracts from such specific causal detail. It is not clear how Salmon's model which is formulated in terms of individual causal processes and their binary interactions, applies to such cases » (Woodward, 1989, pp.362-363).

En outre, le modèle de Salmon ne va pas de soi lorsqu'il est confronté aux explications proposées par la physique quantique¹⁸. Nous n'avons qu'à penser aux corrélations de type EPR¹⁹, qui, pour le dire simplement, impliquent en quelque sorte un transfert d'information physique instantané et donc plus rapide que la vitesse de la lumière. Par exemple, si deux particules sont intriquées et corrélées, c'est-à-dire si elles possèdent deux états superposés (par exemple, si elles ont deux spins inverses superposés avant d'être mesurés) mais nécessairement opposés l'un à l'autre lorsque la mesure physique est prise, alors peu importe la distance entre les deux particules, dès que le spin d'une des deux particules est mesuré, il est possible de connaître immédiatement le spin de l'autre, et ce, avec une corrélation totale, comme si les deux particules se communiquaient de l'information. Selon Salmon, il faudrait évidemment une cause commune, agissant de manière spatiotemporellement

¹⁸ Il n'est pas ici question d'explorer les enjeux épistémologiques soulevés par la physique quantique, ni même d'offrir une explication quelque peu détaillée des phénomènes quantiques. Il s'agit d'un sujet extrêmement complexe qui ne fait l'unanimité ni chez les épistémologues, ni chez les scientifiques; il existe une multitude de manières différentes de comprendre et d'interpréter ces phénomènes. Il s'agit ici uniquement de montrer que le modèle de Salmon s'accorde difficilement avec le genre de phénomènes étudiés en physique quantique.

¹⁹ Ces corrélations découlent de l'expérience de pensée produite par Einstein, Podolsky et Rosen pour montrer les difficultés liées à l'interprétation de Copenhague de la physique quantique.

continue, pour expliquer cette corrélation parfaite. Or, les corrélations de type EPR représentent un genre de saut dans l'espace-temps, car il est impossible d'en appeler à une cause commune mécanique qui expliquerait la corrélation entre les deux particules une fois la mesure empirique prise. Pour Salmon, le fait que de tels phénomènes ne puissent pas cadrer dans son modèle causal mécanique (contrairement, par exemple, au modèle déductif-nomologique, qui est en mesure de rendre compte, avec des lois statistiques, de ce type de phénomènes) représente non pas une faiblesse du modèle, mais plutôt un avantage : « Salmon seems to suggest that it is an explanatory deficiency in QM that it fails to explain such correlations in terms of spatio-temporally continuous processes connected to a common cause and that it is an advantage of his account of explanation that it allows us to recognize this, while the standard DN model would not » (*Ibid.*, p.360).

Cependant, ce qu'il est important de souligner, c'est qu'il ne s'agit pas ici d'une impossibilité pratique due, par exemple, au manque de finesse des équipements des scientifiques ou encore à une limite cognitive imposée par le cerveau humain; les théorèmes constituant le cœur des plus récents développements en physique quantique semblent impliquer qu'il n'est pas possible, par principe, que des causes communes agissant de manière spatiotemporellement continue puissent rendre compte des phénomènes de corrélation observés (*Ibid.*, p.361). Il est donc peut-être plus adéquat, et c'est d'ailleurs l'attitude qu'adoptent la plupart des scientifiques versés en physique quantique, de considérer qu'il n'est pas légitime de demander des explications causales telles que le voudrait Salmon pour rendre compte des phénomènes quantiques, comme le suggère Woodward: « Rather than concluding that there is an explanatory gap in quantum mechanics, those who adopt this view will conclude instead that it is our theory of explanation and our prior expectations that requests for explanation of a certain general kind are always in order which need revising » (*Ibid.*).

Effectivement, Salmon semble présupposer un certain réalisme pour les phénomènes quantiques selon lequel les différents paramètres quantiques des particules auraient des valeurs déterminées avant d'être mesurés, ce qui pourrait permettre l'existence de causes spatiotemporellement déterminées, alors qu'un tel réalisme ne semble aucunement s'accorder avec les résultats empiriques qu'obtiennent les scientifiques. Bref, la question du réalisme en physique quantique et de l'interprétation des résultats empiriques est évidemment très difficile, mais il reste que le modèle causal mécanique que propose Salmon ne semble pas bien adapté aux phénomènes quantiques, ce qui est certainement problématique étant donné que la physique quantique est au fondement de la science du 21^e siècle. Il faut toutefois spécifier que Salmon n'exclut pas la possibilité que son modèle causal mécanique puisse s'appliquer éventuellement à la physique quantique, malgré les

difficultés conceptuelles tout juste présentées²⁰; s'il est peut-être vrai que les explications causales ne sont pas appropriées pour certains types de phénomènes quantiques, il n'en reste pas moins que le modèle de Salmon pourrait très bien s'appliquer de manière limitée en physique quantique (Salmon, 1984, p.298).

2.6 Synthèse critique

Le modèle causal mécanique de l'explication scientifique qu'a développé Salmon, en tant qu'il s'insère dans une conception ontique de l'explication, se distingue surtout des deux autres conceptions épistémiques abordées dans le premier chapitre parce qu'il ne traite pas l'explication comme étant un argument. Au contraire, une explication, pour Salmon, est le dévoilement de la manière dont le phénomène à expliquer s'insère dans le monde naturel. L'intuition fondamentale au cœur du modèle de Salmon est que la notion d'explication est intimement liée à celle de causalité, en ce sens que c'est de manière causale que les phénomènes sont liés entre eux dans le monde. Ainsi, expliquer un phénomène, c'est rendre explicite le réseau causal dont il fait partie; c'est le situer précisément dans une toile causale. Salmon dira d'ailleurs que le passage d'une conception épistémique de l'explication à une conception ontique est en quelque sorte un saut gestaltique (*gestalt-switch*) en ce sens qu'il n'est plus question, avec l'explication scientifique, de l'anticipation des phénomènes (comme c'était le cas avec le modèle hempélien), mais plutôt de comprendre et de dénuder les mécanismes physiques qui sont au cœur des phénomènes (Salmon, 1984, p.301; 1985, p.651). Dans le premier chapitre, il a été dit que Salmon prétendait que le modèle déductif-nomologique était victime du problème des explications sans pertinence à cause de l'absence de la notion de causalité. Or, nous avons vu que le modèle causal mécanique de Salmon, bien qu'il soit centré sur les notions d'interaction et de processus causaux, était lui aussi aux prises avec une forme de ce même problème, à savoir celui de déterminer quelles propriétés étaient pertinentes du point de vue causal pour l'explication d'un événement donné.

De plus, ce modèle est confronté à certaines difficultés assez importantes lorsqu'il est question d'expliquer des phénomènes au sein de systèmes complexes ou de type quantique. Par ailleurs, le modèle de Salmon, parce qu'il est centré sur les mécanismes constituant les phénomènes, et donc sur l'histoire causale actuelle ou effective de ces phénomènes, met de côté, selon Woodward, un aspect épistémique important de l'explication: l'idée qu'une bonne explication devrait limiter, voire supprimer, le caractère arbitraire de l'explanandum, c'est-à-dire qu'une explication devrait être en

²⁰ Par ailleurs, Salmon souligne que personne ne comprend vraiment les mécanismes de réduction du paquet d'ondes, qui sont au cœur des phénomènes des corrélations de type EPR; il n'est donc en rien surprenant qu'il ne soit pas possible, avec son modèle, de les expliquer clairement! (Salmon, 1984, p.298).

mesure de montrer si, à partir des mêmes conditions initiales, plusieurs phénomènes auraient pu se produire, et si oui, pourquoi c'est ce phénomène plutôt qu'un autre qui en a résulté, ou encore si ce phénomène aurait pu être produit à partir d'une autre chaîne causale: « This feature is missed in an account like Salmon's which just focuses on describing the actual causal history leading up to the explanandum-outcome and leaves out questions about whether other possible but not actual histories would have led to a similar outcome » (Woodward, 1989, p.364).

En conclusion, bien que l'intuition selon laquelle la notion d'explication est intimement liée à celle de causalité soit sans doute puissante et conceptuellement fertile, il me semble que le modèle causal mécanique défendu par Salmon n'est pas complètement satisfaisant, peut-être parce qu'il ignore une certaine dimension pragmatico-épistémique de l'explication (entre autres, le rôle des contextes d'explication) sur laquelle insistera justement van Fraassen, comme nous le verrons dans le quatrième et dernier chapitre de ce mémoire. Cependant, il faudra d'abord comprendre l'essentiel de la théorie de la science que propose van Fraassen pour être en mesure de saisir adéquatement sa manière de concevoir l'explication scientifique. Or, alors que Salmon est un ardent défenseur du réalisme scientifique, van Fraassen soutient l'une des versions de l'antiréalisme scientifique les plus étudiées des dernières décennies, soit l'empirisme constructif. Je proposerai justement, dans le prochain chapitre, d'introduire les notions fondamentales qui définissent et qui caractérisent l'empirisme constructif.

3. L'empirisme constructif de Bas van Fraassen

3.1 Introduction

Bas van Fraassen, éminent philosophe des sciences contemporain, occupe une place absolument prépondérante dans les débats épistémologiques actuels. Cela n'est pas surprenant si l'on tient compte du fait qu'il est à l'origine d'une théorie antiréaliste de la science qu'il a nommée « empirisme constructif », théorie qui, selon plusieurs, a redonné un souffle de vie à l'antiréalisme scientifique qu'on croyait mort après le déclin du positivisme logique. Bien que cette doctrine n'ait pas nécessairement une grande quantité de défenseurs, il reste qu'elle est à la source de plusieurs discussions qui contribuent certainement à l'avancement et au développement de la philosophie des sciences contemporaine.

La conception de l'explication scientifique que propose van Fraassen, qui sera étudiée dans le quatrième et dernier chapitre de ce mémoire, s'insère évidemment dans le cadre plus large que définit l'empirisme constructif. C'est pourquoi le présent chapitre sera dédié à une présentation introductive de cette théorie. Bien entendu, il ne sera pas possible de l'étudier de manière détaillée (cela demanderait d'y consacrer un mémoire en entier), mais cela suffira pour bien comprendre le contexte dans lequel van Fraassen développe son modèle de l'explication scientifique. Premièrement, quelques brèves considérations générales sur le projet de van Fraassen seront abordées, afin de mieux saisir le sens général de l'empirisme constructif. En deuxième lieu, je présenterai la définition ou la thèse principale de l'empirisme constructif que propose van Fraassen et je tenterai de l'expliquer et de la développer brièvement en abordant certaines de ses implications importantes, notamment en ce qui concerne la notion capitale d'adéquation empirique. Pour ce faire, il sera sans doute fort utile de montrer en quoi l'empirisme constructif diffère du réalisme scientifique; d'ailleurs, van Fraassen, dans *The Scientific Image*, consacre plusieurs pages à critiquer le réalisme scientifique dans le but de bien faire ressortir ce qui distingue la théorie qu'il propose. En outre, la notion d'observabilité sera explorée, étant donné qu'elle est absolument essentielle pour bien comprendre la portée de l'idée d'adéquation empirique et, par le fait même, de la thèse centrale de l'empirisme constructif. En dernier lieu, je présenterai une objection à l'empirisme constructif soulevée par James Ladyman, qui concerne justement cette notion d'observabilité et le réalisme modal.

3.2 Remarques préalables

Dans l'un de ses ouvrages majeurs, *The Scientific Image*, van Fraassen présente sa doctrine de l'empirisme constructif. Avant d'aborder la définition officielle qu'il propose, il n'est pas sans intérêt de remarquer que van Fraassen définit l'empirisme, de manière générale, comme étant « the philosophical position that experience is a source of information about the world, and our only source » (*Laws and Symmetry*, p.8). Van Fraassen, comme les autres penseurs qu'on considère de la tradition empiriste, veut éviter le plus possible les considérations métaphysiques ou d'ordre *a priori* dans son enquête sur la connaissance, particulièrement en ce qui concerne la philosophie des sciences (Rosen, 1994, p.144). L'une des implications majeures d'une telle vision des choses est que le domaine de la science ne concerne que le monde des phénomènes observables. Bien sûr, les théories scientifiques postulent une multitude de processus et d'entités qui ne sont pas directement observables afin de rendre compte des phénomènes. Le point de van Fraassen est justement de dire que l'empirisme ne requiert de ces théories qu'elles ne soient vraies qu'à propos des choses observables; les structures et entités postulées ne sont qu'un moyen pour arriver à cette fin (van Fraassen, 1980, p.3). Bref, il faut garder en tête que c'est devant cet arrière-plan épistémique que s'insère sa doctrine plus précise qu'est celle de l'empirisme constructif.

En outre, il importe de préciser brièvement le cadre général dans lequel s'insère le développement de l'empirisme constructif par van Fraassen. Dans la préface de *The Scientific Image*, il souligne qu'il a, comme objectif premier, d'offrir une alternative concluante au réalisme scientifique. Pour ce faire, van Fraassen propose de diviser son entreprise en trois volets principaux, dans chacun desquels il développera une sous-théorie constitutive de l'empirisme constructif: « The first concerns the relation of a theory to the world, and especially what may be called its empirical import. The second is a theory of scientific explanation, in which the explanatory power of a theory is held to be a feature which does indeed go beyond its empirical import, but which is radically context-dependent. And the third is an explication of probability as it occurs within physical theory [...] » (1980, vii). Dans le présent chapitre, il s'agira d'aborder la première de ces sous-théories, c'est-à-dire de comprendre la relation qu'entretient une théorie scientifique avec le monde au sein de l'empirisme constructif. Le deuxième volet, celui concernant l'explication scientifique, sera traité dans le chapitre suivant. Nous devons malheureusement laisser tomber le troisième aspect qui intéresse van Fraassen²¹.

²¹ Contrairement au premier volet qui concerne la relation d'une théorie scientifique au monde, cette sous théorie concernant la notion de probabilité physique n'est pas très pertinente pour bien comprendre le modèle de l'explication scientifique que propose van Fraassen.

3.3 Définition et caractérisation de l'empirisme constructif

3.3.1 Thèse du réalisme scientifique

Van Fraassen donne une courte définition de l'empirisme constructif, qu'il va par la suite expliquer et développer dans son ouvrage. Avant même de donner cette définition, van Fraassen propose ce qu'il estime être la thèse principale du réalisme scientifique, qu'il considère comme étant opposée à celle de l'empirisme constructif, justement afin de permettre de bien cerner le contraste entre ces deux visions de la science. Pour bien saisir ce qu'est l'empirisme constructif, en effet, il est primordial de comprendre de quelle manière il diffère assez radicalement d'une vision réaliste de la science; c'est pourquoi je propose, avant de nous pencher plus sérieusement sur la définition de l'empirisme constructif, de nous arrêter quelques instants sur la manière dont van Fraassen comprend ce qu'est le réalisme scientifique.

Selon van Fraassen, le réalisme scientifique, en philosophie des sciences, est une position à propos de la manière dont une théorie scientifique devrait être comprise et de la nature de l'activité scientifique elle-même (*Ibid.*, p.6). Ainsi, le réalisme scientifique concerne directement le premier volet de l'enquête de van Fraassen qui a été évoqué plus haut, à savoir la question du rapport de la théorie scientifique au monde. Voici la thèse qui, selon van Fraassen, définit le réalisme scientifique (*Ibid.*, p.8):

RS: *La science vise à fournir, de par ses théories, une description vraie et littérale de ce qu'est le monde; et l'acceptation d'une théorie scientifique implique la croyance qu'elle est vraie*²².

Cette thèse comporte deux éléments: elle précise ce qu'est le but de la science dans sa production de théories et, par ailleurs, elle spécifie ce qui est impliqué par l'acceptation d'une théorie scientifique. En ce qui regarde le premier élément, il est important de noter que van Fraassen, lorsqu'il parle de la visée de la science, n'a pas en tête les objectifs ou les buts particuliers que poursuivent les scientifiques individuels²³. Il s'agit plutôt de la visée de l'entreprise ou de l'institution qu'est la science elle-même, et c'est une telle visée générale qui permet de déterminer ce qu'est une réussite ou une avancée scientifique. Ainsi, dans la perspective du réalisme scientifique, la science a pour

²² « Science aims to give us, in its theories, a literally true story of what the world is like; and acceptance of a scientific theory involves the belief that it is true ».

²³ Van Fraassen établit une comparaison avec le jeu d'échecs qui me semble éclairante: le but général du jeu est de placer son adversaire échec et mat, mais un joueur particulier peut jouer aux échecs pour toute sorte de raisons (plaisir, argent, etc.).

but de fournir des théories littéralement²⁴ vraies à propos du monde, y compris en ce qui concerne les entités inobservables qu'elle postule, comme les atomes ou les électrons. Ainsi, comme le fait remarquer Ladyman, les réalistes scientifiques croient en l'existence des entités théoriques inobservables, en ce sens qu'ils croient que des tokens des types postulés par les théories scientifiques existent dans les faits (Ladyman, 2000, p.838). Par ailleurs, de dire que l'acceptation d'une théorie scientifique implique la croyance qu'elle est vraie ne revient pas à dire que les réalistes croient que les théories actuelles sont totalement vraies. Van Fraassen pense même qu'un réaliste scientifique n'est pas obligé de penser que les théories scientifiques seront un jour vraies dans tous leurs aspects (1980, p.7). En effet, l'acceptation²⁵, au même titre que la croyance, est une affaire de degré; un certain degré d'acceptation d'une théorie scientifique n'implique qu'un certain de degré de croyance en sa vérité.

3.3.2 Thèse de l'empirisme constructif et la notion d'adéquation empirique

La thèse du réalisme scientifique maintenant brièvement expliquée, il sera plus aisé de comprendre la thèse centrale de l'empirisme constructif, que van Fraassen énonce comme suit (1980, p.12):

EC: La science vise à fournir des théories empiriquement adéquates; et l'acceptation d'une théorie implique seulement comme croyance qu'elle est empiriquement adéquate²⁶.

D'emblée, ce qu'on remarque, c'est que dans cette thèse, la notion d'adéquation empirique semble remplacer la notion de vérité qui est était centrale dans la thèse du réalisme scientifique. De fait, l'adéquation empirique est une idée absolument fondamentale dans l'empirisme constructif. Par ailleurs, les remarques ayant été faites à propos de la visée de la science et des liens entre l'acceptation et la croyance, lorsqu'il était question du réalisme scientifique, sont tout aussi valides dans le cas de l'empirisme constructif (*Ibid.*). Van Fraassen précise aussi que dans l'empirisme constructif, à l'instar du réalisme scientifique, les théories scientifiques doivent être interprétées de

²⁴ Van Fraassen ne s'étend pas trop sur la notion de littéralité, mais il précise qu'il s'agit de distinguer le réalisme scientifique (tout comme l'empirisme constructif, ainsi que j'aurai l'occasion d'y revenir) d'autres approches qui soutiendront que les théories scientifiques produisent des énoncés qui sont vrais ou faux lorsque traduits dans des énoncés de type particulier (par exemples, des énoncés empiriquement vérifiables dans le cas de l'empirisme logique) ou interprétés de manière non standard. Autrement dit, une théorie est interprétée littéralement lorsque les énoncés qui la constituent possèdent une valeur de vérité sans être préalablement traduits dans un langage différent.

²⁵ La notion d'acceptation mise de l'avant par van Fraassen sera explorée davantage un peu plus bas.

²⁶ « Science aims to give us theories which are empirically adequate; and acceptance of a theory involves as belief only that it is empirically adequate ».

manière littérale. Attardons-nous donc à l'idée d'adéquation empirique pour être en mesure de bien comprendre la thèse de l'empirisme constructif.

L'adéquation empirique ne concerne que les phénomènes observables du monde. Alors qu'il a été dit que les partisans du réalisme scientifique accordent une importance particulière à l'idée de vérité, notamment en ce qui concerne les entités théoriques inobservables postulées par une théorie scientifique, l'adéquation empirique n'est relative qu'à la vérité d'une théorie à propos de l'observable²⁷: « a theory is empirically adequate exactly if what it says about the observable things and events in this world, is true — exactly if it 'saves the phenomena' » (*Ibid.*, p. 12). Il est important de noter que l'adéquation empirique ne se limite pas à ce qui est observé actuellement ou encore à ce qui a déjà été ou sera observé, mais elle concerne tous les phénomènes observables, qu'ils soient dans les faits observés ou non: « I must emphasize that this refers to *all* the phenomena; these are not exhausted by those actually observed, nor even by those observed at some time, whether past, present, or future » (*Ibid.*), ce que souligne aussi Gideon Rosen (1994, p.160).

Par ailleurs, van Fraassen précise davantage ce que signifie, pour une théorie, d'être empiriquement adéquate en élaborant une conception sémantique des théories scientifiques, qu'il oppose à une vision syntaxique. Alors que dans cette dernière vision, une théorie consiste en un ensemble de théorèmes énoncés dans un langage particulier, l'approche sémantique considère qu'une théorie est un ensemble de structures qui constituent les modèles de cette théorie (van Fraassen, 1980, p.44). Ces modèles sont eux-mêmes constitués par ce que van Fraassen appelle des substructures empiriques (*empirical substructures*), qui sont en mesure d'être des représentations directes des phénomènes observables. Une théorie empiriquement adéquate en sera donc une qui comporte des modèles qui ont des substructures capables de représenter tous les phénomènes observables, comme l'explique van Fraassen:

To present a theory is to specify a family of structures, its *models*; and secondly, to specify certain parts of those models (the *empirical substructures*) as candidates for the direct representation of observable phenomena. The structures which can be described in experimental and measurement reports we can call *appearances*: the theory is empirically adequate if it has some model such that all appearances are isomorphic to empirical substructures of that model (*Ibid.*, p.64).

Autrement dit, une théorie est empiriquement adéquate non pas quand les phénomènes peuvent être décortiqués de sorte que les théorèmes fondamentaux de ladite théorie puissent en rendre compte (ce qui serait, en gros, la vision syntaxique des théories scientifiques), mais plutôt quand les

²⁷ Le concept d'observabilité est philosophiquement problématique; nous y reviendrons un peu plus bas.

phénomènes peuvent être reconnus (selon des relations isomorphes) dans l'une des substructures empiriques d'un modèle de la théorie.

Il me paraît intéressant de souligner, à ce point, que la notion d'adéquation empirique au cœur de l'empirisme constructif est plus modeste, du point de vue épistémique, que ne l'est la notion de vérité qui fait partie intégrante du réalisme scientifique, c'est-à-dire qu'elle exige un degré de croyance moins important. En effet, alors que la vérité concerne le monde dans sa totalité, l'adéquation empirique, pour le répéter, ne concerne que les phénomènes observables, même si van Fraassen est bien conscient que les théories scientifiques vont bien au-delà de ces phénomènes observables: « My view is that physical theories do indeed describe much more than what is observable, but that what matters is empirical adequacy, and not the truth or falsity of how they go beyond the observable phenomena » (1980, p.64). L'idée de van Fraassen n'est donc aucunement que les scientifiques ne devraient pas construire de théories incorporant des entités inobservables. Il s'agit de dire que la croyance impliquée par l'acceptation d'une théorie scientifique ne concerne que ce qu'elle dit à propos de l'observable. L'empirisme constructif n'implique pas la thèse métaphysique que les entités théoriques inobservables n'existent pas; il s'agit de dire que l'acceptation d'une théorie scientifique n'implique en rien la croyance en l'existence des entités théoriques inobservables qu'elle postule.

J'aimerais remarquer que cette modestie épistémique, si on peut l'appeler ainsi, va tout à fait dans le sens de la philosophie empiriste de van Fraassen, telle que caractérisée au début du présent chapitre. En effet, en s'en tenant à l'adéquation empirique, van Fraassen semble limiter le plus possible le dépassement de l'expérience empirique (en tant que seule source de connaissance) tout en étant capable de rendre compte de l'activité scientifique, comme l'explique Rosen: « Of the available stances towards science, van Fraassen's has some claim to the minimal non-revisionary stance. It is the most cautious position - the position most compatible with empiricism in the broad sense - which at the same time renders the phenomena of scientific activity fully intelligible and ratifiable » (1994, p.162). Bref, la croyance en l'adéquation empirique d'une théorie, étant moins compromettante que la croyance en sa vérité, permet à l'empiriste constructif de se rapprocher le plus possible de l'esprit de la philosophie empiriste tout en rendant compte de façon plausible de l'activité scientifique, notamment parce qu'il limite les considérations métaphysiques invasives (van Fraassen, 1980, p.73). C'est pourquoi, selon van Fraassen, l'empiriste conséquent avec lui-même se doit d'adopter l'empirisme constructif (Rosen, 1994, p.159; Alspector-Kelly, 2001, p.415).

Il faudrait toutefois éviter de penser que la notion d'adéquation empirique respecte à la lettre les limites épistémiques imposées par l'empirisme. En effet, il a déjà été dit que les phénomènes observables, concernés par l'adéquation empirique, regroupent tous les phénomènes observables passés, présents et futurs. Or, évidemment, il est par principe impossible que tous ces phénomènes soient un jour observés (et non plus seulement observables). Il semble donc que tout comme la notion de vérité, la notion d'adéquation empirique va au-delà des frontières épistémiques définies par la philosophie empiriste, comme l'admet d'ailleurs van Fraassen. Cependant, il est très clair à cet effet: même si ces deux notions dépassent le cadre empiriste, il reste que celle d'adéquation empirique est à préférer, car elle est beaucoup plus faible du point de vue épistémique: « In either case we stick our necks out: empirical adequacy goes far beyond what we can know at any given time. (All the results of measurement are not in; they will never all be in; and in any case, we won't measure everything that can be measured.) Nevertheless there is a difference: the assertion of empirical adequacy is a great deal weaker than the assertion of truth, and the restraint to acceptance delivers us from metaphysics » (van Fraassen, 1980, p.69).

Van Fraassen ne donne pas de justification étoffée de cette dernière affirmation selon laquelle la notion d'adéquation empirique est à préférer par rapport à celle de vérité malgré que les deux aillent plus loin que ce qui est strictement permis par l'empirisme. Il écrit toutefois, et cet extrait est désormais fameux dans la communauté épistémologique, que « it is not an epistemological principle that one might as well hang for a sheep as for a lamb » (*Ibid.*, p.72), c'est-à-dire qu'il faut éviter l'attitude selon laquelle il n'est pas plus grave de commettre une faute sérieuse (dans ce cas, prêcher pour la notion de vérité) si l'on sait que l'on commet une faute (dans ce cas, prêcher pour la notion d'adéquation empirique) de toute manière. Les deux notions vont peut-être au-delà des limites épistémiques définies par l'empirisme, mais il reste que celle d'adéquation empirique les dépasse dans une bien moindre mesure, ce qui est loin d'être négligeable lorsque vient le temps de défendre une position épistémologique dans le cadre plus large d'une philosophie empiriste.

Pour clore cette section, nous avons déjà dit quelques mots sur la notion d'acceptation, mais comme elle est centrale dans la définition de l'empirisme constructif que propose van Fraassen, je pense qu'il serait adéquat de l'explorer de manière un peu plus détaillée. L'aspect épistémique de l'acceptation d'une théorie scientifique a déjà été touché lorsqu'il a été question de dire que l'acceptation, dans le cas de l'empirisme constructif, n'impliquait que la croyance que cette théorie était empiriquement adéquate: « To accept the theory involves no more belief, therefore, than that what it says about observable phenomena is correct » (*Ibid.*, p.57). Il a aussi été dit que l'acceptation est une notion qui comporte des degrés, comme c'est le cas de celle de croyance; il est donc possible

d'accepter plus ou moins une théorie scientifique, et donc de croire plus ou moins qu'elle est empiriquement adéquate. Toutefois, il faut comprendre qu'un certain degré de croyance en l'adéquation empirique d'une théorie n'est pas équivalent à la croyance que la théorie est approximativement empiriquement adéquate; cette dernière option, selon van Fraassen, reviendrait à la croyance que seules certaines parties de la théorie sont parfaitement empiriquement adéquates (*Ibid.*, p.9). Bien que van Fraassen ne semble pas préciser davantage ce qu'il entend par « un degré de croyance en l'adéquation empirique d'une théorie », j'ai l'impression qu'il s'agit de dire que nous croyons qu'une théorie scientifique, selon les phénomènes que nous connaissons et selon la moins grande plausibilité des théories rivales, est empiriquement adéquate, mais que nous croyons aussi qu'il est possible que d'autres phénomènes qui n'ont pas encore été étudiés ne cadrent pas parfaitement dans cette théorie.

Cependant, il pourrait être tentant de penser qu'un adepte de l'empirisme constructif qui croirait en la vérité d'une théorie scientifique qu'il accepte se retrouve dans une position intenable. Mais il n'en est rien; c'est ici que la distinction entre l'agnosticisme et le gnosticisme scientifiques que propose van Fraassen prend tout son sens. Un agnostique scientifique croit en l'adéquation empirique de la théorie scientifique qu'il accepte, mais suspend son jugement quant à sa vérité (il ne croit ni qu'elle est vraie, ni qu'elle est fausse), alors qu'un gnostique scientifique croit en la vérité de la théorie scientifique qu'il accepte. L'idée de van Fraassen, c'est qu'un tenant du réalisme scientifique, tel que défini plus haut, sera nécessairement gnostique; pour un réaliste, un agnostique scientifique ne comprend pas adéquatement ce qu'est la nature de la science. Toutefois, un empiriste constructif a bien sûr la possibilité d'être un agnostique scientifique, mais aussi d'être gnostique. En effet, rien n'empêche un défenseur de l'empirisme constructif d'adopter des croyances qui vont au-delà de ce qu'est le but de la science. Autrement dit, un gnostique scientifique qui accepte l'empirisme constructif comprend qu'il adopte des croyances qui dépassent les visées de l'entreprise scientifique, mais cela n'est en soi, selon van Fraassen, aucunement problématique (1998, pp.213-214).

En outre, la notion d'acceptation comporte un aspect pragmatique considérable, expliqué dans le passage suivant:

Theory acceptance has a pragmatic dimension. While the only belief involved in acceptance, as I see it, is the belief that the theory is empirically adequate, more than belief is involved. To accept a theory is to make a commitment, a commitment to the further confrontation of new phenomena within the framework of that theory, a commitment to a research programme, and a wager that all relevant phenomena can be accounted for without giving up that theory (van Fraassen, 1980, p.88).

Qu'est-ce qui motive cet engagement dans un programme de recherche, et cet effort d'utiliser toutes les ressources possibles d'une théorie pour rendre compte du plus de phénomènes possible, si ce n'est uniquement la croyance en sa vérité (pour les réalistes) ou en son adéquation empirique (pour les empiristes constructifs)? Van Fraassen soutient que certaines valeurs pragmatiques, non épistémiques, jouent un rôle prépondérant pour l'acceptation d'une théorie scientifique (*Ibid.*; 2007, p.340). Ces valeurs, en effet, ne concernent pas les rapports entre la théorie et le monde, mais plutôt l'utilité et l'usage de cette théorie, par exemple en ce qui concerne la capacité de s'en servir pour élaborer des explications²⁸. Finalement, van Fraassen ne considère pas que seul l'empirisme constructif soit en mesure d'accorder une place à ces valeurs pragmatiques. Il reconnaît tout à fait que le réaliste scientifique puisse accepter une telle dimension pragmatique de l'acceptation d'une théorie. Seulement, l'engagement épistémique qu'exige le réalisme scientifique (le degré de croyance) étant plus important que celui qu'exige l'empirisme constructif, van Fraassen remarque que les défenseurs de ce dernier (et des autres positions antiréalistes, de manière plus générale) ont tendance à accorder une plus grande place à la dimension pragmatique dans leurs réflexions épistémologiques (1980, p.13).

3.3.3 Observabilité

La notion d'adéquation empirique, nous venons de le voir, concerne l'isomorphie entre les phénomènes observables et certaines substructures constituant les modèles d'une théorie scientifique. Il est donc manifeste que la notion d'observabilité, en étant centrale pour celle d'adéquation empirique, est très importante dans l'empirisme constructif. Mais c'est aussi l'une des notions les plus problématiques développées par van Fraassen; plusieurs critiques faites à l'endroit de l'empirisme constructif sont fondées sur les difficultés conceptuelles soulevées par la manière dont van Fraassen traite de l'observabilité. Ainsi, penchons-nous d'abord sur la notion d'observabilité telle qu'exposée dans *The Scientific Image*.

La première chose à dire est que van Fraassen, dans *The Scientific Image*, ne donne pas de définition claire et précise de l'observabilité. Il propose plutôt la formulation suivante, qui se veut un principe simplificateur et opérateur pour caractériser cette notion: « X is observable if there are circumstances which are such that, if X is present to us under these circumstances, then we observe it » (1980, p.16). Si van Fraassen ne propose pas de définition comme telle, c'est parce qu'il considère que « observable » est un prédicat vague, c'est-à-dire qu'il y a un continuum entre

²⁸ D'ailleurs, le quatrième chapitre de ce mémoire sera consacré au modèle de l'explication scientifique chez van Fraassen, qui accorde une grande importance à la pragmatique dans l'explication.

l'observable et l'inobservable, et qu'il n'est donc pas possible de séparer radicalement ces deux notions. Cependant, il suffit qu'il y ait des cas clairs d'observation, ainsi que des contre-exemples évidents, pour que cette notion soit utilisable, et van Fraassen soutient bien sûr que c'est le cas. Il faut aussi préciser dès maintenant que pour un empiriste constructif, une observation se fait toujours sans l'aide d'outils pour affiner nos perceptions (comme un microscope ou un télescope). Toutefois, est-ce à dire que quelqu'un qui regarderait, par exemple, les lunes de la planète Jupiter à l'aide d'un télescope, regarderait des objets inobservables? Non, « since astronauts will no doubt be able to see them as well from close up » (*Ibid*), c'est-à-dire qu'il serait en principe tout à fait possible pour les astronautes d'observer les lunes de Jupiter sans outil d'aide à la perception s'ils étaient à une distance appropriée de ces lunes. De même, l'étude d'une paramécie avec un microscope ne constitue pas une observation de la paramécie; il s'agit plutôt de l'observation d'une image de la paramécie. C'est pourquoi Gabriele Contessa affirme que dans cette perspective, l'invention d'outils tels que le microscope n'a pas permis d'augmenter l'étendue des phénomènes observables, mais plutôt le nombre de phénomènes observables à expliquer (comme l'image de la paramécie) (Contessa, 2006, p.456).

En outre, van Fraassen soutient que la notion d'observabilité est toujours relative à une communauté épistémique donnée: « what counts as an observable phenomenon is a function of what the epistemic community is (that *observable* is *observable-to-us*) » (1980, p.19). Dans le cas qui nous concerne, la communauté épistémique en question est celle formée par l'ensemble des êtres humains; mais il n'y a rien qui empêcherait, par principe, l'intégration d'autres entités dans notre communauté épistémique sur la base de considérations idéologiques ou morales (*Ibid.*, p.18). Dans un tel cas, il serait possible que l'extension du concept d'observabilité ne soit plus exactement la même, et donc, par le fait même, que la notion d'adéquation empirique ne concerne pas exactement les mêmes phénomènes. La thèse centrale de l'empirisme constructif étant, pour le rappeler, que l'acceptation d'une théorie scientifique n'implique que la croyance qu'elle est empiriquement adéquate, il s'ensuit que nos attitudes épistémiques vis-à-vis d'une théorie scientifique vont reposer, en fin de compte, sur la distinction entre l'observable et l'inobservable²⁹. Cela s'accorde d'ailleurs avec la position empiriste de van Fraassen: l'expérience étant notre seule source de connaissance, les phénomènes que l'on peut expérimenter (i.e., les phénomènes observables) vont nécessairement influencer très fortement nos attitudes épistémiques. Il n'est toutefois pas ici question de donner une portée ontologique à la distinction observable-inobservable, par exemple d'affirmer que seules les

²⁹ Dans chaque cas particulier, nos attitudes épistémiques par rapport à une théorie scientifique à propos d'un phénomène donné vont résulter de l'application concrète de cette distinction entre l'observable et l'inobservable pour ce phénomène particulier.

entités observables sont vraiment réelles. Van Fraassen remarque que cette idée selon laquelle les limites de notre perception devraient être prises en considération pour déterminer nos attitudes épistémiques à l'endroit des théories scientifiques est habituellement rejetée par les réalistes (1985, p.258), pour qui la science nous dit la vérité à propos du monde indépendamment de nos limites perceptives.

Par ailleurs, les membres de la communauté épistémique, le « nous » à qui s'adressent les théories scientifiques, font eux-mêmes partie des objets auxquels s'intéressent ces théories (1980, p.59). Ainsi, les limites perceptives des sujets épistémiques peuvent être étudiées par la science. C'est donc dire que c'est la théorie scientifique *T* elle-même, acceptée par la communauté épistémique et décrivant le monde, qui définira les limites de ce qui est observable, comme l'explique van Fraassen: « If there are limits to observation, these are a subject for empirical science, and not for philosophical analysis. [...] To find the limits of what is observable in the world described by theory *T* we must inquire into *T* itself [...] » (1980, p.57). Autrement dit, les limites de ce qui est observable ne peuvent aucunement être déterminées *a priori*, et c'est donc la science qui, en fin de compte, délimitera le domaine de l'observable, comme le soulignent Stephen Leeds: « it is our scientific theories themselves which tell us [...] what parts of the world we can observe » (1994, p.187) et Rosen: « the bounds of the observable are determined by science itself, and not a priori by some transempirical means » (1994, p.171). En effet, ce sont les sciences comme la biologie et la physique qui sont aptes à décrire nos capacités et limitations en tant qu'appareils de mesure (van Fraassen, 1980, p.17). Cependant, il ne faudrait pas confondre cette idée avec celle selon laquelle la détermination de ce qui est observable est *dépendante* de la théorie scientifique acceptée elle-même. Il s'agit plutôt de dire que la question de savoir ce qui est observable est une question empirique factuelle qui doit être étudiée par la science, en ce qu'elle dépend des facultés perceptuelles des membres de la communauté épistémique, mais elle reste indépendante de la théorie. Il faut donc distinguer « être dépendant de la théorie » et « être descriptible par la théorie ». Van Fraassen est très clair à ce sujet dans ce passage:

To delineate what is observable, however, we must look to science—and possibly to that same theory—for that is also an empirical question. This might produce a vicious circle if what is observable were itself not simply a fact disclosed by theory, but rather theory-relative or theory-dependent. It will already be quite clear that I deny this; I regard what is observable as a theory-independent question. It is a function of facts about us *qua* organisms in the world, and these facts may include facts about the psychological states that involve contemplation of theories—but there is not the sort of theory-dependence or relativity that could cause a logical catastrophe here (1980, pp.57-58).

Dans cette citation, van Fraassen évoque la possibilité d'un cercle vicieux qui pourrait naître de l'idée que ce soit les théories scientifiques qui déterminent l'extension du domaine de l'observable, alors même que la croyance en la vérité d'une théorie scientifique ne concerne justement que ce qui est observable. Dans un tel cas, non seulement la notion d'observabilité deviendrait inutile³⁰, mais la notion d'adéquation empirique perdrait de son intérêt, car elle varierait selon la théorie scientifique concernée, et il deviendrait donc impossible de comparer l'adéquation empirique de deux théories concurrentes. Van Fraassen rejette cette possibilité étant donné que la détermination de l'observable reste tout de même une question factuelle indépendante des théories scientifiques (seulement, cette question doit être étudiée par la science).

3.4 Objection de Ladyman: observabilité et réalisme modal

Plusieurs auteurs³¹ ont soulevé des objections à van Fraassen concernant le concept d'observabilité ; en fait, c'est une des notions développées par van Fraassen les plus critiquées. L'une de ces objections, il me semble, est particulièrement intéressante: il s'agit de celle mise de l'avant par James Ladyman (2000), selon laquelle la manière de traiter de l'observabilité proposée par van Fraassen implique un certain réalisme modal, réalisme qui devrait certainement être rejeté par les empiristes constructifs. Rappelons-nous que van Fraassen caractérise l'observabilité en disant que X est observable s'il existe des circonstances dans lesquelles si nous étions en la présence de X, alors nous observerions X.

L'argument de Ladyman est assez complexe, et pour être présenté dans le détail, il faudrait d'abord proposer une analyse fine de la modalité chez van Fraassen, ce qui dépasserait le propos du présent chapitre. Il est toutefois possible d'en comprendre les grandes lignes avec ce qui a été expliqué dans les paragraphes précédents. En effet, Ladyman débute son argument en soulevant deux questions: « (a) is it a theory-independent fact that if X was present to us we would observe it? ; and (b) if so how can we know such a fact? » (2000, p.850). Lorsqu'il était question de la notion d'observabilité, nous avons vu que van Fraassen considérait que la question de savoir ce qui est observable est empirique, en ce sens que c'est la science qui peut nous éclairer à ce sujet, mais que c'était tout de même une question indépendante de la théorie, car ce qui observable dépend de faits qui ne dépendent pas de la théorie; seulement, la théorie scientifique peut décrire ces faits. Ainsi, il faut répondre à (b) en expliquant que ce sont les sciences, comme la biologie et la physique, qui peuvent nous décrire ce qui est observable en étudiant certains faits du monde (nos organes perceptifs, etc.).

³⁰ Car la notion d'observabilité deviendrait alors radicalement dépendante de la théorie scientifique étudiée.

³¹ Par exemple, Churchland (1985) et Hacking (1985).

Et il faut bien sûr répondre à (a) par l'affirmative, car il s'agit d'un fait objectif et indépendant de la théorie que si nous étions en la présence de X, alors nous l'observerions (Ladyman, 2000, p.851).

Pour bâtir son argument, Ladyman propose des exemples d'objets que van Fraassen considérerait comme observables, mais qu'on pourrait observer dans des circonstances qui n'existeraient jamais, par exemple, les dinosaures; ce sont donc des exemples qui peuvent être formulés de manière contrefactuelle (*Ibid*, p.851). Ladyman soutient donc que les empiristes constructifs, s'ils veulent à la fois rendre compte de l'observable indépendamment ce qui a été observé et respecter l'importance épistémologique de la notion d'observabilité³², sont obligés de croire en au moins certains énoncés contrefactuels qu'impliquent les théories scientifiques. C'est exactement ce que remarque aussi Dicken: « [t]here are some states of affairs that will never actually be observed, yet which constructive empiricism would still classify as observable. The constructive empiricist is thus committed to believing certain *counterfactual* claims made by his theories [...] and thus believing some of the modal implications of his theories » (2007, p.605). Or, cela est problématique, car van Fraassen refuse d'accorder une valeur de vérité objective aux énoncés contrefactuels, car selon lui, leur vérité dépend toujours du contexte particulier dans lequel ils sont formulés³³ (van Fraassen, 1980, p.13).

Par ailleurs, il est ici question d'implications modales, en ce sens que les énoncés contrefactuels qui sont problématiques sont ceux qui portent sur des phénomènes qui n'existeront jamais dans le monde actuel³⁴, mais qui doivent tout de même être considérés comme vrais par les empiristes constructifs. Il semble donc que van Fraassen soit contraint de croire en un certain réalisme modal, selon lequel certains énoncés contrefactuels sont objectivement vrais, ce qui s'accorde très mal avec son souci empiriste de limiter le plus possible les considérations métaphysiques. Ladyman résume son argument comme suit: « To summarize, although the constructive empiricist need not be committed to the full truth of theories to demarcate the observable, she is committed to believe in some of their modal implications. In other words, she is committed to belief in more than just what theories say about what is observable and actual, in order to discern what they do say about what is observable and actual » (2000, pp.851-852).

³² Cette importance épistémologique découle du fait que l'observabilité est cruciale pour la définition de la notion d'adéquation empirique.

³³ Un peu plus bas, nous reviendrons sur l'analyse des énoncés contrefactuels proposée par van Fraassen.

³⁴ « Actuel » ne signifie pas ici « dans le temps présent », mais se dit d'un phénomène tel qu'il a existé, existe ou existera, ce qui s'oppose à un phénomène qui *aurait pu* se produire dans des circonstances différentes. Van Fraassen écrit d'ailleurs, à propos des phénomènes actuels, qu'ils sont « what does happen, and not, what would happen under different circumstances » (1980, p.60).

Bref, Ladyman voit une incohérence, car la thèse centrale de l'empirisme constructif est que l'acceptation d'une théorie scientifique n'implique que la croyance en la vérité de ce qu'elle dit à propos des phénomènes observables. Or, les énoncés contrefactuels tels que « Si un dinosaure était à deux mètres de moi, je l'observerais » ne portent pas sur des phénomènes observables, car les circonstances qui feraient que les dinosaures puissent être observés (entre autres, l'existence simultanée des dinosaures et des êtres humains) sont contrefactuelles et non actuelles. La croyance en de tels énoncés contrefactuels, pour Ladyman, « is totally incompatible with constructive empiricism [...] since that would amount to allowing that scientific theories tell us about more than the actual phenomena » (*Ibid.*, p.852).

Van Fraassen et son collègue Bradley Monton proposent une réponse à cette objection de Ladyman. En effet, ils insistent sur le fait que l'observabilité, contrairement à ce qu'affirme Ladyman, ne doit pas être comprise comme une propriété modale (2003, p.411). En outre, van Fraassen soutient effectivement que les énoncés contrefactuels n'ont pas de valeur de vérité objective, car ils sont toujours dépendants d'un contexte particulier: « The sense in which counterfactuals are here held not to have an objective truth value is that they are in general context-dependent » (*Ibid.*, p.410). Les énoncés contrefactuels peuvent donc avoir une valeur de vérité une fois les données contextuelles adéquatement spécifiées. Il devient alors possible de les comprendre comme des énoncés conditionnels non modaux. Par exemple, dans le cas des énoncés contrefactuels au sujet de l'observabilité, il devient possible, en spécifiant leur contexte (notamment les informations à propos de la communauté épistémique de l'observateur), de les comprendre de manière non modale: « But in principle, once the epistemic community is specified, the answer to the question is implied by facts about the world. ('Facts' should be understood to be empirical, non-modal, and theory-independent) » (*Ibid.*, p.414). Autrement dit, une fois l'énoncé contrefactuel compris comme un simple conditionnel dont le contexte est fixé, il devient possible d'accorder une valeur de vérité à un énoncé de type « X est observable »³⁵. L'idée de Monton et van Fraassen est donc qu'un énoncé contrefactuel à propos de l'observable perd sa nature modale une fois le contexte déterminé; il devient alors un énoncé factuel qui peut être étudié par la science. Monton et van Fraassen résument leur réponse à Ladyman de cette façon:

Consider the claim 'if the moons of Jupiter were present to us (in the right kind of circumstances) then we would observe them'. Ladyman ([2000], p. 851) thinks van Fraassen would analyse the claim in terms of possible worlds. [...] The way to understand the claim is to note that, even though it is a counterfactual, it is entailed by facts about the world: facts that the moons of Jupiter are constituted in a certain way,

³⁵ Monton et van Fraassen (2003)

and facts that we are constituted in a certain way. These facts can be disclosed by empirical research (*Ibid.*, p.415).

Reconsidérons l'énoncé « Si un dinosaure était à deux mètres de moi, je l'observerais » pour bien saisir la réponse de van Fraassen. Dans ce cas, il faudrait prendre en compte les données scientifiques dont on dispose à propos des dinosaures. Par exemple, à partir de leurs ossements fossilisés et des données dont nous disposons à propos de la biologie évolutive, il est possible d'établir de manière scientifique que plusieurs dinosaures avaient une taille comparable aux gros reptiles actuels, et qu'ils étaient souvent plus gros encore. Cela, mis en conjonction avec les données scientifiques que nous avons à propos de la vision de l'être humain, permet de conclure de manière factuelle que nous serions en mesure d'observer un dinosaure s'il était à une distance de deux mètres de nous.

3.5 Synthèse critique

L'empirisme constructif a été développé par van Fraassen dans le but d'offrir une alternative viable au réalisme scientifique qui s'accordait mieux avec la philosophie empiriste, selon laquelle seule l'expérience est une source de connaissance. En effet, alors que le réalisme scientifique stipule que la science a pour but de décrire le monde de manière littéralement vraie, et que l'acceptation d'une théorie scientifique implique la croyance en sa vérité, l'empirisme constructif, bien qu'il considère aussi que les théories scientifiques doivent être interprétées de manière littérale, est plus modeste du point de vue épistémique. Selon l'empirisme constructif, la science a comme objectif de fournir des théories empiriquement adéquates, et l'acceptation d'une théorie scientifique n'implique que la croyance qu'elle est empiriquement adéquate. Or, une théorie est empiriquement adéquate, en résumé, si elle est vraie à propos des phénomènes observables. Autrement dit, selon les empiristes constructifs, l'acceptation d'une théorie scientifique n'implique pas de croire en la vérité de ce qu'elle affirme à propos des entités théoriques inobservables, comme les virus ou les électrons. Par ailleurs, il est aussi important de rappeler que l'acceptation d'une théorie scientifique, pour van Fraassen, comporte une dimension pragmatique importante, en ce sens qu'au-delà de sa valeur épistémique (la croyance en son adéquation empirique), elle implique, par exemple, un engagement dans un programme de recherche et une certaine confiance que les phénomènes qui seront observés dans le futur pourront être expliqués par cette théorie.

Par ailleurs, la notion d'observabilité est au cœur de celle d'adéquation empirique et est donc, par le fait même, absolument centrale dans l'empirisme constructif. Sans définir en bonne et due forme l'observabilité, van Fraassen considère qu'un objet X est observable s'il est possible pour nous de l'observer dans certaines circonstances. Il faut préciser que l'observabilité est toujours relative à une

certaine communauté épistémique spécifiée (dans le cas qui nous intéresse, celle formée par les êtres humains). De plus, bien que ce qui est observable ne dépende pas de la théorie scientifique, c'est cette dernière qui pourra délimiter et décrire le domaine de l'observable. La question de savoir ce qui est observable est donc, en fin de compte, une question empirique qui doit être étudiée par la science, car elle dépend de faits du monde ainsi que de faits à propos de nous en tant que nous sommes des machines perceptuelles.

Vu l'importance épistémique de la notion d'observabilité pour l'empirisme constructif, ainsi que la tension qui découle de la manière dont van Fraassen la développe (l'adéquation empirique repose sur la notion d'observabilité, mais le domaine de l'observable est délimité par la science elle-même), il n'est pas étonnant que ce soit à ce sujet que de nombreux auteurs aient décidé de critiquer van Fraassen et son empirisme constructif. C'est le cas, par exemple, de Ladyman, qui affirme que certains énoncés contrefactuels découlant de la notion d'observabilité telle que mise de l'avant par van Fraassen impliquent la croyance en un certain réalisme modal, qui semble inconsistant avec la thèse principale de l'empirisme constructif. Toutefois, Monton et van Fraassen rejettent cette objection en soutenant que les énoncés contrefactuels en question doivent être compris de manière factuelle et non modale. Bref, l'empirisme constructif est certainement l'une des versions de l'antiréalisme scientifique les plus prometteuses pour qui veut adhérer à une philosophie empiriste, mais elle comporte certaines difficultés considérables, notamment concernant la notion d'observabilité, qui occupent encore aujourd'hui une place importante dans les débats épistémologiques. Il aura sans doute été constaté que rien n'a été dit, dans le présent chapitre, à propos du modèle de l'explication scientifique au sein de l'empirisme constructif. En effet, le dernier chapitre du présent travail sera consacré à ce sujet, étant donné que la théorie de l'explication scientifique défendue par van Fraassen a suscité de nombreux débats en philosophie des sciences dans les dernières années, et que cette question mérite donc d'être étudiée plus en détail.

4. La théorie pragmatique de l'explication de Bas van Fraassen

4.1 Introduction

L'une des sections les plus influentes de l'ouvrage majeur de van Fraassen dont il a été question dans le chapitre précédent, *The Scientific Image*, est le cinquième chapitre dans lequel il pose les fondements d'une théorie pragmatique de l'explication scientifique, qui s'oppose en quelque sorte aux conceptions de l'explication scientifique que nous avons abordées jusqu'à maintenant. En effet, bien que Hempel, Kitcher ou Salmon ne nient absolument pas que l'explication scientifique comporte certaines dimensions pragmatiques, ces auteurs partagent tous une intuition assez similaire, à savoir qu'une théorie de l'explication scientifique devrait concerner d'abord et avant tout le cœur de la notion d'explication scientifique, celui-ci n'étant pas de nature pragmatique. D'ailleurs, nous avons vu dans le premier chapitre que Hempel rejetait plusieurs critiques ayant été adressées à son modèle justement parce qu'elles concernaient des aspects pragmatiques de l'explication, que son modèle ne tenait pas en compte.

Bien que nous verrons plus précisément le sens de ce terme « pragmatique », l'idée générale des auteurs dont il a été question jusqu'à présent est ainsi que l'aspect central de la notion d'explication scientifique peut être (et devrait être) compris comme étant indépendant des caractéristiques psychologiques des individus concernés par l'explication et des divers contextes dans lesquels sont fournies ces explications. Autrement dit, Hempel, Kitcher et Salmon veulent identifier et théoriser les caractéristiques de l'explication scientifique de manière formelle et suffisamment abstraite de sorte qu'on puisse comprendre la notion d'explication scientifique indépendamment des contextes particuliers dans lesquels émergent les occurrences d'explication scientifique.

Or, c'est justement à cette intuition centrale que s'opposent les défenseurs d'une théorie pragmatique de l'explication comme van Fraassen. Pour eux, non seulement la notion d'explication scientifique comporte un aspect pragmatique absolument essentiel qui devrait ainsi occuper une place centrale dans toute théorie de l'explication scientifique, mais les difficultés pour ainsi dire insolubles auxquelles font face les conceptions de l'explication scientifique traditionnelles sont justement dues au fait qu'elles ont évacué cette dimension pragmatique. Ainsi, selon les tenants des théories pragmatiques de l'explication, les théories de l'explication scientifique sont vouées à l'échec tant qu'elles ne rendront pas compte de cet aspect pragmatique fondamental de la notion d'explication.

Dans un premier temps, je préciserai ce qu'il faut comprendre par l'idée d'une théorie pragmatique de l'explication et je présenterai certaines caractéristiques générales de ce type de théorie. Par la suite, nous nous pencherons plus spécifiquement sur la théorie pragmatique de l'explication telle que proposée par van Fraassen, qu'il présente notamment comme étant une théorie des *questions-pourquoi* (*why-questions*). Finalement, j'essaierai de montrer en quoi on retrouve l'esprit de l'empirisme constructif dans cette théorie pragmatique de l'explication, et donc que la théorie de l'explication scientifique présentée par van Fraassen, bien que pouvant être comprise et défendue indépendamment de l'empirisme constructif, en découle tout de même assez naturellement.

4.2 Considérations générales à propos des théories pragmatiques de l'explication

Comme le terme « pragmatique » est polysémique, il est avant tout primordial de préciser sa signification pertinente dans le cadre de l'étude des théories pragmatiques de l'explication. D'abord, il faut comprendre qu'il n'est pas question de faire référence à la tradition philosophique du pragmatisme telle qu'incarnée notamment par Charles Sanders Peirce, John Dewey et William James, selon laquelle, très grossièrement, les problèmes et les concepts philosophiques doivent être abordés et étudiés du point de vue de leurs conséquences pratiques dans le monde. D'ailleurs, l'emploi du terme « pragmatique » dans le langage courant n'est pas sans se rapprocher de cet usage philosophique, en ce sens qu'il concerne le succès ou la réussite dans l'action, en opposition avec la valeur théorique ou idéologique. Woodward (2014) remarque qu'il est intuitif de comprendre le mot « pragmatique » à peu près de cette manière: « On one natural understanding of this notion, a pragmatic consideration is one that has to do with utility or usefulness in the service of some goal connected to human interests, where these interests are in some relevant sense “practical” ». Or, ce n'est pas non plus à cette signification qu'il faut nous attarder dans notre enquête.

En effet, pris en ce sens, on pourrait qualifier plusieurs théories de l'explication traditionnelles d'au moins partiellement pragmatiques, notamment le modèle déductif-nomologique de Hempel. De fait, nous avons vu dans le premier chapitre que Hempel liait de manière assez forte les notions d'explication et de compréhension scientifiques avec celle de prévision. La prévision scientifique, bien qu'elle puisse concerner d'abord et avant tout le développement ou la confirmation de théories ou la vérification d'énoncés, est tout de même une notion qui a une valeur pragmatique évidente, en ce sens que c'est à partir de prévisions, entre autres, qu'il devient possible de développer des applications utiles aux découvertes scientifiques. Bref, cette signification du terme « pragmatique », liée à l'utilité et aux conséquences pratiques, est tout à fait compatible avec l'esprit des théories traditionnelles de l'explication scientifique.

Lorsqu'il est question de théories pragmatiques de l'explication, il s'agit d'insister sur l'idée que les occurrences d'explications sont toujours fortement dépendantes du contexte précis dans lequel elles sont données. Par ailleurs, les considérations psychologiques liées à celui ou celle qui explique et à celui ou celle qui reçoit l'explication (par exemple, les croyances des individus, leurs intérêts, etc.) constituent une part importante de ce contexte d'explication (Woodward, *Ibid.*). Bien évidemment, à première vue, cela peut ne pas paraître incompatible avec les conceptions traditionnelles de l'explication, car les défenseurs de ces théories peuvent tout à fait accepter que des facteurs pragmatiques aient une influence sur les occurrences particulières d'explications. Comme il a été mentionné dans l'introduction, Hempel, par exemple, reconnaît explicitement l'existence de considérations pragmatiques pour rejeter certaines critiques lui ayant été dirigées. Cependant, les défenseurs des théories traditionnelles, bien qu'ils reconnaissent l'existence de facteurs pragmatiques dans l'interaction qui consiste à fournir et à recevoir une explication, pensent que ces facteurs ne font pas partie des constituants essentiels ou fondamentaux de la notion d'explication. C'est précisément cette idée que rejettent les défenseurs des théories pragmatiques de l'explication.

Les tenants des théories traditionnelles de l'explication considèrent que la notion d'explication peut être définie et caractérisée indépendamment de ces facteurs pragmatiques, c'est-à-dire qu'ils ont comme objectif d'élaborer une notion d'explication suffisamment générale et formelle (par exemple, en termes de déduction logique, de relations statistiques, etc.) pour ne pas varier selon les contextes d'explication. Autrement dit, pour les défenseurs de ces approches, il est légitime de parler de « bonne explication » ou d'« explication correcte », en ce sens que l'explication en question respecte le modèle structurel de la théorie, peu importe les facteurs pragmatiques qui entrent en jeu. Au contraire, non seulement les défenseurs d'une théorie pragmatique de l'explication accordent une plus grande importance auxdits facteurs pragmatiques lorsque vient le temps de considérer les occurrences réelles d'explication, mais ils pensent que cette dimension pragmatique est constitutive de la notion théorique d'explication. Qu'est-ce à dire? Il s'agit de soutenir que cette dimension pragmatique de l'explication ne peut aucunement se réduire ou être intégrée dans une conception formelle et générale de l'explication telle que défendue dans les théories traditionnelles de l'explication. Ainsi, si les conceptions de l'explication traditionnelles ont toutes, jusqu'à maintenant, fait face à des difficultés pour ainsi dire insurmontables, c'est justement parce que cet aspect pragmatique situé au cœur de la notion d'explication a été ignoré, comme l'explique Woodward (2014):

[w]hat is distinctive about pragmatic approaches to explanation is not just the bare idea that explanation has a “pragmatic dimension” but rather the further and much stronger claim that that the traditional project of constructing a model of explanation

pursued by Hempel and others has so far been unsuccessful (and perhaps is *bound* to be unsuccessful) and that this is so *because* pragmatic or contextual factors play a central and ineliminable role in explanation in a way that resists incorporation into models of the traditional sort.

Bref, l'apport original des tenants des théories pragmatiques de l'explication n'est pas simplement d'affirmer l'existence de facteurs pragmatiques liés aux occurrences d'explication, mais plutôt de soutenir qu'il est nécessaire de théoriser cette dimension pragmatique de sorte qu'elle fasse elle-même partie intégrante de la notion d'explication. Je pense qu'il est possible de comprendre la distinction ainsi: alors que pour les partisans des théories traditionnelles de l'explication, les facteurs pragmatiques viennent se greffer aux explications lors de leur occurrence³⁶, les défenseurs des théories pragmatiques considèrent que ces facteurs pragmatiques doivent être intégrés dans la notion même d'explication. Dans cette dernière optique, ignorer la dimension pragmatique de l'explication reviendrait donc simplement à ne pas saisir correctement l'essentiel de la notion d'explication.

4.3 La théorie pragmatique de l'explication de van Fraassen: une théorie des questions-pourquoi

4.3.1 Mise en contexte : l'explication et les théories scientifiques

Avant de nous pencher sur la manière dont van Fraassen rend compte de la notion d'explication en élaborant une théorie des questions-pourquoi, je pense qu'il serait à propos de dire quelques mots concernant la place qu'il accorde à l'explication au sein de l'entreprise scientifique et aux liens que la notion d'explication entretient avec celle de théorie scientifique. Il a été mentionné dans le chapitre précédent que certaines vertus pragmatiques, selon van Fraassen, jouaient un rôle considérable eu égard à l'acceptation d'une théorie scientifique, et que cette dimension pragmatique de l'acceptation ne concerne pas la relation entre la théorie et le monde³⁷, mais plutôt l'utilité de ladite théorie indépendamment des considérations sur sa vérité ou son adéquation empirique.

C'est après avoir établi cette distinction entre les vertus pragmatiques et les vertus épistémiques (liées aux propriétés sémantiques) d'une théorie scientifique que van Fraassen aborde la question de savoir ce qui rend rationnelle la recherche d'explications en science (van Fraassen, 1980, p.88). Pour y répondre, il faut d'abord comprendre que les vertus épistémiques que doit posséder une théorie

³⁶ C'est-à-dire, au moment de la formulation d'une explication dans un contexte donné.

³⁷ Van Fraassen qualifie les propriétés qui concernent la relation entre la théorie et le monde de « sémantiques ». Les deux propriétés sémantiques principales sont celles de vérité et d'adéquation empirique. (van Fraassen, 1980, p.90).

scientifique pour être utilisée dans une explication ne sont pas particulières, c'est-à-dire qu'il s'agit exactement des vertus épistémiques dont il a été question dans le chapitre précédent, notamment celle d'adéquation empirique. Cependant, ces vertus épistémiques ne suffisent pas; il faut quelque chose de plus pour véritablement parler d'explication, et c'est ici que la dimension pragmatique devient cruciale, comme nous le constaterons plus bas. L'idée de van Fraassen est que la poursuite de l'explication en science, même si elle comporte elle-même un aspect pragmatique important, se rapporte, en fin de compte, à la poursuite des vertus épistémiques primaires des théories scientifiques. En effet, il a été souligné dans le chapitre précédent que selon l'empirisme constructif, l'objectif central de la science est de produire des théories empiriquement adéquates. La poursuite de l'explication va justement permettre la construction de théories scientifiques de plus en plus empiriquement adéquates, ce qui fait de l'explication scientifique le moyen par excellence de recherche des vertus épistémiques fondamentales qui constituent l'objectif de la science: « But in the pursuit of explanation we pursue *a fortiori* those more basic merits, which is what makes the pursuit of explanation of value to the scientific enterprise as such. [...] the pursuit of explanatory power is the best means to serve the central aims of science » (*Ibid.*, pp.88-89).

Pour illustrer cette affirmation quelque peu contre-intuitive, dans la mesure où, à première vue, on serait tenté de penser que le pouvoir explicatif d'une théorie scientifique est une qualité désirable en elle-même, van Fraassen mentionne certains exemples historiques de recherches d'explication qui ont effectivement conduit à des théories scientifiques empiriquement plus adéquates. C'est le cas, par exemple, de la recherche d'une théorie mécanique compatible avec le modèle astronomique copernicien qui a conduit à la théorie mécanique newtonienne ou encore de la recherche d'une explication du phénomène des spectres atomiques distincts qui poussa à la formation d'un nouveau modèle de la structure atomique (*Ibid.*, p.93). Ainsi, il y a toujours eu certaines raisons de douter de l'adéquation empirique des théories scientifiques existantes, et ce sont précisément ces raisons qui motivent les recherches de meilleures explications. Autrement dit, le pouvoir explicatif n'est pas une qualité irréductible des théories scientifiques, comme l'explique van Fraassen: « [t]he search for explanation is valued in science because it consists *for the most part* in the search for theories which are simpler, more unified, and more likely to be empirically adequate. This is not because explanatory power is a separate quality *sui generis* which, mysteriously, makes those other qualities more likely, but because having a good explanation *consists* for the most part in having a theory with those other qualities » (*Ibid.*, pp.93-94).

Cela appelle quelques réflexions supplémentaires concernant les rapports entre l'adéquation empirique ou la vérité des théories scientifiques et la capacité explicative de ces théories.

Effectivement, il est justifié de se demander en quoi la relation qu'une théorie scientifique entretient avec le monde peut être reliée à ses capacités explicatives. Ce qui vient d'être dit dans les derniers paragraphes pourrait insinuer qu'une théorie scientifique donnée ne peut être explicative que si elle est empiriquement adéquate, étant donné que selon van Fraassen, *avoir* une bonne explication d'un phénomène *T* revient en grande partie à posséder une théorie empiriquement adéquate à propos du phénomène *T*. Or, les propos de van Fraassen, à cet égard, peuvent sembler nébuleux, voire contradictoires, en ce sens qu'il affirme à plusieurs reprises qu'il n'est aucunement requis qu'une théorie donnée soit empiriquement adéquate pour affirmer que cette même théorie a un pouvoir explicatif. Encore une fois, on peut se référer à de multiples exemples historiques pour mettre l'accent sur cette idée: il n'est en rien controversé d'affirmer que la théorie mécaniste de Newton expliquait efficacement le phénomène des marées ou le mouvement des planètes, ni que la théorie de Huygens expliquait clairement le phénomène de la diffraction lumineuse. Pourtant, aujourd'hui, on sait très bien que ces théories ne sont pas empiriquement adéquates. Comment donc expliquer cette tension entre les propos de van Fraassen? Quelle est donc, finalement, la relation entre l'adéquation empirique d'une théorie scientifique et son pouvoir explicatif?

La clef de la réponse à ces interrogations se trouve dans une distinction importante, mais subtile, que fait van Fraassen entre l'affirmation « nous *avons* une explication » et « cette théorie explique ce phénomène ». Dans le premier cas, lorsqu'il est question pour un individu quelconque d'avoir une explication, cela implique que la théorie scientifique utilisée par l'individu pour expliquer le phénomène soit acceptable, notamment en ce qui concerne son adéquation empirique. Ainsi, dans un tel cas, l'adéquation empirique est en quelque sorte une condition à l'explication: nous ne disons pas que nous avons une explication de tel phénomène si nous n'avons pas une théorie acceptable quant à son adéquation empirique à laquelle nous référer pour expliquer (*Ibid.*, p.95). Ce n'est toutefois pas à ce premier sens que s'intéresse principalement van Fraassen.

De fait, cela est différent pour le deuxième cas; la simple affirmation selon laquelle une théorie donnée explique un phénomène n'implique en rien l'acceptabilité générale de cette théorie: « The important point is that the mere statement "Theory T explains fact E" does not carry any such implication: not that the theory is true, not that it is empirically adequate, and not that it is acceptable » (*Ibid.*, p.98). Il est cependant important de noter que cette formulation est imprécise, car une théorie, en elle-même, ne peut jamais expliquer quelque phénomène que ce soit, comme nous le comprendrons un peu plus bas. Il serait plus juste d'affirmer qu'il est possible d'utiliser une théorie pour expliquer un phénomène sans pour autant que cette théorie ne soit empiriquement adéquate dans son ensemble. Le présent point est de montrer qu'il est possible d'utiliser une théorie

pour expliquer un phénomène particulier sans que l'on puisse dire légitimement que *nous avons* une explication: c'est le cas si cette théorie n'est pas acceptable. Par exemple, on peut utiliser la théorie de Newton pour expliquer les marées, mais il ne serait pas justifiable d'affirmer que nous avons alors une explication des marées, car nous savons aujourd'hui que cette théorie n'est pas empiriquement adéquate, et donc pas acceptable.

Ainsi, en résumé, ce qui a été dit plus haut, à savoir que les vertus épistémiques qui donnent à une théorie scientifique un pouvoir explicatif considérable sont précisément les mêmes qui constituent la visée de la science, est valide dans le contexte où l'on affirme que l'on possède une explication d'un phénomène quelconque. Les explications sont donc recherchées dans le but de créer des théories scientifiques les plus acceptables possible. Cependant, si l'on ne s'intéresse qu'à la relation entre la théorie scientifique elle-même et le pouvoir explicatif, l'adéquation empirique globale de la théorie n'est en rien conditionnelle à sa capacité d'expliquer tel ou tel phénomène. En fait, la relation d'explication, en tant que telle, tient entre différents faits et une théorie, et cette relation est tout à fait indépendante de l'adéquation empirique de cette théorie, ce qui implique par ailleurs que la relation d'explication est visible ou identifiable avant même de nous pencher sur la question de l'adéquation empirique de ladite théorie: « Hence, to say that a theory explains some fact or other, is to assert a relationship between this theory and that fact, which is independent of the question whether the real world, as a whole, fits that theory » (*Ibid.*, p.98). C'est à cette relation d'explication entre la théorie scientifique elle-même et les phénomènes à expliquer que van Fraassen s'intéresse principalement dans sa théorie pragmatique de l'explication.

À cet égard, il est crucial de préciser que cette relation d'explication entre la théorie scientifique et les faits n'est pas une relation binaire entre ces deux éléments. Toute la théorie des questions-pourquoi que développe van Fraassen, comme nous le verrons plus bas, consiste justement à montrer qu'il existe un troisième terme absolument déterminant dans la relation d'explication, à savoir les données contextuelles qui entourent l'occurrence de l'explication. Une simple relation entre une théorie scientifique et un fait ou un phénomène est au plus une relation descriptive, et c'est justement l'erreur des partisans des modèles traditionnels de l'explication que d'avoir confondu cette relation descriptive avec une véritable relation explicative. Toutefois, la similarité entre ces deux relations permet de comprendre plus clairement en quoi l'adéquation empirique générale d'une théorie n'est pas requise afin que celle-ci soit utilisée pour expliquer un phénomène.

En effet, décrire un phénomène avec une théorie scientifique implique que la théorie soit empiriquement adéquate par rapport à ce phénomène, étant donné qu'une théorie est empiriquement

adéquate lorsqu'un phénomène peut être intégré dans l'une des substructures empiriques de la théorie, comme nous l'avons vu au chapitre précédent. Décrire un phénomène, c'est donc montrer qu'il s'intègre dans l'une de ces substructures, d'où l'idée que la relation de description est bel et bien une relation binaire entre la théorie et le phénomène. Mais il est possible de décrire un phénomène avec une théorie sans pour autant qu'elle soit empiriquement adéquate de manière globale, c'est-à-dire sans que tous les phénomènes qui appartiennent à son champ d'études puissent être reconnus dans l'une ou l'autre de ses substructures empiriques. Il en est de même pour la relation d'explication: expliquer un phénomène avec une théorie implique bien sûr que le phénomène puisse être décrit par la théorie, et donc intégré dans l'une de ses substructures empiriques, mais l'adéquation empirique générale de la théorie n'est nullement nécessaire. C'est ce que soulignait van Fraassen dans la citation ci-haut: la question de savoir si la théorie permet de rendre compte de tous les phénomènes du monde qui appartiennent à son champ d'études est indépendante de la question de savoir si la théorie peut être utilisée pour expliquer un phénomène particulier.

4.3.2 Importance du contexte dans les relations causales et les questions-pourquoi

Van Fraassen, pour élaborer sa théorie de l'explication, commence par faire un bref retour sur différentes théories qui ont été formulées au vingtième siècle, notamment celles de Hempel et de Salmon. L'une des conclusions à laquelle arrive van Fraassen est que la notion de causalité semble intimement liée à celle d'explication, et que Salmon l'a utilisée entre autres pour résoudre le problème de l'asymétrie de l'explication dont il a été question dans les deux premiers chapitres. Or, bien que Salmon élabore toute une théorie de la causalité centrée autour de la notion de processus, van Fraassen remarque que dans le langage ordinaire, la relation de causalité est souvent exprimée à l'aide d'énoncés conditionnels contrefactuels. D'ailleurs, l'une des théories les plus influentes de la causalité, formulée par David Lewis, définit la causalité en termes de conditionnels contrefactuels (van Fraassen, 1980, p.114). Prenons l'exemple suivant: « Si la plante n'avait pas été aspergée de poison, elle ne serait pas morte ». Dans cet exemple, il est tout à fait intuitif de dire que le fait d'avoir aspergé la plante de poison est la cause de sa mort. Cependant, van Fraassen note que ce conditionnel contrefactuel implique une clause *ceteris paribus* implicite qui est tout à fait importante, qui se traduit explicitement comme suit: « Si la plante n'avait pas été aspergée de poison (*et que tout le reste était resté identique*), elle ne serait pas morte » (*Ibid.*, p.115). En effet, si la plante n'avait pas été aspergée de poison, mais qu'elle avait été placée dans un congélateur, elle serait tout de même morte. La clause *ceteris paribus* est donc tout à fait essentielle à la vérité de ce conditionnel contrefactuel.

Le point de van Fraassen, c'est qu'il est impossible de déterminer le contenu exact de cette clause, car elle dépend du contexte dans lequel le conditionnel est énoncé: « There is therefore a contextual variable – determining the content of that tacit *ceteris paribus* clause – which is crucial to the truth-value of the conditional statement » (*Ibid.*, p.116). C'est en jumelant cette remarque avec la théorie de l'explication de Salmon (qui, rappelons-nous-en, est centrée autour de la notion de toile causale) que van Fraassen peut introduire peu à peu sa propre théorie pragmatique de l'explication. En effet, Salmon soutient qu'expliquer scientifiquement un événement, c'est le situer dans la toile causale pour ainsi identifier les processus qui l'ont précédé. Mais van Fraassen précise qu'il faut parler des facteurs pertinents (*salient factors*), déterminés en partie par la clause *ceteris paribus* contextuelle, qui ont mené à l'événement que l'on cherche à expliquer. Ce sont ces facteurs pertinents que l'on identifiera comme les causes de l'événement (*Ibid.*, p.124). Les facteurs en question sont pertinents pour une certaine personne en raison des données contextuelles, par exemple les intérêts ou les connaissances préalables de cette personne. Van Fraassen résume ainsi son idée: « To sum up: no factor is explanatory relevant unless it is scientifically relevant; and among the scientifically relevant factors, context determines explanatory relevant ones » (*Ibid.*, p.126). Autrement dit, c'est le contexte qui va déterminer sur quelle partie ou portion de la toile causale notre regard sera porté pour expliquer l'occurrence d'un événement particulier. C'est cette intuition qui sera développée dans la théorie des questions-pourquoi que propose van Fraassen.

Avant d'enfin plonger dans cette théorie comme telle, une dernière (mais brève) incursion s'impose. Effectivement, van Fraassen s'inspire de Sylvain Bromberger³⁸ dans son analyse des questions-pourquoi, en tant que ce type de question est en fait une demande d'explication, et il importe d'avoir en tête ce qu'il en retire précisément pour élaborer sa propre théorie. Une question-pourquoi, par exemple « Pourquoi l'eau bout-elle à 100 degrés Celsius? », prend normalement la forme « Pourquoi (est-ce le cas que) *P*? », où *P* est un énoncé quelconque. La première chose que constate Bromberger est que cette question ne peut survenir de manière sensée que si l'eau bout bel et bien à 100 degrés Celsius. Ainsi, pour toute question de type « Pourquoi *P*? », on dira que *P* est la présupposition de cette question (*Ibid.*, p.127). Van Fraassen, cependant, remarque que les choses ne sont pas aussi simples. Prenons la question « Pourquoi Basile porte-t-il un chapeau? ». Cette question peut être reformulée de plusieurs façons différentes, par exemple « Pourquoi est-ce Basile plutôt qu'une autre personne qui porte un chapeau? » ou « Pourquoi est-ce un chapeau que Basile porte? », ces reformulations représentant des demandes d'explication distinctes les unes des autres.

³⁸ Bromberger (1966).

En effet, alors qu'on pourrait répondre à la question qu'est la première reformulation par « Parce que seul Basile est chauve et il ne veut pas avoir froid », cette réponse ne serait pas adéquate pour la question de la deuxième reformulation, qui demanderait plutôt une réponse du type « Parce que Basile n'aime pas les autres sortes de couvre-tête ». En effet, la différence principale entre ces deux reformulations est qu'elles ne pointent pas vers les mêmes alternatives contrastantes (*Ibid.*); dans le premier cas, il s'agit de demander pourquoi est-ce Basile, *plutôt qu'une autre personne*, qui porte un chapeau, alors que dans le deuxième cas, il s'agit de demander pourquoi Basile porte un chapeau *plutôt qu'un autre couvre-tête*. C'est pourquoi van Fraassen pense qu'il est crucial de préciser la formulation de la structure des questions-pourquoi proposée par Bromberger comme suit: « Pourquoi (est-ce le cas que) P par opposition à (d'autres membres de) X ? », où X est la classe contrastante qui contient plusieurs alternatives. Une réponse adéquate à une question de ce type en sera donc une qui apportera de l'information qui favorise P par rapport aux autres alternatives contenues dans X . Par ailleurs, il faut noter que dans la majorité des situations, la classe contrastante n'est pas explicitement mentionnée dans la question, car elle est bien souvent implicite selon le contexte de la question (*Ibid.*, p.128).

Résumons ce qui a été dit jusqu'à maintenant dans cette sous-section. Il a été question de deux aspects de l'explication dans lesquels le contexte est primordial. Premièrement, la relation de causalité (en tant qu'elle est constituante de l'explication), selon van Fraassen, ne peut être comprise indépendamment du contexte, car c'est selon le contexte qu'il sera possible de déterminer sur quelle portion de la toile causale il faudra porter notre regard pour expliquer un événement donné. Bref, c'est le contexte qui va permettre de déterminer la pertinence des différents processus constituant la toile causale, et qui va donc proprement permettre de déterminer ce que sont les causes d'un événement (qui ne seront donc pas les mêmes selon le contexte): « But salience is context-dependent, and the selection of the correct 'important' factors depends on the range of alternatives contemplated in that context » (*Ibid.*, p.129). Deuxièmement, le contexte va permettre d'identifier la classe contrastante X qui contient des alternatives différentes à l'événement P que l'on cherche à expliquer. C'est par opposition à ces alternatives que l'explication de P sera convaincante. Par ailleurs, on pourrait penser que le fait de requérir une explication *scientifique* d'un événement détermine par le fait même la pertinence des différents facteurs causaux ainsi que la classe contrastante, mais il n'en est rien. Le fait de demander une explication scientifique implique que l'explication soit fondée sur une théorie scientifique et sur des données expérimentales rigoureuses (*Ibid.*), mais le contexte demeure crucial pour la détermination de la pertinence des facteurs et de la classe contrastante, comme l'explique van Fraassen: « The description of some account as an

explanation of a given fact or event, is incomplete. It can only be an explanation with respect to a certain *relevance relation* and a certain *contrast-class*. These are contextual factors, in that they are determined neither by the totality of accepted scientific theories, nor by the event or fact for which an explanation is requested » (*Ibid.*, p.130).

4.3.3 L'explication selon van Fraassen: une théorie des questions-pourquoi

Nous détenons maintenant tous les outils nécessaires pour comprendre la théorie de l'explication que propose van Fraassen. Nous nous en doutons: pour lui, une explication, c'est une réponse à une question-pourquoi, et une théorie de l'explication en sera donc une qui portera sur la compréhension de ce type de question et de la structure de la réponse qui peut lui être donnée. Reprenons un exemple de question-pourquoi, « Pourquoi le sang circule-t-il dans les vaisseaux sanguins? ». Van Fraassen identifiera comme le thème (*topic*) de cette question la proposition selon laquelle le sang circule dans les vaisseaux sanguins. Ce thème est ainsi exactement la présupposition centrale de la question, si l'on reprend la terminologie de Bromberger abordée plus tôt. Par ailleurs, la classe contrastante X impliquée par cette question sera constituée par différentes alternatives, qui sont en fait différentes propositions impliquant le thème de la question. La classe contrastante pourrait être constituée, par exemple, par des propositions comme « le *sang* circule dans les vaisseaux sanguins », « de l'*eau pure* circule dans les vaisseaux sanguins », « la *lymphe* circule dans les vaisseaux sanguins », etc. En dernier lieu, il faut prendre en considération « the respect-in-which a reason is requested » (*Ibid.*, p.142), qui permet en fait d'identifier le champ des facteurs explicatifs possibles, c'est-à-dire la relation de pertinence explicative. Cette dernière idée se rapproche bien sûr de ce dont il a été dit ci-haut à propos de la pertinence concernant les relations causales. Dans notre même exemple, la demande d'explication pourrait avoir trait aux fonctions biologiques remplies par le sang, à l'histoire évolutive du processus de circulation sanguine, etc.

Ainsi, plusieurs questions-pourquoi peuvent porter sur exactement le même thème, mais impliquer des classes contrastantes distinctes ou des relations de pertinence explicative bien différentes. Il ne serait donc pas adéquat d'affirmer que telle proposition est pertinente en regard de ce thème, ou encore que telle proposition est pertinente concernant telle classe contrastante. En effet, il faudrait plutôt affirmer que telle proposition est pertinente, dans un contexte donné, à propos de ce thème en considérant une certaine classe contrastante. En résumé, une question-pourquoi Q exprimée dans un contexte donné dépendra de trois facteurs: le thème P_k de la question, la classe contrastante $X = \{P_1, \dots, P_k, \dots\}$ et la relation de pertinence R (*Ibid.*, p.143). Ainsi, une proposition A pourra être dite pertinente par rapport à Q si elle entretient la relation R par rapport au couple (P_k, X) , c'est-à-dire si elle permet de sélectionner P_k parmi tous les autres membres de X , étant donné la relation R

impliquée par la question. Van Fraassen dira alors que la proposition A est une *raison*. Bref, formellement, une réponse à la question Q sera de type « P_k (contrairement à tous les autres membres de X) à cause de A ». Cette réponse implique plusieurs choses: que P_k est vrai, que les autres membres de X sont faux, que A est vraie et que A est une raison, donc qu'elle est pertinente par rapport à Q .

Prenons l'exemple de quelqu'un qui demanderait pourquoi telle personne s'est fracturé le genou. Hors contexte, cette question est très vague. Dans un certain contexte, la personne qui pose la question pourrait s'intéresser aux événements qui ont mené à la fracture du genou, par exemple à la chute lors d'une descente en skis. Dans ce cas, la question aurait comme thème « telle personne s'est fracturé le genou », la classe contrastante serait constituée par une multitude d'événements (boire un café, embarquer dans la voiture, etc.) et la relation R serait une relation concernant la chaîne d'événements ayant mené à la fracture. Une réponse adéquate³⁹ pourrait donc « Elle s'est fracturé le genou parce qu'elle est tombée lors d'une descente en skis », dans laquelle la proposition A est « elle est tombée lors d'une descente en ski ». On voit bien que dans ce cas, A est bel et bien une raison, car elle montre en quoi P_k , par rapport aux autres membres de la classe contrastante, s'accorde avec la relation R impliquée par la question. Cette relation aurait pu être autre: quelqu'un aurait pu demander pourquoi cette même personne s'est fracturé le genou en s'intéressant à l'aspect physiologique de la fracture. La réponse, dans une telle situation, aurait pu prendre l'allure de « Elle s'est fracturé le genou parce que les tissus ligamentaires et cartilagineux autour de sa rotule n'étaient pas suffisamment forts pour maintenir l'articulation en place ». On voit bien que A , ici, est complètement d'un autre genre, c'est-à-dire qu'elle n'est pas du tout la même sorte de raison que dans la réponse à la première version de la question.

Il est maintenant possible de comprendre la caractérisation formelle que van Fraassen propose pour comprendre ce qu'est une réponse à une question-pourquoi: « B is a *direct answer*⁴⁰ to question $Q = \langle P_k, X, R \rangle$ exactly if there is some proposition A such that A bears relation R to $\langle P_k, X \rangle$ and B is the proposition which is true exactly if (P_k ; and for all $i \neq k$, not P_i ; and A) is true where, as before, $X = \{P_1, \dots, P_k, \dots\}$ » (*Ibid.*, p.144). En outre, van Fraassen remarque qu'on peut souvent répondre à une question-pourquoi uniquement par « Parce que A », ce pour quoi il dira que A est le cœur (*core*) de la réponse B , alors que « P_k ; and for all $i \neq k$, not P_i » sera la présupposition centrale de la question

³⁹ Nous reviendrons dans quelques lignes sur l'évaluation des réponses aux questions-pourquoi.

⁴⁰ Van Fraassen considère qu'une réponse directe à une question est une réponse qui fournit assez d'information pour répondre complètement à la question posée, mais pas plus (1980, p.138).

Q (*Ibid.*, p.145). Van Fraassen donne aussi quelques pistes de réflexion concernant l'évaluation des réponses aux questions-pourquoi. En effet, une réponse peut être évaluée sous trois aspects différents. Premièrement, il faut bien sûr s'assurer que la proposition *A* est vraie ou acceptable. Deuxièmement, il faut aussi examiner si *A* favorise bel et bien le thème P_k par rapport aux autres membres de la classe contrastante⁴¹. Troisièmement, il faut comparer la réponse partielle « Parce que *A* » aux autres réponses qui auraient pu être possibles. Ce dernier aspect implique trois choses: il faut d'abord évaluer la probabilité de *A* par rapport aux autres raisons possibles. Il faut aussi voir en quoi *A*, même si elle favorise le thème P_k , le favorise plus ou moins par rapport aux autres raisons qui auraient pu être évoquées. Finalement, il est possible que la réponse donnée à une question puisse devenir plus ou moins caduque lorsqu'on la compare à d'autres réponses plus complètes ou plus précises (*Ibid.*, pp.146-147).

Il est alors bien aisé de comprendre en quoi la théorie de l'explication de van Fraassen est une théorie pragmatique de l'explication. En effet, la validité de l'explication, en tant qu'elle est une réponse à une question-pourquoi, dépend très fortement du contexte dans lequel la question est posée, et ce, sous deux rapports. Premièrement, c'est le contexte qui permet de déterminer le contenu de la classe contrastante dans laquelle se retrouve le thème de la question P_k . Deuxièmement, la relation *R* dépend elle aussi fortement du contexte, notamment en ce qui concerne la nature de l'information qui est recherchée par celui ou celle qui pose la question. Le contexte déterminant la question, et donc la réponse à cette question, de deux manières certes liées, mais tout de même distinctes, on peut conclure qu'une explication, selon la théorie de van Fraassen, est intrinsèquement rattachée aux données contextuelles qui l'entourent. Dans cette optique, l'explication est donc une notion fondamentalement pragmatique, qui ne se résume aucunement à une relation entre une théorie et un fait ou un phénomène; le contexte est absolument essentiel pour comprendre ce qu'est l'explication, comme l'explique van Fraassen pour conclure ses réflexions dans son cinquième chapitre de *The Scientific Image*:

The discussion of explanation went wrong at the very beginning when explanation was conceived of as a relationship like description: a relation between theory and fact. Really it is a three-term relation, between theory, fact, and context. [...] Being an explanation is essentially relative, for an explanation is an *answer*. [...] Since an explanation is an answer, it is evaluated *vis-à-vis* a question, which is a request for information. But exactly what is requested, by means of the interrogative 'Why is it the case that *P*?', differs from context to context. In addition, the background theory plus

⁴¹ Ici, on pourrait penser que la notion de pertinence statistique de Salmon, abordée dans le deuxième chapitre, est à propos.

data relative to which the question is evaluated, as arising or not arising, depends on the context. (*Ibid.*, p.156)

Par ailleurs, il a été mentionné très rapidement, plus haut, qu'une explication scientifique n'est en fait qu'une explication fondée sur les théories scientifiques. En effet, pour van Fraassen, rien ne distingue l'explication scientifique des autres explications du point de vue de sa structure: « [van Fraassen] explicitly denies that there is anything distinctive about the category of scientific explanation that has to do with its structure or form—instead, a scientific explanation is simply an explanation that makes use of information that is (or at least that is treated as) grounded in a “scientific” theory » (Woodward, 2014). Bref, toute explication, qu'elle soit scientifique ou non, a la même structure en tant qu'elle est une réponse à une question-pourquoi, ce qui implique que l'aspect pragmatique est donc constituant de la notion d'explication.

4.4 La théorie de l'explication de van Fraassen au sein de l'empirisme constructif

Pour conclure ce chapitre, j'aimerais soutenir la thèse que même si la théorie pragmatique de l'explication de van Fraassen peut certainement se comprendre et se défendre de manière relativement indépendante de l'empirisme constructif, il existe des liens conceptuels manifestes entre les deux théories, de sorte que cette théorie de l'explication semble s'inscrire naturellement dans la théorie de l'empirisme constructif. En premier lieu, si je pense que la théorie de l'explication proposée par van Fraassen peut être comprise et défendue de manière assez indépendante de l'empirisme constructif, c'est parce que cette théorie de l'explication n'est pas fondée sur les notions centrales de l'empirisme constructif. En effet, la théorie de l'explication comme théorie des réponses aux questions-pourquoi peut tout à fait se comprendre indépendamment des notions d'adéquation empirique ou d'acceptabilité des théories scientifiques, qui sont, à mon avis, les deux notions les plus importantes de l'empirisme constructif. De fait, les développements théoriques à propos de la structure des questions-pourquoi et de leurs réponses ne reposent en rien sur ces notions. Par ailleurs, la manière dont van Fraassen montre que le contexte est absolument fondamental dans la détermination de ces questions-pourquoi (et donc, bien entendu, de leurs réponses) ne fait aucunement référence au cadre théorique de l'empirisme constructif.

Cependant, les choses sont quelque peu différentes si l'on se penche sur les réflexions d'ordre plus général que van Fraassen élabore à propos de l'explication scientifique, notamment en ce qui concerne les rapports entre la quête de l'explication et les théories scientifiques. Effectivement, il a été expliqué, dans la sous-section 4.3.1, en quoi van Fraassen pensait que la recherche d'explications est le moteur principal de la formation de théories scientifiques de plus en plus empiriquement

adéquates. De plus, comme il a été dit, avoir une explication scientifique valable, selon van Fraassen, implique la possession d'une théorie scientifique acceptable sur laquelle fonder cette explication. Ainsi, de ce point de vue plus général, une compréhension au moins partielle de l'empirisme constructif est nécessaire pour saisir les liens entre les notions d'explication et de théorie scientifique. Mais si notre attention ne porte que sur la théorie pragmatique de l'explication en tant que telle, c'est-à-dire sur la définition formelle de l'explication, il me semble qu'elle peut être correctement comprise sans connaissances préalables de la théorie de l'empirisme constructif.

Il n'en reste pas moins qu'il existe des liaisons conceptuelles serrées entre la théorie de l'empirisme constructif et la théorie pragmatique de l'explication telles que présentées par van Fraassen. D'abord, en effet, le cadre général déterminé par l'empirisme constructif défendu par van Fraassen limite les possibilités des conceptions de l'explication scientifique admissibles. L'empirisme constructif est une version de l'antiréalisme scientifique, c'est-à-dire une théorie de la science selon laquelle la science n'a pas pour objectif de décrire le monde tel qu'il est fondamentalement, autrement dit de décrire le monde de manière véridique dans toutes ses parties (rappelons-nous que la science, selon van Fraassen, ne concerne que la vérité à propos de l'observable). Nous avons vu que dans le cas précis de l'empirisme constructif, il s'agit de dire que la science a pour objectif de produire des théories empiriquement adéquates. Or, cet antiréalisme est incompatible avec bon nombre de théories de l'explication scientifique. D'abord, c'est le cas avec le modèle déductif-nomologique de Hempel (et ses autres versions similaires), dans lequel la notion de loi de la nature, nous l'avons vu, est absolument capitale. Bien qu'il ait été dit, dans le premier chapitre, que la notion de loi de la nature était complexe et difficilement définissable, on peut affirmer sans trop se tromper que l'idée de loi de la nature en tant que généralisation universellement valide est assez près de l'idée d'une proposition universellement vraie. Ainsi, un modèle d'explication fondé sur la notion de loi générale de la nature semble difficilement compatible avec une théorie de la science qui rejette l'idée que la science a pour objectif de produire des propositions vraies à propos du monde⁴². Bref, la notion de loi de la nature étant en opposition avec l'esprit antiréaliste de l'empirisme constructif, il est clair qu'un modèle de l'explication fondé sur une telle notion est incompatible avec cette théorie de la science.

Par ailleurs, un modèle de l'explication scientifique purement causal, comme le propose Salmon, semble lui aussi incompatible avec l'empirisme constructif, et ce, pour une raison assez similaire: une théorie de l'explication selon laquelle expliquer un phénomène, c'est le situer dans la toile

⁴² D'ailleurs, dans son ouvrage *Laws and Symmetry* (1989), van Fraassen argumente contre l'existence des lois de la nature.

causale du monde, implique que la science est en mesure d'identifier correctement la toile causale du monde. En d'autres mots, une telle manière de concevoir l'explication scientifique dépend de l'idée que la science devrait être en mesure de nous dire la vérité à propos de la toile causale qui constitue le monde en général. Encore une fois, cela semble incompatible avec l'empirisme constructif. Une théorie de l'explication telle que défendue par Salmon est donc nécessairement dépendante d'une vision réaliste de la science, en ce sens qu'expliquer un phénomène, c'est dévoiler la vérité à propos des liens causaux qui ont mené à son existence, comme le souligne Hitchcock: « So, according to Salmon, explanations do reflect the causal structure of the world [...] » (1992, p.154).

C'est pourtant exactement cette thèse, selon laquelle une théorie antiréaliste de la science est incompatible avec un modèle causal de l'explication, que propose de rejeter Hitchcock. Selon lui, la position que l'on adopte concernant la question du réalisme ou de l'antiréalisme scientifique n'est pas limitative en ce qui concerne les théories de l'explication. Plus précisément, il soutient qu'une théorie causale de l'explication, traditionnellement associée avec le réalisme scientifique, est tout à fait compatible avec l'antiréalisme de van Fraassen (*Ibid.*, p.151). Prenons, par exemple, l'affirmation suivante, qui est en fait une explication de type causal: « La structure moléculaire de l'eau explique sa capacité à dissoudre le sel de table ». Naturellement, on serait portés à déduire que si l'on accepte cette explication, cela implique l'existence des atomes et des molécules, et donc qu'une telle explication est incompatible avec une vision antiréaliste de la science. Mais, selon Hitchcock, cela n'est pas le cas: pourquoi l'empiriste constructif ne pourrait-il pas interpréter cette proposition explicative comme il interprète les théories scientifiques, c'est-à-dire de manière littérale⁴³, mais sans pour autant endosser l'attitude épistémique qui consiste en croire en sa vérité: « [the anti-realist] can treat this explanatory claim exactly as he did the earlier descriptive claim, interpreting it realistically, but adopting an anti-realist epistemic attitude toward the claim. There is no prima facie reason why the anti-realist cannot adopt any interpretation of explanatory claims that is available to the realist » (*Ibid.*, p.155). Autrement dit, selon Hitchcock, on peut accepter une explication causale sans pour autant croire en sa vérité, comme on peut accepter une théorie scientifique sans croire en sa vérité.

Cependant, d'aucuns soutiennent qu'au contraire, on ne peut accepter une explication causale sans croire en la vérité des causes impliquées. C'est le cas de Nancy Cartwright, selon laquelle « To the

⁴³ Il a été dit, au chapitre précédent, que l'empirisme constructif est une théorie antiréaliste dans laquelle les théories scientifiques doivent être interprétées littéralement, c'est-à-dire sans en traduire le contenu dans un langage différent appartenant à un cadre conceptuel particulier.

extent that we find the causal explanation acceptable, we must believe in the causes described » (1983, p.5). Il ne sera pas ici question de nous pencher en détail sur les arguments de Cartwright, mais elle soutient cette thèse en endossant entre autres l'idée qu'une explication causale décrit des phénomènes ou des états de fait *actuels* pour rendre compte d'un événement *particulier*, contrairement à une explication qui utilise un modèle abstrait qui peut être utilisé pour expliquer divers phénomènes. Ainsi, selon Cartwright, il n'existe qu'une seule explication causale valide pour chaque phénomène, comme le note Hitchcock: « in the case of causal explanation, there is no tolerance for a diversity of explanations » (1992, p.159).

Pour défendre son point *contra* Cartwright, Hitchcock utilise un argument fondé sur certaines explications qu'on retrouve en physique quantique. Il utilise l'expérience paradigmatique des deux fentes, dans laquelle des électrons sont bombardés sur une surface qui comporte deux fentes microscopiques à une certaine distance l'une de l'autre, derrière laquelle se trouve un écran pour détecter la présence des électrons qui traversent lesdites fentes. L'expérience vise à déterminer, entre autres, la répartition des collisions sur l'écran pour comprendre la trajectoire des électrons qui traversent les fentes. Or, Hitchcock remarque que cette distribution des électrons sur l'écran est expliquée par des explications causales les plus standards qui soient, qui font intervenir, par exemple, la distance entre les deux fentes et la longueur d'onde des électrons bombardés. Cependant, cela est étrange, car ces explications impliquent que les électrons aient une position et un momentum déterminés, ce qui est rejeté par les physiciens actuels:

[the] proffered explanations contradict almost every interpretation of quantum mechanics, and as such, would not be believed as literally true by any but the most stubborn believer in hidden variable theories. The explanations given assume that the electron has definite position and momentum values, and that the uncertainty relations express our ignorance of these values, which results from our inability to measure position accurately while controlling the amount of momentum introduced into the system. According to quantum mechanical orthodoxy (and most quantum mechanical heresy), however, the electron does not have a simultaneously sharp position value and momentum value (*Ibid.*, p.171).

Autrement dit, Hitchcock tente de montrer qu'une explication causale n'implique pas la croyance en la vérité des causes qu'elle fait intervenir, étant donné que de telles explications sont fréquemment utilisées en physique quantique et que les physiciens les acceptent sans croire en la vérité de leurs implications. C'est ce que semble soutenir clairement Hitchcock: « These causal explanations, then, provide good candidates for counterexamples to the claim that acceptance of a causal explanation involves commitment to belief in the truth of the causal story. The causal explanations of the uncertainty relations would seem to be accepted, but not believed » (*Ibid.*, p.172).

Cet argumentaire développé par Hitchcock ne me semble pas particulièrement convaincant. À mon avis, il est quelque peu exagéré de déduire des conclusions d'ordre conceptuel aussi importantes (à savoir que l'acceptation d'une explication de type causal n'implique pas une interprétation réaliste de cette explication) à partir d'un exemple tiré de la physique quantique. Pourquoi? Parce que la physique quantique, simplement, est terriblement problématique pour nombre de notions de philosophie des sciences, et il me semble inadéquat de vouloir tirer des conclusions générales à propos de notions épistémologiques sur la base d'exemples foncièrement en tension avec les notions traditionnelles de philosophie des sciences. Bien sûr, la physique quantique constitue une part importante, sinon la part principale, de la physique moderne, mais les difficultés épistémologiques (et même scientifiques) qui lui sont liées sont tellement importantes qu'il est trop facile d'établir des conclusions générales contre-intuitives sur la base d'expériences de physique quantique. D'ailleurs, dans le deuxième chapitre, il a été noté que Salmon reconnaissait les difficultés auxquelles faisait face son modèle causal de l'explication lorsqu'appliqué à la physique quantique; or, cela ne l'a pas empêché de défendre son modèle, qui lui semblait adéquat dans le champ de la physique classique. Bref, sans vouloir nier la pertinence de l'argument de Hitchcock, je pense qu'il ne permet pas de rejeter si rapidement les intuitions de Cartwright et des autres philosophes qui voient un lien assez fort entre l'explication causale et le réalisme scientifique.

Que conclure de tout cela? L'empirisme constructif, en tant que théorie antiréaliste de la science, semble bel et bien limiter les conceptions de l'explication scientifique qui peuvent lui être jumelées. Par exemple, nous avons vu que le modèle déductif-nomologique semblait incompatible avec l'antiréalisme scientifique⁴⁴ tel que soutenu par van Fraassen à cause de la notion de loi de la nature, et au moins en forte tension avec un modèle purement causal de l'explication tel que proposé par Salmon. Est-ce à dire que l'empirisme constructif implique une théorie pragmatique de l'explication scientifique? Je ne pense pas; il existe certainement une multitude de conceptions de l'explication compatibles avec l'empirisme constructif, et plus généralement avec l'antiréalisme scientifique. Mais il n'en reste pas moins que la théorie pragmatique de l'explication proposée par van Fraassen est en adéquation avec l'esprit général de l'empirisme constructif. Il ne s'agit pas, évidemment, d'affirmer simplement que la théorie pragmatique des questions-pourquoi est compatible avec

⁴⁴ Cela peut paraître très étrange, alors que Hempel était un ardent défenseur de l'empirisme logique, qui est souvent considérée comme une vision antiréaliste de la science, en ce que ses défenseurs voulaient traduire les énoncés scientifiques faisant intervenir des entités inobservables en des énoncés vérifiables empiriquement. Les défenseurs de cette théorie de la science voulaient donc comprendre les théories scientifiques de manière non littérale, ce qui est exactement en opposition avec l'empirisme constructif de van Fraassen. Ainsi, je pense que le modèle déductif-nomologique est incompatible avec une théorie antiréaliste de la science qui prône la compréhension littérale des théories scientifiques.

l'empirisme constructif, ce qui ne serait pas d'un grand intérêt philosophique. Il s'agit plutôt de dire que cette théorie pragmatique de l'explication, en elle-même, incarne certaines idées fondamentales de la théorie de la science de van Fraassen.

En effet, on peut certainement affirmer que l'empirisme constructif est une théorie de la science qui fait preuve d'une souplesse épistémologique considérable, en ce sens que la vertu épistémique première qu'il incorpore est celle d'adéquation empirique, cette dernière étant épistémiquement plus faible que la notion de vérité. Par ailleurs, comme nous l'avons vu, l'empirisme constructif accorde une grande importance à l'aspect pragmatique lié à l'acceptation des théories scientifiques. La théorie pragmatique de l'explication comme théorie des questions-pourquoi paraît donc s'harmoniser remarquablement bien avec l'empirisme constructif. Non seulement le volet pragmatique de l'explication lié à la contextualité des questions-pourquoi fait écho à la dimension pragmatique de l'empirisme constructif en ce qui concerne les théories scientifiques, mais encore le fait qu'un même phénomène puisse être expliqué de différentes manières selon les paramètres de la question-pourquoi rappelle de manière intéressante la modestie épistémique de l'empirisme constructif. De fait, de même que la science ne vise pas à dévoiler la vérité sur le monde, de même il n'existe pas une seule explication véridique ou valide pour chaque phénomène, mais plutôt une pluralité d'explications possibles selon le point de vue que l'on adopte pour rendre compte du phénomène. C'est pourquoi la théorie pragmatique de l'explication des questions-pourquoi s'intègre parfaitement au sein de l'empirisme constructif.

Conclusion

Le premier chapitre de ce mémoire a été consacré à la présentation de deux des modèles de l'explication scientifique les plus influents de la seconde moitié du vingtième siècle, soit le modèle déductif-nomologique de Hempel et le modèle unificationniste de Friedman et de Kitcher. Dans le premier de ces deux modèles, une explication scientifique est conçue comme un argument dans lequel les prémisses, qui constituent l'*explanans*, doivent inclure une loi universelle de la nature et la description de circonstances empiriques particulières qui permettent de déduire l'*explanandum*, qui lui est le phénomène qu'on cherche à expliquer. Bref, l'explication d'un phénomène consiste à montrer que ce phénomène est une instanciation spatio-temporelle particulière d'une loi de la nature. Par ailleurs, nous avons vu que dans ce modèle, la notion de compréhension scientifique était intimement liée à celle d'anticipation ou de prévision: nous comprenons un phénomène lorsque nous pouvons l'anticiper en appliquant une loi universelle de la nature dans un contexte spatio-temporel particulier. Cependant, plusieurs problèmes sont rattachés à une telle manière de concevoir l'explication scientifique. Non seulement la notion même de loi universelle de la nature est controversée, mais le modèle déductif-nomologique engendre certaines difficultés conceptuelles, notamment celle des explications non pertinentes et l'incapacité de rendre compte de l'asymétrie de certaines explications.

Le modèle unificationniste tel que mis de l'avant par Friedman repose en grande partie sur l'intuition qu'un modèle de l'explication scientifique devrait être en mesure de lier fortement la notion de compréhension scientifique à celle d'explication. Friedman voit justement dans le concept d'unification la clef pour lier de manière intime ces deux dernières notions. Kitcher va reprendre cette intuition de Friedman pour élaborer un modèle unificationniste plus détaillé, selon lequel, comme dans le modèle déductif-nomologique hempélien, une explication est conçue comme un argument. Or, Kitcher voit la science comme étant un réservoir d'arguments explicatifs qui peuvent être utilisés pour expliquer différents phénomènes du monde. C'est en diminuant le plus possible le nombre de patrons argumentatifs qui constituent ce réservoir d'arguments que la science parvient à une unification des connaissances scientifiques et ainsi à une meilleure compréhension du monde. Autrement dit, c'est en utilisant un nombre limité de patrons argumentatifs pour expliquer une multitude de phénomènes qu'il devient possible de parler d'unification. Cette manière de comprendre l'explication permet, par exemple, d'apporter une solution au problème des explications non pertinentes: même s'il peut y avoir plusieurs explications possibles pour un phénomène donné, les explications que nous jugeons non pertinentes sont justement celles qui sont fondées sur des patrons argumentatifs qui ne réalisent pas une unification optimale des phénomènes. Il en est de

même pour le problème des explications asymétriques que le modèle hémipélicien ne permettait pas de résoudre. Cependant, le modèle unificationniste n'est pas sans impliquer lui aussi certaines difficultés, notamment celle de l'unification fallacieuse, qui revient à expliquer un phénomène à partir de lui-même en utilisant le patron argumentatif auto-déductif. En outre, on peut se demander si ce modèle, tel que développé par Kitcher, n'implique pas une sorte de rigueur extrême, en ce sens que seule l'explication construite sur la base du patron argumentatif qui réalise la meilleure unification possible serait réellement valide. Pourtant, il semble que l'explication soit une affaire de degrés (certaines explications sont meilleures que d'autres), et la position de Kitcher est difficilement conciliable avec une telle intuition.

Par ailleurs, les deux modèles de l'explication étudiés dans le premier chapitre ne font pas intervenir la notion de causalité, du moins pas de manière importante. Cette notion est pourtant non négligeable dans notre compréhension du monde, et elle deviendra absolument essentielle dans le modèle de l'explication causal mécanique que propose Wesley Salmon, qui a fait l'objet du deuxième chapitre. L'intuition centrale de Salmon est qu'une explication scientifique doit décrire la toile causale qui entoure l'*explanandum*. Pour développer davantage cette idée, il est donc nécessaire que Salmon se fonde sur une théorie de la causalité. Salmon a justement développé une théorie de la causalité centrée sur les notions de processus causal et d'interaction causale, par opposition à la notion plus traditionnelle d'événement. Selon le modèle de l'explication causal mécanique, contrairement aux modèles déductif-nomologique et unificationniste, une explication n'est pas conçue comme un argument, mais plutôt comme la conversion des relations de pertinence statistique entre attributs ou propriétés en des relations causales, ce qui devient possible, entre autres, grâce au concept de cause commune. Dit de manière concise, pour Salmon, l'explication d'un phénomène consiste en la mise en évidence de la portion de la toile causale, formée de processus causaux et d'interactions causales, dans laquelle se trouve le phénomène. Toutefois, à l'instar des deux modèles précédents, le modèle causal mécanique est à la source de quelques problèmes conceptuels. D'abord, il semble que ce modèle soit aussi victime de la difficulté des explications non pertinentes, en ce sens que les notions de processus causal et d'interaction causale ne permettent pas vraiment d'identifier les propriétés pertinentes du point de vue causal pour l'explication d'un phénomène donné. De plus, le modèle de Salmon semble difficilement applicable dans le cas des systèmes complexes constitués par une multitude de phénomènes et dans le cas de certains phénomènes étudiés en physique quantique.

Dans le troisième chapitre, nous avons temporairement laissé de côté les réflexions à propos des modèles de l'explication scientifique comme tels pour nous tourner vers la théorie de l'empirisme

constructif, fondée par Bas van Fraassen. Cette théorie antiréaliste de la science a comme principale thèse que la science vise à fournir des théories empiriquement adéquates (par opposition à la vision du réalisme scientifique, selon lequel la science a pour but de dire la vérité à propos du monde), et que l'acceptation d'une théorie scientifique implique seulement, du point de vue épistémique, la croyance en son adéquation empirique. Cette dernière notion d'adéquation empirique, étant au cœur de la caractérisation de l'empirisme constructif, demande d'être clairement saisie. Nous avons vu que l'adéquation empirique ne concerne que les phénomènes observables: une théorie est empiriquement adéquate lorsqu'elle est vraie à propos des phénomènes observables. Plus précisément, van Fraassen soutient une vision sémantique des théories scientifiques selon laquelle une théorie est empiriquement adéquate lorsque tous les phénomènes observables peuvent être reconnus dans l'une des substructures empiriques des modèles qui constituent cette théorie. Par ailleurs, l'acceptation d'une théorie scientifique comporte aussi une dimension pragmatique, en ce sens qu'accepter une théorie scientifique, c'est, par exemple, s'engager dans un certain programme de recherche et tenter de rendre compte des phénomènes particuliers nouvellement observés avec cette théorie.

Par ailleurs, la notion d'observabilité, sur laquelle s'appuie celle d'adéquation empirique, a aussi été abordée. Si van Fraassen ne définit pas rigoureusement cette notion, il considère qu'une entité X est observable s'il existe des circonstances dans lesquelles il nous serait possible d'observer X. Le « nous », ici, fait référence à l'ensemble des êtres humains; d'ailleurs, van Fraassen soutient que la notion d'observabilité est toujours relative à une communauté épistémique particulière. Ainsi, l'ensemble formé par toutes les choses observables n'est pas nécessairement le même selon la communauté épistémique concernée. En outre, les membres de la communauté épistémique pouvant faire eux-mêmes partie des objets sur lesquels portent les théories scientifiques, il en découle que ce sera, en fin de compte, les théories scientifiques qui seront en mesure de délimiter le domaine de l'observable selon les caractéristiques physiques et biologiques des membres de la communauté épistémique. Autrement dit, selon van Fraassen, la question de savoir ce qui est observable n'est pas une question qui se répond *a priori*, mais bien une question d'ordre empirique. Van Fraassen prend toutefois bien soin de montrer en quoi cette réciprocity entre la théorie scientifique et l'observabilité n'est pas logiquement problématique, en précisant que le domaine de l'observable, même s'il est descriptible par la science, n'est pas pour autant dépendant des théories scientifiques. Il faut toutefois noter que la notion d'observabilité telle que présentée par van Fraassen est loin de faire l'unanimité. Ladyman, entre autres, a construit une critique de l'empirisme constructif fondée sur cette notion d'observabilité en soutenant qu'elle est à la source d'énoncés modaux qui sont, selon lui,

clairement incompatibles avec l'empirisme constructif, en ce qu'ils impliquent la croyance en la vérité d'énoncés conditionnels qui portent sur des phénomènes qui n'existeront jamais de manière actuelle. Cette brève incursion dans l'empirisme constructif nous aura servi de propédeutique en ce qui concerne les développements effectués dans le dernier chapitre du mémoire.

En effet, nous avons poursuivi notre étude, dans le quatrième chapitre, en nous intéressant à la théorie de l'explication que défend van Fraassen, qui fait partie de la famille des théories pragmatiques de l'explication, selon lesquelles les facteurs pragmatiques (liés au contexte d'explication) sont constituants de la notion théorique d'explication. Par ailleurs, nous avons aussi abordé la question des liens généraux entre une théorie scientifique et sa capacité explicative, en soulignant notamment qu'une théorie scientifique n'a pas besoin d'être empiriquement adéquate, de manière globale, pour faire partie d'une explication. Effectivement, van Fraassen s'intéresse principalement à la relation d'explication entre la théorie scientifique et les faits (relation qui inclut aussi, de manière primordiale, le contexte d'explication défini par la personne qui donne l'explication, la personne qui la reçoit, etc.), ce qui ne concerne pas l'acceptabilité générale de la théorie scientifique en question.

Par la suite, nous avons vu que le contexte est intimement lié à la relation d'explication, et cela, de deux manières différentes. D'abord, la notion de causalité, qui est importante, comme le soutenait Salmon, pour la notion d'explication, comporte un aspect contextuel important. Les relations causales pouvant être explicitées à l'aide d'énoncés conditionnels, et ces énoncés conditionnels étant toujours accompagnés d'une clause *ceteris paribus* implicite, il s'ensuit que les facteurs pertinents qui peuvent entrer dans une explication, qui seront identifiés comme les causes d'un événement, dépendent directement du contexte dans lequel les énoncés conditionnels décrivant la relation de causalité sont inscrits. Deuxièmement, en reprenant et en précisant la théorie des questions-pourquoi de Bromberger, van Fraassen montre que l'identification de la classe contrastante, qui déterminera le champ des réponses possibles à une question-pourquoi, dépend aussi très fortement du contexte dans lequel est posée la question.

Il nous a alors été possible de comprendre la nature d'une explication selon van Fraassen: une explication est une réponse à une question-pourquoi, cette question étant elle-même caractérisée par son thème, la classe contrastante qu'elle implique et la relation de pertinence explicative qu'elle soulève. Une réponse à une question-pourquoi (donc, une explication) sera évaluée en tenant compte de la vérité de sa proposition principale, du fait qu'elle favorise ou non le thème de la question par rapport aux autres membres de la classe contrastante et en la comparant aux autres

réponses possibles qui auraient pu être données pour répondre à la question. Ainsi, la théorie de l'explication de van Fraassen est clairement une théorie pragmatique en ce que le contexte est absolument essentiel pour comprendre et définir ce qu'est une explication. Il faut aussi se rappeler qu'une explication scientifique n'est pas fondamentalement différente, selon van Fraassen, de toute autre forme d'explication: une explication scientifique est simplement une explication qui repose sur des informations qui proviennent de théories scientifiques.

J'ai finalement soutenu la thèse que la théorie pragmatique de l'explication de van Fraassen, tout en étant certainement compréhensible de façon relativement indépendante de l'empirisme constructif, est fidèle à la modestie épistémique de l'empirisme constructif, de sorte que ces deux théories s'harmonisent de manière manifeste. J'ai construit mon argumentaire à partir de l'idée que l'empirisme constructif, en tant que théorie antiréaliste de la science, limitait clairement les conceptions de l'explication scientifique qui peuvent lui être rattachées. J'ai défendu l'idée que le modèle déductif-nomologique hempélien ainsi que le modèle causal mécanique de Salmon, par exemple, ne sont pas compatibles avec l'antiréalisme de l'empirisme constructif. Cependant, certains auteurs, notamment Hitchcock, nient que la défense de l'antiréalisme scientifique par opposition au réalisme, ou vice-versa, implique quoi que ce soit en ce qui regarde l'adoption d'une conception particulière de l'explication scientifique. Il faudrait donc étudier en profondeur cette question pour être en mesure de défendre rigoureusement la position que j'ai voulu soutenir. L'étude sérieuse d'une notion particulière de philosophie des sciences, ici celle d'explication scientifique, me semble ainsi nous ramener inévitablement, en fin de compte, à la question fondamentale qu'est celle de la définition de la science.

Bibliographie

Alspector-Kelly, Marc (2001). « Should the Empiricist Be a Constructive Empiricist », *Philosophy of Science*, 68, 4, pp.413-431.

Bromberger, Sylvain (1966). « Why-Questions », in *Mind and Cosmos*, R. G. Colodny (ed.), Pittsburgh, University of Pittsburgh Press, pp.86-108.

Cartwright, Nancy (1983). *How the Laws of Physics Lie*, Oxford, Oxford University Press.

Chakravartty, Anjan (2005). « Causal Realism: Events and Processes », *Erkenntnis*, 63, pp.7-31.

Contessa, Gabriele (2006). « Constructive Empiricism, Observability, and Three Kinds of Ontological Commitment », *Studies in History and Philosophy of Science Part A*, 37, 3, pp.454-468.

Churchland, Paul & Clifford A. Hooker (eds.) (1985). *Images of Science: Essays on Realism and Empiricism*, Chicago, University of Chicago Press.

Churchland, Paul (1985). « The Ontological Status of Observables: In Praise of the Superempirical Virtues », in P. Churchland & C. Hooker (eds.), 1985, pp.35-47.

De Regt, Henk W., Sabrina Leonelli & Kai Eigner (eds.) (2009). *Scientific Understanding: philosophical perspectives*, Pittsburgh, University of Pittsburgh Press.

Dicken, Paul (2007). « Constructive Empiricism and the Metaphysics of Modality », *British Journal for the Philosophy of Science*, 58, pp.605-612.

Friedman, Michael (1974). « Explanation and Scientific Understanding », *The Journal of Philosophy*, 71, 1, pp.5-19.

Gijssbers, Victor (2007). « Why Unification is neither Necessary nor Sufficient for Explanation », *Philosophy of Science*, 74, 4, pp.481-500.

Hacking, Ian (1985). « Do We See Through a Microscope », in P. Churchland & C. Hooker (eds.), 1985, pp.132-152.

Hempel, Carl G. (1965). *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, New York, Free Press.

--- (2001). *The Philosophy of Carl G. Hempel*, James H. Fetzer (ed.), New York, Oxford University Press.

Hitchcock, Christopher Read (1992). « Causal Explanation and Scientific Realism », *Erkenntnis*, 37, 2, pp.151-178.

--- (1995). « Salmon on Explanatory Relevance », *Philosophy of Science*, 62, 2, pp.304-320.

Kitcher, Philip (1981). « Explanatory Unification », *Philosophy of Science*, 48, 4, pp. 507-531.

--- (1989). « Explanatory Unification and the Causal Structure of the World », in *Scientific Explanation*, P. Kitcher & W. Salmon (eds.), Minneapolis, University of Minnesota Press, pp.410-505.

Ladyman, James (2000). « What's Really Wrong with Constructive Empiricism? Van Fraassen and the Metaphysics of Modality », *British Journal for the Philosophy of Science*, 51, pp.837-856.

--- (2004). « Constructive Empiricism and Modal Metaphysics: A Reply to Monton and van Fraassen », *British Journal for the Philosophy of Science*, 55, pp.755–765.

Leeds, Stephen (1994). « Constructive Empiricism », *Synthese*, 101, 2, pp.187-221.

Monton, Bradley & Bas van Fraassen (2003). « Constructive Empiricism and Modal Nominalism », *British Journal for the Philosophy of Science*, 54, pp.405-422.

Reichenbach, Hans (1956). *The Direction of Time*, Berkeley, University of California Press.

Rosen, Gideon (1994). « What Is Constructive Empiricism », *Philosophical Studies*, 74, 2, pp.143-178.

Russell, Bertrand (1948). *Human Knowledge: Its Scope and Limits*, London, G. Allen & Unwin.

Salmon, Wesley (1971). « Statistical Explanation », in *Statistical Explanation and Statistical Relevance*, W. Salmon (ed.), Pittsburgh, University of Pittsburgh Press, pp.29-87.

--- (1984). *Scientific Explanation and the Causal Structure of the World*, Princeton, Princeton University Press.

--- (1984). « Scientific Explanation: Three Basic Conceptions », *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, pp.293-305.

--- (1985). « Conflicting Conceptions of Scientific Explanation », *The Journal of Philosophy*, 82, 11, pp.651-654.

--- (1989). *Four Decades of Scientific Explanation*, Minneapolis, University of Minnesota Press.

--- (1998). « Comets, Pollen, and Dreams: Some Reflections on Scientific Explanation », *Causality and Explanation*, New-York, Oxford University Press, pp.50-66.

Sober, Elliott (1984). « Common Cause Explanation », *Philosophy of Science*, 51, 2, pp.212-241.

Strevens, Michael (2009). « An Argument Against the Unification Account of Explanation » (non publié), <http://www.strevens.org/research/expln/Unificatio.pdf>

Van Fraassen, Bas (1980). *The Scientific Image*, Oxford, Clarendon Press.

--- (1985). « Empiricism in the Philosophy of Science », in P. Churchland & C. Hooker (eds.), 1985, pp.245-308.

--- (1989). *Laws and Symmetry*, Oxford, Oxford University Press.

--- (1998). « The Agnostic Subtly Probabilified », *Analysis*, 58, 3, pp.212-220.

--- (2007). « From a View Of Science to a New Empiricism », in *Images of Empiricism: Essays in Science and Stances, with a Reply from Bas C. van Fraassen*, B. Monton (ed.), Oxford, Oxford University Press, pp.337-383.

Weber Erik, Jeroen Van Bouwel & Leen De Vreese (eds.) (2013). *Scientific Explanation*, Springer.

Woodward, James (1989). « The Causal Mechanical Model of Explanation », in *Scientific Explanation*, P. Kitcher & W. Salmon (eds.), Minneapolis, University of Minnesota Press, pp.357-383.

--- (2011), « Scientific Explanation », *The Stanford Encyclopedia of Philosophy* (Winter 2011 Edition), Edward N. Zalta (ed.), <http://plato.stanford.edu/archives/win2011/entries/scientific-explanation/>.

--- (2014), « Scientific Explanation », *The Stanford Encyclopedia of Philosophy* (Winter 2014 Edition), Edward N. Zalta (ed.), <http://plato.stanford.edu/archives/win2014/entries/scientific-explanation/>.