

Université de Montréal

Causalité et explication causale chez James Woodward

par

Vincent Roy

Département de Philosophie  
Faculté des arts et des sciences

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

Août 2015

© Vincent Roy

## RÉSUMÉ

Certains philosophes affirment que les relations causales sont fondées sur les lois de la nature. Cette conception cadre mal avec la réalité des sciences biomédicales et des sciences humaines. Pour se rapprocher de la pratique réelle des diverses sciences, James Woodward propose une conception de la causalité et de l'explication causale fondée sur une relation beaucoup moins exigeante que celle de loi de la nature, qu'il appelle l'invariance. Le but de ce mémoire est de présenter le concept d'invariance et les autres concepts causaux qui s'y rattachent et, d'identifier certaines difficultés, dans le but de cerner l'usage approprié de cette famille de concepts.

La conception causale de Woodward suppose que le but de la recherche des causes est pratique plutôt que simplement épistémique : il s'agit pour les agents de s'appuyer sur les causes pour modifier les phénomènes. Cette conception est également non-réductive; elle utilise des contrefactuels et reflète les méthodes expérimentales des diverses sciences.

La cohérence de cette conception avec les généralisations causales réelles des sciences fait en sorte qu'elle abandonne l'objectif d'universalité rattaché à la notion de loi de la nature, en faveur d'un objectif de fiabilité temporaire. De plus, comme le critère d'invariance est peu exigeant, d'autres critères doivent lui être ajoutés pour identifier, parmi les relations causales (c'est-à-dire invariantes), les relations les plus susceptibles d'être employées pour modifier les phénomènes de façon fiable.

Mots clés : Philosophie, épistémologie, causalité, explication, loi de la nature, intervention, invariance, Woodward, science.

## **ABSTRACT**

Some philosophers claim that causal relations are based on the laws of nature. This view is not consistent with the actual causal relations found in the biomedical and social sciences. In order to better reflect the actual practice in the various sciences, James Woodward puts forward a view of causation and causal explanation based on a much less demanding relation than that of law of nature, which he calls 'invariance'. This essay presents the concept of invariance and other related causal concepts, and identifies certain problems, in order to outline the proper use of this group of concepts.

Woodward's conception assumes that the goal of causal inquiry is practical rather than merely epistemic: agents use causal relationships to modify outcomes. The conception is also non reductive; it uses counterfactuals and is a reflection of the experimental methods of the various sciences.

The fact that this conception is consistent with the actual causal generalizations found in the various sciences implies that the goal of universal truth associated with the notion of law of nature is set aside and is replaced by an objective of temporary reliability. In addition, since the invariance criterion is not very demanding, other criteria must be added to identify, among causal (i.e. invariant) relations, those relations that can be relied upon to reliably modify outcomes.

Keywords: Philosophy, epistemology, causation, explanation, law of nature, intervention, invariance, Woodward, science.

## TABLE DES MATIÈRES

<b>CHAPITRE 1 : INTRODUCTION .....</b>	<b>1</b>
<b>CHAPITRE 2 : LES CONCEPTS MINIMAUX DE CAUSALITÉ CHEZ WOODWARD .....</b>	<b>6</b>
2.1 Le projet de Woodward.....	6
1) Le projet .....	6
2) Trois illustrations préliminaires.....	9
2.2 Les concepts causaux minimaux chez Woodward .....	16
1) La formalisation.....	16
2) Les concepts causaux : cause totale, cause directe, cause contributoire et intervention.....	22
2.3 Une conclusion sur les concepts causaux minimaux.....	29
<b>CHAPITRE 3 : LES CONCEPTS CAUSAUX RENFORCÉS.....</b>	<b>32</b>
3.1 Un détour par la notion d'explication .....	33
3.2 Deux modèles nomologiques de l'explication scientifique : les modèles DN et IS.....	34
3.3 Les lois de la nature dans les modèles DN et IS et les problèmes de ces modèles.....	39
3.4 La conception manipulative de l'explication causale .....	49
1) Les explications se rattachent à des changements.....	52
2) Les explications satisfaisant EXP donnent souvent une compréhension limitée .....	54
3) Les possibilités sérieuses.....	56
4) Le contraste.....	58
5) Le niveau de l'explication.....	60
3.5 L'invariance.....	61
3.6 L'invariance et les contrefactuels.....	67
3.7 Réévaluation de certains problèmes des théories nomologiques de l'explication.....	69
3.8 Les degrés d'invariance .....	78
1) La stabilité .....	81
2) La proportionnalité ou le choix du niveau approprié d'explication .....	84
3) La spécificité.....	86

<b>CHAPITRE 4 : ÉLÉMENTS DE CRITIQUE.....</b>	<b>91</b>
4.1 Des gains et au moins une perte .....	91
4.2 Échec dans le cas de théories physiques appliquées à l'ensemble de l'univers .....	93
4.3 Les mécanismes.....	95
4.4 L'explication dans les situations de causalité minimale.....	99
4.5 Les possibilités sérieuses .....	104
<b>CHAPITRE 5 : CONCLUSION.....</b>	<b>105</b>
<b>BIBLIOGRAPHIE.....</b>	<b>109</b>

## LISTE DES FIGURES ET DES TABLEAUX

Fig. 2.1 Un bloc sur un plan incliné.....	14
Fig. 2.2 Un graphe orienté.....	20
Fig. 2.3 Deux relations différentes ont le même graphe orienté.....	21
Fig. 2.4 X n'est pas une cause directe de Y.....	24
Fig. 3.1 Concepts causaux minimaux et renforcés.....	52

## LISTE DES PRINCIPALES CARACTÉRISATIONS

2.1.3 Intervention (caractérisation qualitative) .....	12
2.2.3 Cause totale .....	23
2.2.4 Cause directe .....	24
2.2.9 Cause contributoire .....	26
2.2.10 Intervention (caractérisation technique) .....	28
3.2.1 Modèle DN.....	35
3.2.2 Modèle IS.....	37
3.4.1 Explication .....	50
3.5.3 Intervention de testing .....	63
3.5.4 Généralisation invariante .....	64
3.5.5 Relation causale générale ('type') .....	64
3.8.1 Relation causale générale ('type') (version alternative).....	79
3.8.3 Stabilité .....	81
3.8.11 Proportionnalité .....	85
3.8.13 Spécificité.....	87
4.4.1 Généralisation de cause possible .....	99

## REMERCIEMENTS

Je me suis engagé dans l'exercice de rédiger le présent mémoire en suivant un parcours intellectuel un peu tortueux. Certaines personnes m'ont orienté et influencé de façon particulière.

Je remercie d'abord certains professeurs de philosophie à la retraite qui partagent leur passion pour cette discipline en donnant des cours et des conférences destinées au grand public. Ce sont eux qui ont rallumé en moi la flamme de l'interrogation philosophique. Je remercie particulièrement Jacques Senécal qui, par ses assertions provocatrices sur la liberté, m'a en quelque sorte poussé vers les problèmes rattachés à la causalité.

Je remercie principalement mon directeur de mémoire, Daniel Laurier. Il a d'abord été une inspiration pour moi, en vertu du fait qu'il aborde de façon exemplaire les problèmes philosophiques dans la tradition de la philosophie analytique. Cette approche me semble la plus naturelle, ou du moins correspond à mes propres intuitions philosophiques. M. Laurier m'a encouragé à entreprendre la rédaction de ce mémoire et il m'a laissé trouver la voie qui me convenait en matière de causalité, qui s'est ultimement trouvée être plus proche de la philosophie des sciences que de la métaphysique. Je le remercie pour sa disponibilité et pour ses commentaires toujours pertinents.



## CHAPITRE 1. INTRODUCTION

Je présente dans ce mémoire la conception de James Woodward de la causalité et de l'explication causale. Je me fonde principalement sur son livre de synthèse «Making Things Happen. A Theory of Causal Explanation», publié en 2003, ainsi que sur divers articles de Woodward et d'autres auteurs sur des aspects particuliers de la causalité et de l'explication.

Woodward propose sa conception en réaction aux difficultés bien connues des conceptions nomologiques de l'explication scientifique. Selon les conceptions nomologiques, l'explication d'un phénomène particulier est un raisonnement, déductif ou inductif<sup>1</sup>, qui comprend deux genres de prémisses : (1) des données décrivant les conditions initiales prévalant avant (ou pendant) le phénomène à expliquer, et (2) des lois, déterministes ou probabilistes. Le phénomène à expliquer (*l'explanandum*) s'énonce alors comme le résultat d'une inférence déductive ou inductive découlant des prémisses. La principale difficulté des conceptions nomologiques provient du fait que le paradigme des lois envisagées est celui des lois fondamentales de la nature en physique, comme les lois de Maxwell en électromagnétisme. Ces lois fondamentales s'appliqueraient de façon universelle dans le temps et l'espace, et n'auraient pas d'exception. L'application universelle et l'absence d'exception font en sorte que les conclusions du raisonnement explicatif sont très fiables, voire certaines. Le caractère universel et sans exception des lois devient une difficulté à partir du moment où l'on constate que les généralisations des sciences biomédicales et des sciences humaines ne sont pas d'application universelle et ont de nombreuses exceptions. Le tenant d'une des conceptions nomologiques se trouve alors face à un dilemme : ou bien, il accepte qu'il n'y a pas d'explication valable dans le domaine biomédical et dans les sciences humaines, ou bien il édulcore d'une manière ou d'une autre la notion de loi. Une seconde difficulté (connexe à la première) des conceptions nomologiques est qu'il n'y a pas d'entente, ni chez les philosophes, ni chez les scientifiques, sur les caractéristiques qui permettent de distinguer clairement l'énoncé d'une loi, par opposition à l'énoncé d'une généralisation vraie de façon accidentelle. Un exemple d'énoncé difficile à classer est : «Tous les corbeaux sont noirs». S'agit-il d'une loi ou d'une généralisation vraie de façon accidentelle ?

---

<sup>1</sup> Selon un sens particulier du mot «inductif» qui sera précisé plus loin

L'objectif principal de Woodward, face aux difficultés des conceptions nomologiques, est de concevoir différemment les causes et les explications, scientifiques et autres, de façon à mieux comprendre, et ainsi à prescrire, les conditions des bonnes explications, valables pour les sciences dites «dures» tout autant qu'en biologie et en sciences sociales.

Woodward s'attaque à cette tâche avec certains présupposés. (1) Il existe des relations causales dans l'univers. Il n'est pas nécessaire de démontrer cette assertion. Le but d'un exercice de clarification des concepts causaux est plutôt de donner à l'agent humain des outils pour identifier les phénomènes véritablement causaux là où ils se trouvent. (2) La recherche des causes chez les humains et chez certains animaux vise un objectif d'abord pratique, plutôt qu'un objectif principalement épistémique. C'est pourquoi, il arrive fréquemment que les agents se contentent de savoir qu'il existe un lien de cause à effet entre un premier aspect de la réalité et un second aspect, et ensuite d'exploiter ce lien pour modifier la réalité, même s'ils ne connaissent pas le détail ou le mécanisme en vertu duquel la relation fonctionne. On a par exemple identifié et utilisé la prise d'aspirine pour atténuer la douleur bien avant de connaître le mécanisme en vertu duquel l'aspirine agit sur certaines douleurs. (3) Dans la mesure où il est possible et intéressant d'établir et d'utiliser l'existence d'une relation répétable entre tel type de «chose» (A), à titre de cause, et tel autre type de «chose» (B), à titre d'effet, même si on ne connaît pas le mécanisme qui les relie, il est essentiel de pouvoir distinguer entre les situations où A et B sont «simplement corrélés» (au sens où ils sont souvent ou toujours présents ensemble sans être en relation de cause à effet), et les situations où A et B sont véritablement en relation de cause à effet.

Ces trois présupposés se reflètent chez Woodward par la production de deux conceptions de la relation causale. Premièrement, une conception «minimale» qui n'exige pas de connaissance des détails de la relation causale, mais qui permet néanmoins de distinguer les véritables relations causales des simples corrélations. Je présenterai cette conception au chapitre 2.

Comme la première conception n'exige que peu ou pas d'information sur le détail de la relation entre les causes et les effets, elle est souvent peu satisfaisante au plan de l'explication. C'est pourquoi Woodward s'intéresse ensuite aux situations où nous disposons d'informations plus importantes sur la relation entre les causes et les effets. Woodward produit à cet égard une

seconde conception causale «renforcée» qui inclut explicitement des généralisations décrivant les relations entre les causes et les effets. Ces relations peuvent être déterministes ou probabilistes. Woodward développe cette seconde conception référant à des généralisations à titre de substitut aux conceptions fondées sur les lois (i.e. les conceptions nomologiques) mentionnées plus haut. L'univers des relations causales telles que les conçoit Woodward peut donc être envisagé en trois sous-groupes épistémiques :

- 1) Les situations causales où nous disposons d'un modèle déterministe, caractérisé grossièrement par le fait que la généralisation décrivant la relation causale n'admet pas d'exceptions dans son domaine d'application.
- 2) Les situations causales où nous disposons d'un modèle probabiliste, caractérisé grossièrement par le fait que la généralisation décrivant la relation causale admet des exceptions dans son domaine d'application.
- 3) Les situations causales où nous ne disposons pas de suffisamment d'informations pour produire une généralisation probabiliste stable. Ces situations sont celles où les concepts causaux minimaux s'appliquent.

Je présenterai les conceptions nomologiques de l'explication causale, et ensuite les conceptions causales «renforcées» de Woodward au chapitre 3 de ce mémoire. Le chapitre 4 portera sur certaines limites des conceptions de Woodward. Le chapitre 5 présentera les conclusions du mémoire.

Avant d'aborder le détail des conceptions de Woodward, je veux attirer l'attention du lecteur sur un aspect particulier de la perspective de ce philosophe : quelle est la matière première de son analyse conceptuelle des notions d'explication et de causalité ? Il me semble que les philosophes disposent d'au moins deux avenues à cet égard. Soit ils abordent ces questions principalement à partir d'expériences de pensée fondées sur nos intuitions causales dans la vie quotidienne. Ils proposeront, à titre d'exemple, la situation de deux tireurs qui tirent en même temps sur une cible humaine en vue de tuer la personne. Il s'agit alors d'un cas de surdétermination causale et la question se pose : étant donné une caractérisation proposée de la relation causale, y a-t-il, dans ce cas, une cause du décès, deux causes, ou encore aucune réponse à la question ? Cette approche par l'analyse d'expériences de pensée est celle qui est retenue par, entre autres, les penseurs de la

causalité et de l'explication qui travaillent dans la tradition inaugurée par David Lewis. Soit, au contraire, les philosophes abordent ces questions principalement à partir des diverses sciences, de la physique aux sciences humaines. On étudie les résultats de ces disciplines pour classer les relations de dépendance qu'on y trouve et identifier les conditions de relations de dépendance véritablement causales. Un exemple d'une telle analyse porte sur les questions entourant l'assertion : «Fumer cause le cancer». Quels étaient la signification et le statut de vérité de l'assertion à l'époque où l'on ne connaissait pas le mécanisme de l'action du tabac sur les cellules du poumon ? Quels sont la signification et le statut de vérité de l'assertion aujourd'hui, alors qu'on connaît ce mécanisme, mais qu'on sait aussi que de nombreuses personnes fument sans être atteintes du cancer ?

Woodward, à titre de philosophe des sciences, appartient au groupe qui se fonde principalement sur les résultats des sciences. Il faut dire que les tenants des conceptions nomologiques de l'explication font également ce choix. J'exprimerai ici, sans le défendre de façon approfondie, un préjugé favorable à l'égard de ce choix. Je crois que ce choix est défendable pour au moins deux raisons. Premièrement, il me semble que la recherche conceptuelle devrait procéder en commençant avec des situations simples pour ensuite s'attaquer aux situations plus complexes. Or les objets qu'on étudie afin d'y déceler les relations causales (ou non-causales) sont plus simples dans les sciences comme la physique et la chimie, dont les objets fondamentaux sont des particules, des atomes et des molécules, ou même comme la biologie, dont les objets de base sont des cellules et des êtres vivants souvent simples. Par contre, les agents des expériences de pensée des philosophes de l'école de Lewis sont des agents humains auxquels on rattache des capacités de décision et d'action complexes, comme par exemple, dans le cas d'agents qui tirent sur une cible humaine, la capacité de retarder ou de devancer un tir si un agent voit (ou simplement prévoit) qu'un second agent vise (ou visera) la même cible. La complexité des capacités des agents dans les expériences de pensée permet de multiplier presque sans fin les exemples et les contre-exemples, ce qui réduit la possibilité d'arriver à des conclusions même minimalement nettes. Deuxièmement, et c'est peut-être une façon d'exprimer autrement la première raison, j'estime que la méthode par les expériences de pensée laisse une part trop grande à la raison (au sens de faculté de raisonnement humain). Je souscris aux intuitions méthodologiques de Kant qui pensait que la raison, si elle ne s'appuie pas sur l'expérience, tend à s'égarer. C'est pourquoi il me semble, en ce

qui concerne l'analyse conceptuelle de la causalité et de l'explication causale, que l'approche fondée principalement sur les résultats des diverses sciences est à privilégier.

## CHAPITRE 2. LES CONCEPTS MINIMAUX DE CAUSALITÉ CHEZ WOODWARD

Je débiterai ce chapitre (section 2.1) en apportant certaines précisions sur la portée philosophique du projet de Woodward. S'agit-il d'un projet conceptuel, visant à caractériser les concepts causaux, ou d'un projet ontologique, visant, en plus d'une caractérisation conceptuelle, à préciser la nature des liens causaux, ou encore à résoudre la question du déterminisme : les événements du monde sont-ils ou non entièrement déterminés par des causes ? Je présenterai ensuite trois illustrations de situations potentiellement causales, qui permettront d'introduire certains des concepts causaux abordés dans la section suivante. La section 2.2 caractérisera les concepts causaux dans une perspective minimale, c'est-à-dire sans présumer que l'utilisateur connaît les détails de la relation entre les causes et les effets. Dans la section 2.3, j'aborderai un problème ou une tension qui affecte selon moi les concepts causaux minimaux présentés dans la section 2.2.

### 2.1 Le projet de Woodward

#### 1) Le projet

Les caractérisations de la causalité peuvent être de plusieurs types, selon l'objectif visé. Une caractérisation peut être principalement conceptuelle. Il s'agit alors, en partant d'une notion vague du langage ordinaire, de préciser la signification que les utilisateurs rattachent (ou devraient rattacher) aux expressions causales. Une caractérisation conceptuelle peut être réductive, comme lorsque l'on affirme que le concept de causalité se ramène à un concept non-causal plus fondamental. David Lewis a par exemple tenté de montrer que le concept de causalité se réduit au concept de dépendance contrefactuelle, où la dépendance contrefactuelle est considérée comme un concept non-causal. Mais une caractérisation conceptuelle peut aussi être non-réductive. On tente alors de préciser la signification que les utilisateurs rattachent au vocabulaire causal sans exiger que les termes employés soient non-causaux. Une telle caractérisation non-réductive pourrait sembler entachée de circularité. Mais cette circularité apparente ne fait pas nécessairement en sorte que la caractérisation n'ait aucune valeur.

Une seconde grande catégorie de caractérisations est celle des caractérisations ontologiques. Il s'agit alors de dire ce qui, dans le monde réel, constitue une relation causale, ou encore ce qui, dans le monde réel, est le vérifacteur ('truth-maker') de cette relation. Certains métaphysiciens

soutiennent par exemple que la relation causale contient un élément de nécessité, que cette nécessité est exprimée par des énoncés généraux qu'on appelle parfois «lois de la nature», et que l'élément du monde réel qui fonde ces lois de la nature (leur vérificateur) est le fait qu'il y a des relations nécessaires entre certaines propriétés. Dans certaines autres conceptions de la relation causale, la relation causale est d'abord un processus, et le vérificateur de la relation causale est une quantité physique conservée dans le déroulement de ce processus. Les caractérisations ontologiques réductives sont les plus ambitieuses, car elles proposent une caractérisation unique et précise des relations causales, alors que l'expérience du monde réel semble indiquer une grande diversité d'instances de la causalité. Il va de soi aussi que si un philosophe adopte une caractérisation ontologique de la causalité, il adopte par le fait même une caractérisation conceptuelle parallèle.

Cette brève présentation de certains types de caractérisation me permet maintenant de situer le projet de Woodward, et de réduire la confusion qui découle souvent, au moment d'aborder une conception particulière de la causalité, de la diversité des conceptions concurrentes. Woodward ne propose pas une conception ontologique de la causalité. Il admet certes qu'il y a probablement un ou des vérificateurs (qu'il appelle parfois des «mécanismes») dans chaque situation causale particulière, (sauf peut-être dans les situations véritablement non-déterministes), mais il ne tente pas de caractériser directement la causalité par une description générale de son/ses vérificateurs. De plus, comme je l'ai indiqué au chapitre 1, il croit que les humains peuvent identifier et utiliser les relations causales dans une situation particulière sans nécessairement connaître le ou les mécanismes qui entrent en jeu dans cette situation.

Woodward considère plutôt son projet comme celui d'une caractérisation (i) conceptuelle et (ii) non-réductive, de la causalité et de l'explication causale. Voici quelques-unes de ses affirmations à l'égard de son projet entendu comme caractérisation conceptuelle.

«My enterprise is, roughly, to provide an account of the meaning or content of just those qualitative causal notions that Pearl [un autre chercheur travaillant sur la causalité et l'explication] [...] takes as primitive.»<sup>2</sup>

Dans un article sur la causalité en biologie, Woodward écrit :

---

<sup>2</sup> Woodward, J. (2003), p. 38

«Although I have attempted to provide biological illustrations of these causal notions, my primary interest is in the content of the notions themselves and less in the empirical details of the illustrations. [Par exemple, à propos du caractère «spécifique» de la relation causale entre l'ADN et la protéine qu'il encode,] «my concern is with what this claim means – with the empirical features that biologists believe this relationship to possess which leads them to think of it as specific – and only secondarily with the complicated and controversial question of whether the relationship in fact possesses these features.»<sup>3</sup>

D'autre part, Woodward justifie comme suit le caractère non-réductif de sa conception de la causalité et de l'explication causale. Je présente d'abord en quelques mots la conception minimale de la causalité de Woodward, dans le but de montrer en quoi elle est non-réductive.

(2.1.1) Un facteur A est une cause du facteur B et figure correctement dans une explication causale de B si deux conditions sont satisfaites. :

- (i) une variation de A est ou serait associée à une variation de B (autrement dit, il y a une dépendance entre A et B); et
- (ii) la relation de dépendance entre A et B est ou serait maintenue dans une procédure de test idéalisée, consistant principalement en une variation contrôlée du facteur A, que Woodward appelle une «intervention».

La notion de test ou d'intervention sur le facteur A est elle-même causale, et cela fait en sorte que la conception proposée est non-réductive. Woodward ne considère pas que le caractère non-réductif de cette conception doit être défendu à tout prix. Il accueillerait favorablement toute conception réductive qui serait plus adéquate que la sienne. Le test ultime des diverses conceptions demeure néanmoins leur capacité à rendre compte des intuitions causales bien fondées des humains dans la vie courante et dans les activités scientifiques. Si par exemple une conception réductive comme la conception mécaniste de Salmon classifie comme non-causaux les phénomènes de gravitation, parce qu'ils font appel à de l'action à distance, alors que la conception non-réductive de Woodward les inclut parmi les phénomènes causaux, le choix entre les deux conceptions ne devrait pas dépendre du caractère réductif ou non-réductif des conceptions visées.<sup>4</sup>

---

<sup>3</sup> Woodward, J. (2010), p. 288

<sup>4</sup> Woodward, J. (2003), p. 20-21



Un deuxième argument général pour justifier les conceptions non-réductives est la constatation que, à l'intérieur de la famille des notions causales comme «cause», «loi de la nature», «explication», «intervention», certaines notions peuvent être plus fondamentales que d'autres. Le philosophe fait face à des choix non-évidents quand il tente de conceptualiser les relations entre ces notions. Woodward soutient par exemple que :

«the counterfactuals on which the philosophical tradition has tended to focus in elucidating these [causal] notions are the wrong counterfactuals for this purpose».

Le choix des bons concepts et de leurs relations ne semble pas dépendre en première analyse du caractère réductif ou non-réductif de la conception générale de la causalité. Troisièmement, la circularité qui semble être présente dans la conception de Woodward n'invalide pas cette conception. Pour élucider l'assertion que le facteur  $C_1$  est une cause du facteur  $E_1$ , on peut légitimement faire appel au savoir antérieur que le facteur  $C_2$  est une cause (en vertu d'une relation causale différente) du facteur  $C_1$ . La conception de la causalité proposée possède cette structure. On élucide la relation causale (possible) entre un facteur A et un facteur B en faisant appel à une autre relation causale déjà connue entre un facteur I (pour intervention) et le facteur A.<sup>5</sup>

Le projet de présenter une caractérisation conceptuelle non-réductive de la causalité et de l'explication causale ne se limite pas à une simple description de l'usage des termes causaux. Woodward se propose d'examiner aussi les pratiques d'inférence (causale) qui ont cours dans la vie courante et dans l'activité scientifique. Il annonce qu'il veut effectuer des distinctions et classifications supplémentaires à l'intérieur de la classe générale des énoncés causaux. Il se préoccupe des objectifs visés par les utilisateurs des énoncés causaux. Son projet a une composante normative importante : il présente des significations recommandées des termes causaux; il s'intéresse à la cohérence entre le concept causal retenu et les buts poursuivis; il examine la capacité des utilisateurs à vérifier les assertions causales.<sup>6</sup>

## 2) Trois illustrations préliminaires

Je voudrais maintenant commenter préliminairement la conception minimale de Woodward (que j'appellerai à l'occasion «conception interventionniste», à cause du rôle qu'y jouent les

---

<sup>5</sup> Woodward, J. (2003), p. 21-22

<sup>6</sup> Woodward, J. (2003), p. 7-9

interventions) et en donner trois illustrations<sup>7</sup>. Ces commentaires et ces illustrations visent surtout à attirer l'attention sur des intuitions importantes qui situent, et justifient en partie, la conception interventionniste.

Premièrement, il y a lieu de préciser un peu le sens des termes «explication» et «explication causale» que Woodward cherche à conceptualiser. (Ces précisions sont particulièrement importantes dans le cas d'une conception non-réductive comme la sienne. Puisque la caractérisation des causes et des relations causales contiendra des termes de concepts eux-mêmes causaux, il semble nécessaire de décrire préliminairement la famille des notions causales, ne serait-ce que pour identifier les termes qui, dans la caractérisation ultérieure plus précise, feront en sorte que cette caractérisation sera non-réductive.) Le verbe «expliquer» en langue anglaise (et en français) peut signifier «décrire en détail» comme dans «expliquer comment cuire une tarte aux pommes», ou il peut aussi signifier «justifier» comme lorsqu'on explique pourquoi on a décidé de faire telle chose. Woodward retient un troisième sens plus restreint. Expliquer consiste en l'activité d'élucider pourquoi un résultat particulier, ou un pattern général de résultats, ont lieu. Un élément distinctif de ces explications, selon Woodward :

«is that they show how what is explained depends on other, distinct factors, where the dependence in question has to do with some relationship that holds as a matter of empirical fact, rather than for logical or conceptual reasons»<sup>8</sup>.

Ainsi, l'explication en mathématiques ne fait pas partie du domaine visé. Cette caractérisation a aussi pour but d'exclure du domaine visé les activités purement descriptives. Selon Woodward, les scientifiques distinguent couramment les activités descriptives des activités d'explication. Ainsi en biologie (en botanique), l'activité de recueillir des spécimens de plantes et de les classifier en genres et en espèces, est une activité descriptive. Par contre, la biologie moléculaire est une des nouvelles branches de la biologie qui serait principalement explicative au sens visé ici. On y emploie des techniques qui permettent de relier des variations de certains facteurs à des variations de certains résultats.<sup>9</sup>

---

<sup>7</sup> Les illustrations se trouvent dans Woodward J. (2003), p. 12-17, et p. 312-313

<sup>8</sup> Woodward, J. (2003), p. 4-5

<sup>9</sup> Woodward, J. (2003), p. 4 et p. 9

Woodward circonscrit également le sens de l'expression «explication causale» visée dans son projet. Plusieurs penseurs caractérisent les causes et les explications de manière telle qu'il y a de la place pour des explications non-causales. Woodward, pour sa part, minimise la différence entre les explications en général, telles que caractérisées plus haut, et les explications causales. Il retient une acception large de la notion de causalité en vertu de laquelle toute explication qui fonctionne en montrant qu'un résultat dépend d'autres facteurs (où la dépendance en question n'est pas logique ou conceptuelle), est considérée comme une explication causale. Il suggère de plus qu'une conséquence de cette conception de l'explication causale est que les informations qui y apparaissent sont potentiellement utilisables par les humains aux fins de manipuler (les facteurs causaux) et de contrôler les résultats. C'est pourquoi il réfère parfois à sa conception en utilisant l'expression 'the manipulability conception (of causal explanation)' que je traduis un peu maladroitement par «conception manipulative»<sup>10</sup>.

Les remarques du paragraphe précédent sur l'utilisation possible des causes pour manipuler les résultats sont très importantes. Si la conception interventionniste ou manipulative s'avère acceptable, au sens où elle rendrait compte de la plupart de nos intuitions bien fondées sur la causalité (ce qui reste à démontrer), cette conception bénéficiera alors d'une justification majeure. Elle sera en accord avec le sens commun qui considère que les causes sont avant tout des leviers pour affecter les résultats. La recherche des causes chez les humains sert l'objectif pratique de contrôler la réalité. La conception sera aussi en accord avec la distinction courante, en sciences, entre les activités descriptives et les activités explicatives, à laquelle j'ai fait brièvement allusion plus haut en comparant les activités descriptives de la botanique et les activités causalement explicatives de la biologie moléculaire. Enfin, la conception manipulative se rattache aisément à ce qu'on appelle la méthode expérimentale en sciences. Expérimenter en laboratoire, c'est essentiellement examiner un phénomène dans des conditions qui permettent d'isoler les facteurs causaux potentiels les uns des autres et d'observer l'effet individuel de chacun d'entre eux sur un certain résultat.

Woodward souligne que la justification pratique rattachée à la conception manipulative se distingue des objectifs rattachés par les philosophes à certaines autres conceptions de l'explication

---

<sup>10</sup> Woodward, J. (2003), p. 9

causale. Certains justifient la notion de causalité en lui attribuant l'objectif de faire des prédictions, ou encore un objectif d'unification des théories. Je n'entrerai pas dans le détail de ces autres justifications, mais je voudrais simplement indiquer la différence importante qui existe entre prédire un résultat et manipuler ou contrôler la réalité. Cette différence se rattache souvent à la distinction entre (i) des phénomènes où les facteurs sont simplement corrélés et (ii) des phénomènes où un facteur constitue la véritable cause d'un autre facteur. (Deux facteurs d'un système donné sont dits «corrélés» s'il existe une dépendance entre eux. Les statisticiens ont développé des mesures de corrélation entre des facteurs. Ces mesures exigent que les facteurs soient représentés par des variables portant des valeurs numériques. Lorsque des facteurs sont simplement (au sens de seulement) corrélés, ils dépendent l'un de l'autre, mais cette relation n'est pas causale.) La connaissance d'une simple corrélation permet de faire des prédictions, tandis que la présence d'une véritable relation causale permet la manipulation et le contrôle.

On peut illustrer cela avec l'exemple du baromètre (qui est la première des trois illustrations préliminaires de cette section).

(2.1.2) Le niveau (N) du baromètre est corrélé avec la présence d'une tempête (T).

Il y a une dépendance entre N et T. Lorsque le niveau du baromètre est bas, il y a du mauvais temps; lorsque le niveau du baromètre est haut, il y a du beau temps. Mais en présence d'une telle dépendance observée, la question se pose : s'agit-il d'une dépendance causale ? La conception interventionniste (2.1.1) suggère qu'on doit tester la relation entre le niveau du baromètre et la présence d'une tempête au moyen d'une intervention. Nous verrons plus loin les détails de cette notion, mais l'idée générale peut se comprendre comme suit : on part de l'intuition que les phénomènes du monde réel (au sens de 'actual world') sont influencés par une foule de facteurs inter-reliés. On présuppose également qu'une grande partie des influences agissent de façon répétable.

(2.1.3) Caractérisation qualitative d'une intervention

Une intervention sur un facteur (possiblement causal) (F) à l'égard d'un facteur-résultat (R) change la valeur de F de telle manière que si la valeur de R change par la suite, cela n'a lieu qu'en vertu du changement de F, et pas en vertu du changement d'un facteur autre que F. Un peu plus précisément, l'identification du statut causal d'un facteur donné (F) à l'égard d'un résultat (R) s'effectue (via une expérimentation réelle ou idéalisée) (1) en brisant les

liens en vertu desquels F serait causé par d'autres facteurs du phénomène étudié, (2) en faisant ensuite varier artificiellement la valeur de F (c'est l'intervention), (3) en tenant les autres facteurs fixes sauf si leur valeur est affectée par la variation artificielle de F, et (4) en observant la réaction du facteur R. Selon la conception interventionniste, une relation de dépendance générale (au sens de 'type') entre F et R est causale s'il y a au moins une combinaison des valeurs des autres facteurs telle qu'une intervention sur F entraîne une variation de R.

Je précise que même si la description intuitive que je viens de donner fait référence à des activités humaines, Woodward définit plus loin l'intervention sans référence à des activités humaines.

Lorsque l'on cherche à déterminer si le niveau du baromètre est une cause de la présence d'une tempête, étant donné la dépendance observée entre ce niveau et la présence d'une tempête, on brise les liens de causalité entre le niveau du baromètre et les autres facteurs du phénomène qui pourraient l'influencer. On fait artificiellement varier le niveau du baromètre en tenant fixes les autres facteurs connus (dont notamment la pression atmosphérique), sauf si ces facteurs sont affectés par la variation artificielle du niveau du baromètre. On observe alors, sans surprise, que les variations sous intervention du niveau du baromètre ne sont pas associées à des changements de la situation climatique. Le niveau du baromètre n'est donc pas une cause de la présence de tempêtes, selon la conception interventionniste, ce qui correspond au jugement du sens commun.

L'exemple du baromètre constitue une application négative de la conception interventionniste. Il montre comment cette conception identifie un facteur qui est corrélé à un résultat, sans en être une cause. Je présente maintenant, à titre de seconde illustration préliminaire, un exemple «positif» de la conception interventionniste, provenant lui aussi de Woodward<sup>11</sup>. Dans cet exemple, les facteurs corrélés à un résultat sont aussi des causes de ce résultat. Ce deuxième exemple se distingue également par le fait que la relation causale est représentée explicitement par une équation mathématique. À cause de cela, cet exemple n'est pas une illustration de la conception minimale de la causalité, qui se caractérise par l'absence de description précise de la relation causale, mais constitue plutôt un exemple de la conception renforcée de la causalité. Mais cet exemple est utile parce que, grâce à la présence de la description précise de la relation causale,

---

<sup>11</sup> Woodward, J. (2003), p. 12-15

les notions d'intervention et d'indépendance des mécanismes (que nous discuterons plus loin), sont bien élucidées.

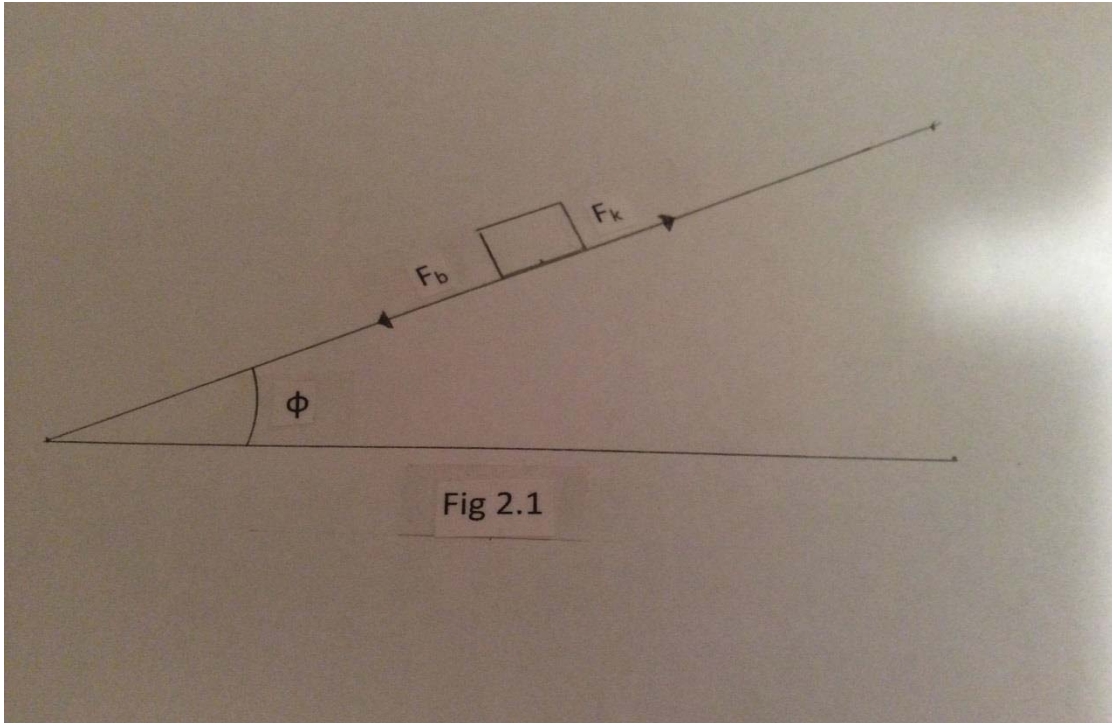


Fig 2.1

(2.1.4) Un bloc est en mouvement sur un plan incliné selon un angle  $\phi$  par rapport à l'horizontale (voir figure 2.1). Le mouvement du bloc est affecté par deux forces, et aussi par la contrainte que constitue le plan incliné (le plan empêche le bloc de tomber verticalement au sol). La force nette  $F_{\text{net}}$  est la somme de (1)  $F_b$  la force de gravitation tirant le bloc vers bas le long du plan, et (2)  $F_k$  la force de friction entre le bloc et la surface du plan, qui s'oppose au mouvement du bloc vers le bas.

$$(2.1.5) F_{\text{net}} = F_b - F_k$$

Je donne sans la justifier la relation mathématique exprimant  $F_{\text{net}}$ .<sup>12</sup>

$$(2.1.6) F_{\text{net}} = mg \sin \phi - \mu_k mg \cos \phi$$

(où  $m$  est la masse du bloc,  $g$  est l'accélération subie par une masse unitaire à la surface de la Terre,  $\phi$  est l'angle (inférieur à 90 degrés) du plan par rapport à l'horizontale,  $\sin \phi$  et  $\cos \phi$  sont des fonctions de  $\phi$  prenant des valeurs entre 0 et 1 selon la valeur de  $\phi$ , et  $\mu_k$  est une constante représentant le taux de frottement)

<sup>12</sup> Voir Woodward, J. (2003), p. 13, pour cette justification

Éliminant la masse en vertu du principe  $F = ma$ , l'accélération «a» du bloc est donnée par :

$$(2.1.7) \quad a = g \sin \phi - \mu_k g \cos \phi$$

Dans le système de ce bloc sur un plan incliné, le mouvement du bloc (i.e. son accélération «a») est corrélé aux divers facteurs qui apparaissent dans l'équation (2.1.7). Cela signifie que le mouvement du bloc dépend de la valeur des divers facteurs :  $g$  (l'attraction gravitationnelle de la Terre au niveau du sol),  $\phi$  (l'angle du plan par rapport à l'horizontale), et  $\mu_k$  (le facteur de frottement). Mais les facteurs représentés par  $g$ ,  $\phi$  et  $\mu_k$  sont-ils des causes du mouvement du bloc ? Selon la conception interventionniste, ces facteurs constituent des causes du mouvement, et l'équation (2.1.7) est une relation causale, si on peut faire varier les facteurs  $g$ ,  $\phi$  et  $\mu_k$  en respectant les normes d'une intervention. Dans le cas du facteur  $\phi$  par exemple, on peut affirmer que  $\phi$  est effectivement une cause du mouvement du bloc parce qu'on peut varier l'angle  $\phi$  du plan en respectant les normes de l'intervention (i.e. en changeant artificiellement  $\phi$  pour briser tout lien causal par lequel  $\phi$  aurait pu être déterminé par  $g$  ou  $\mu_k$ , et en observant l'effet du changement «isolé» de  $\phi$  sur le mouvement du bloc). Puisque l'accélération du mouvement du bloc respecte toujours l'équation (2.1.7),  $\phi$  est bien une cause du mouvement du bloc.

En plus de montrer comment le statut causal de certains facteurs dépend du résultat d'un test par intervention, l'exemple du bloc illustre certains autres aspects fréquemment présents dans les situations causales. Premièrement, le monde réel est complexe, et un modèle quantifié comme les équations (2.1.5) à (2.1.7) constitue une simplification du monde réel. Dans la situation du bloc, on exclut certains facteurs causaux réels mais secondaires : la résistance au mouvement due à la résistance de l'air, la force électromagnétique.

Deuxièmement, la formulation mathématique d'une relation causale ne garantit pas que les facteurs qui y apparaissent soient causaux. Il faut, pour qu'ils soient causaux, qu'elle satisfasse la condition de l'invariance sous intervention, qui sera discutée plus avant au chapitre 3.

Dans une troisième illustration, je cherche à donner une idée de ce que peut être une situation de causalité (épistémiquement) minimale, où il existe une relation de cause à effet mais où l'on ne connaît pas la relation précise entre cette cause et cet effet.

(2.1.8) Il s'agit de la relation entre l'habitude de fumer et le cancer du poumon, telle qu'on l'appréhendait en 1957, avant que l'on ne dispose d'informations détaillées sur le mécanisme biochimique qui relie l'habitude de fumer au cancer du poumon. Woodward tire cet exemple d'une étude publiée par Cornfield et alii, en 1959, où les auteurs affirment que fumer est une cause du cancer, sans tenter de décrire le mécanisme de fonctionnement de la maladie, ou de donner une généralisation chiffrée détaillée reliant la cause supposée et son effet.

Ce que les auteurs donnent à l'appui de leur thèse consiste plutôt en une série de corrélations associant, sous divers angles, l'habitude de fumer à la hausse de la prévalence de cancer. La prévalence du cancer du poumon est plus élevée chez les fumeurs que chez les non-fumeurs pour toutes sortes de sous-groupes : chez les hommes et chez les femmes, dans différents groupes ethniques, chez des personnes ayant différentes habitudes alimentaires, dans différents milieux socio-économiques. De plus d'autres données montrent un pattern d'influence : plus la consommation quotidienne de tabac augmente, plus l'incidence du cancer est élevée; les personnes qui fument depuis plus longtemps ont une incidence plus élevée; etc.; ces patterns tiennent quel que soit le groupe social des sujets.

L'accumulation de ces corrélations positives, dans différents contextes, suffisait selon les auteurs et selon Woodward, à démontrer l'existence d'un lien causal, malgré l'absence d'une description précise du mécanisme ou de la relation rattachant l'habitude de fumer au cancer.<sup>13</sup>

## 2.2 Les concepts causaux minimaux chez Woodward

### 1) La formalisation

Le but de cette section 2.2 est de caractériser les concepts causaux minimaux dans la conception manipulative. Avant cela, il me paraît utile de décrire et de justifier au moins partiellement l'appareil formel que Woodward emploie : (i) dans ses caractérisations de concepts causaux (dans les versions «minimale» et «renforcée») et (ii) au besoin, dans ses descriptions de situations causales. Voici quelques éléments de cette formalisation.

---

<sup>13</sup> Woodward, J. (2003), p. 312-313



1. Les termes (les causes et les effets) de la relation causale sont représentés comme des variables (plutôt que d'être considérés comme des objets ou des types d'objets, ou encore comme des événements ou des types d'événements, comme le conçoit le sens commun, et comme c'est le cas dans plusieurs autres conceptions de la causalité).
2. Les relations entre les termes sont, si elles sont connues et si c'est utile, exprimées sous forme d'équations.
3. Certains aspects des relations entre les termes sont aussi représentés au moyen de graphes orientés ('directed graph').

Je commente d'abord l'emploi des variables. Il est clair que, dans la conception manipulative comme dans les autres, les termes réels de la relation causale sont des «choses», au sens le plus large. Ces choses peuvent être des objets physiques, ou des entités non matérielles (comme une monnaie), ou des événements. Ces choses sont des porteurs de propriétés. (Je laisse ici de côté les situations plus complexes de classes de choses qui seraient des porteuses de propriétés). Dans la conception manipulative toutefois, ces choses doivent être «manipulables» (au moins en théorie) au sens où elles (ou leurs propriétés) doivent être susceptibles de passer d'un état à un autre. On peut attribuer des valeurs numériques à ces états. Nous avons donc, dans le monde réel, des choses individuelles, leurs propriétés et les états de ces choses ou propriétés, et dans le domaine de la représentation, les variables qui sont des étiquettes représentant une chose ou une de ses propriétés, et les valeurs des variables, qui représentent les différents états possibles de cette chose ou de l'une de ses propriétés. La raison première de l'emploi du formalisme «variables, valeurs de variables» est de pouvoir bénéficier des concepts statistiques dans la caractérisation des concepts causaux. Ainsi, le concept causal d'intervention fera appel à la notion de non-corrélation de certaines variables (que je crois proche de la notion d'indépendance statistique). Également à titre d'exemple, plusieurs des concepts de cause que nous verrons plus loin feront appel à l'idée d'une intervention sur une variable (causale)  $X$  qui change la valeur d'une variable (effet)  $Y$ , ou la distribution de probabilité de  $Y$ .

Je passe maintenant à des commentaires et justifications de l'emploi d'équations. Il est d'abord évident et trivial (1) que l'emploi de variables pour représenter les termes de la relation (causale ou non) entre ces variables, se rattache facilement à l'emploi d'équations pour représenter les

relations entre ces variables, et (2) que l'emploi du formalisme variables/équations permet d'appliquer des techniques mathématiques à l'étude de la causalité. Mais il y a de nombreuses formes d'équation, et Woodward, suivant en cela les recherches de J. Pearl<sup>14</sup>, retient une forme particulière d'équation pour représenter la plupart des relations entre les variables dans un contexte de causalité. Woodward suggère que les situations causales peuvent être représentées dans un modèle causal fonctionnel qui est un système d'équations de forme :

$$X_i = F(Pa_{i1}, Pa_{i2}, \dots, Pa_{in}, U_i)$$

où un modèle comprend habituellement plus d'une équation, et

où, dans chaque équation,  $X_i$  représente un facteur effet résultant d'autres facteurs (possiblement causaux) antérieurs qui sont explicitement mentionnés dans le modèle (les divers  $Pa_{i1\dots in}$ ), et  $U_i$  est une variable d'erreur qui représente l'impact combiné de tous les facteurs qui sont des causes de  $X_i$  mais qui ne sont pas mentionnés dans le modèle.<sup>15</sup>

Je n'ai pas les compétences requises pour commenter tous les aspects de cette formalisation. Je retiens toutefois trois aspects. Premièrement, cette formalisation intègre l'idée qu'il y a un ordre dans les dépendances entre facteurs. La relation causale est asymétrique temporellement au sens où un facteur «cause» doit soit précéder, soit être simultanément à son «effet»<sup>16</sup>. Cela apparaît dans la formalisation via : (1) la convention selon laquelle le facteur effet ( $X_i$ ) apparaît du côté gauche de l'équation (et pas du côté droit); les facteurs se trouvant sur le côté droit se manifestent donc avant ou en même temps que le facteur effet. Et via (2) le fait que les facteurs  $Pa_{i1}\dots Pa_{in}$  soient aussi désignés comme les «parents» ou les «causes directes» du facteur effet ( $X_i$ ). Une cause directe d'un effet est un facteur qui influence cet effet sans passer par un facteur intermédiaire. Nous verrons que Woodward propose une caractérisation distincte des causes directes.

Un deuxième aspect que je ne fais que mentionner est celui de l'interprétation des équations (lorsque celles-ci sont disponibles). Selon Pearl, interprété par Woodward,

---

<sup>14</sup> Pearl, J. (2000)

<sup>15</sup> Woodward, J. (2009), p. 243

<sup>16</sup> Je reconnais que l'idée d'une asymétrie temporelle de la relation causale est controversée. Je ne l'aborde pas dans ce mémoire, et je me contente d'accepter les intuitions du sens commun à cet égard.

«each equation [of a functional causal model should] represent a distinct ‘causal mechanism’ which is understood to be ‘autonomous’ in the sense in which that notion is used in econometrics; this means roughly that it is possible to interfere with or disrupt each mechanism (and the corresponding equation) without disrupting any of the others»<sup>17</sup>.

L’idée qu’il soit possible de modifier une relation particulière d’un système causal sans affecter les autres relations est fondamentale en sciences et aussi dans la conception interventionniste de la causalité. Mais cette supposition de l’autonomie des mécanismes est fortement mise en doute par divers penseurs, dont Nancy Cartwright.

Un dernier aspect digne de mention est que la formalisation par l’emploi de variables et d’équations est très lourde. Woodward lui-même ne l’utilise qu’au besoin, lorsque nos intuitions nous égarent ou sont insuffisantes pour éclaircir une situation. Woodward souligne par contre qu’elle s’applique largement, y compris dans les situations de causalité singulière étudiées par les philosophes qui favorisent d’autres formes d’analyse contrefactuelles de la causalité<sup>18</sup>. Une situation de dépendance contrefactuelle entre des événements peut facilement être représentée par des équations ou des relations simples reliant des variables binaires correspondant à l’occurrence ou à la non-occurrence de ces événements. Par exemple, l’assertion que les courts-circuits causent des incendies peut être représentée par une relation entre deux variables C et I, dont chacune prend deux valeurs. C prend les valeurs 1 ou 0 selon qu’un court-circuit a ou n’a pas lieu, et I prend les valeurs 1 ou 0 selon qu’un incendie a ou n’a pas lieu<sup>19</sup>.

Je commente en dernier lieu l’emploi de graphes orientés. Ces graphes ont pour but de représenter visuellement une partie de l’information sur les relations asymétriques entre les facteurs causaux d’un système représenté par des équations. Il s’agit de montrer quel facteur est une cause directe de quel autre facteur dans le système. (La notion de cause directe sera caractérisée plus loin, mais on peut la comprendre intuitivement comme suit : un facteur A est une cause directe de B si aucun autre facteur n’intervient entre A et B pour influencer B.) Un graphe orienté comprend deux types d’objets : (1) des sommets, où chaque sommet correspond à une des variables (correspondant à un

---

<sup>17</sup> Woodward, J. (2009), p. 243-244

<sup>18</sup> Il s’agit notamment des philosophes qui s’inscrivent dans la tradition de recherche lancée par David Lewis. Voir par exemple Paul, L. A. et Hall, N. (2013).

<sup>19</sup> Woodward, J. (2009), p. 248

des facteurs) des équations décrivant le système. (2) des flèches reliant certains sommets. Chaque flèche correspond à une relation de causalité directe entre les sommets / variables / facteurs qu'elle réunit. (S'il n'y a pas de flèche entre deux sommets, cela signifie qu'il n'y a pas de lien de causalité directe entre les deux facteurs correspondants.) Voici un exemple illustratif :<sup>20</sup> un système est représenté par deux équations linéaires.

$$(2.2.1) Y = aX + cZ$$

$$(2.2.2) Z = bX$$

La convention dans les équations est que les effets se trouvent à gauche et les causes à droite. Les majuscules X, Y, et Z sont les variables (i.e. les facteurs causaux) tandis que les minuscules a, b, et c représentent des constantes, possiblement des nombres.

Le graphe orienté correspondant au système décrit par ces deux équations est le suivant.

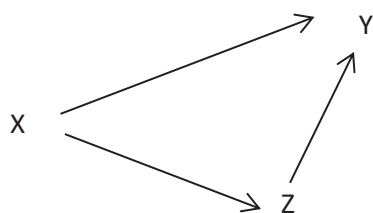


Fig. 2.2 Un graphe orienté

Chacune des variables présentes dans les équations apparaît comme un sommet. Le graphe (contrairement aux équations) montre les causes à gauche et les effets à la droite de leurs causes. X est la fois une cause directe de Z et de Y. C'est pourquoi deux flèches partent de X, une vers Z et l'autre vers Y. Y pour sa part est influencé directement par X et par Z (en vertu de l'équation (2.2.1)). C'est pourquoi Y reçoit deux flèches. X, à titre de facteur causal direct de Y et de Z, peut être appelé un «parent» de Y et Z, et Y et Z peuvent être appelés des «descendants» de X.

Lorsque deux variables sont reliées entre elles par une seule ou plusieurs flèches, on dit qu'elles sont reliées entre elles par un chemin orienté ('directed path'). Woodward va utiliser la notion de chemin orienté dans certaines de ses caractérisations causales.

---

<sup>20</sup> Exemple tiré de Woodward, J. (2009), p. 249

Il est important de noter que les graphes orientés peuvent être trompeurs, en ce sens qu'ils ne font que montrer s'il existe ou non des liens de causalité directe entre des variables, mais sans distinguer entre les types de relation causale directe. Voici une version simplifiée des exemples proposés par Woodward.

Exemple 1 : Un effet Z est fonction de deux facteurs X et Y selon une relation additive exprimée par l'équation :

$$Z = aX + bY, \text{ où } a \text{ et } b \text{ sont des constantes.}$$

En vertu de la relation d'addition entre les deux facteurs X et Y, l'effet «marginal» sur Z de chacun des facteurs X et Y est indépendant de la valeur prise par l'autre facteur. Notamment, si Y prend la valeur 0, cela n'empêchera pas qu'une variation dx de X ait un effet adx sur Z.

Exemple 2 : Un effet Z est fonction de deux facteurs X et Y selon une relation multiplicative exprimée par l'équation :

$$Z = aX*Y, \text{ où } a \text{ est une constante.}$$

Dans ce cas de relation multiplicative entre les facteurs X et Y, l'effet «marginal» sur Z de chacun des facteurs X et Y n'est pas indépendant de la valeur prise par l'autre facteur. Si Y prend la valeur 0, l'effet sur Z d'une variation dx de X sera nul.

Mais les équations des exemples 1 et 2, et les relations causales associées, sont représentées par le même graphe orienté (figure 2.3).

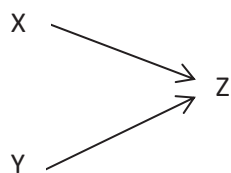


Fig. 2.3 Deux relations différentes ont le même graphe orienté

Les graphes orientés indiquent quelles variables sont en relation directe avec d'autres variables, mais ne disent rien du type de relation causale (additive ou multiplicative dans nos exemples) qui prévaut.<sup>21</sup>

## 2) Les concepts causaux : cause totale, cause directe, cause contributoire et intervention

L'intuition de départ dans la conception manipulative est qu'un facteur X est une cause du facteur Y dans une situation générale donnée si et seulement si une variation appropriée de X a un effet non nul sur la valeur de Y (ou sur la probabilité que Y prenne cette valeur). Cette intuition de départ aurait besoin d'être précisée à plusieurs égards. Que signifie, par exemple, «variation appropriée»? J'y reviendrai dans un prochain paragraphe. Mais l'intuition initiale a besoin d'être immédiatement corrigée. Il faut distinguer entre, d'une part, l'effet total du facteur X sur le facteur Y et, d'autre part, la contribution du facteur X au résultat en Y.

Un facteur X peut en effet n'avoir aucun effet net sur le résultat Y même s'il affecte ce résultat. Woodward donne l'exemple suivant visant un système représenté par les équations déjà employées (2.2.1) et (2.2.2) et la figure 2.2 :

$$(2.2.1) Y = aX + cZ$$

$$(2.2.2) Z = bX$$

Si le paramètre «a» prend la valeur  $-bc$ , alors l'influence causale directe de X (via le chemin direct de X vers Y) sera annulée par l'influence indirecte de X sur Y (sur le chemin XZY). Dans ce cas, aucune variation de la variable X seule n'aura d'effet sur la valeur de Y, et X n'est donc pas une cause totale de Y. On peut néanmoins penser que X a, en un sens à préciser, une influence causale sur Y.

Pour faire face à ce contre-exemple, la conception manipulative distingue trois concepts de cause : (i) cause totale, (ii) cause directe, et (iii) cause contributoire. Le facteur X est une cause totale (générale) du facteur Y si et seulement si X a un effet total non-nul sur Y. Plus précisément :

---

<sup>21</sup> Woodward, J. (2003), p. 44-45

### (2.2.3) Cause totale ou «TC»

[A variable] X is a total cause of [a variable] Y if and only if there is some intervention on X alone (and no other variables) such that for some value of other variables besides X, this intervention on X will change the value of Y<sup>22</sup>.

Le concept de cause totale s'oppose à la notion de «cause contributoire» qui vise à refléter l'intuition qu'un facteur X a une influence causale sur un facteur Y même si, par un effet d'annulation, X n'a pas d'effet total sur Y.

La caractérisation (2.2.3) inclut une référence à la notion d'intervention. Je donne quelques précisions sur l'objectif et le sens intuitif de cette notion, avant de la caractériser plus techniquement plus loin. En gros, ce ne sont pas tous les changements d'une variable X qui sont appropriés pour établir son statut causal à l'égard d'une autre variable Y. L'objectif de la notion d'intervention est donc de caractériser les variations appropriées pour établir ce statut causal. Une intervention sur X à l'égard de Y (dans un système donné) est un processus causal exogène (relativement au système) qui change X de telle manière que si un changement de Y se produit (à la suite du changement de X), le changement de Y a lieu uniquement en vertu de la relation entre X et Y.<sup>23</sup> Par exemple, dans le système composé par la pression atmosphérique (P), le niveau d'un baromètre (N) et la présence d'une tempête (T), il est clair que P est une cause commune de N et de T. Par conséquent, une variation de N est généralement accompagnée d'une variation de T et on pourrait penser que N est une cause de T. Une intervention effectuée (ou imaginée) pour déterminer si N est une cause de T, doit être telle qu'elle montrera si un changement de N seul occasionnerait un changement de T. C'est pourquoi le changement du niveau du baromètre doit consister en un processus causal exogène par rapport au système visé. Ainsi une intervention sur N pourrait être l'action d'un agent qui impose manuellement un changement du niveau du baromètre, indépendamment de la pression atmosphérique.

Je reviens maintenant à la caractérisation de la notion de cause contributoire. Woodward caractérise cette notion en deux étapes. Il établit d'abord le concept de cause directe, et caractérise ensuite la cause contributoire en recourant au concept de cause directe.

---

<sup>22</sup> Woodward, J. (2009), p. 249-250

<sup>23</sup> Woodward, J. (2003), p. 47

On a besoin du concept de cause directe pour bien identifier et traiter les situations de chaîne de facteurs causaux. Dans la figure 2.2 apparaissant plus haut, X est une cause directe de Y sur un des deux chemins vers Y, celui où le facteur X influence Y sans aucun intermédiaire, et n'est pas une cause directe de Y sur l'autre chemin, celui où l'influence de X passe par le facteur Z. Il y a dans ce deuxième cas une chaîne causale allant de X à Z et ensuite à Y. Woodward caractérise comme suit le concept de cause directe (que je désigne également avec l'étiquette «DC») :

(2.2.4) Cause directe ou DC

A necessary and sufficient condition for X to be a direct cause of Y with respect to some variable set V is that there be a possible intervention I on X that will change Y (or the probability distribution of Y) when all other variables in V besides X and Y are held fixed at some value by additional interventions that are independent of I.<sup>24</sup>

En vertu de cette caractérisation, X est une cause directe de Y dans le système représenté par les équations (2.2.1) et (2.2.2) (et la figure 2.2) parce que si on fixe la valeur de Z à la constante m (c'est l'intervention qui fixe la valeur de Z, à titre de variable de V autre que X et Y), alors (2.2.2) se lit :

$$(2.2.5) Z = m$$

et (2.2.1), où le symbole  $\Delta x$  représente l'intervention sur X, se lit :

$$(2.2.6) \Delta y = a\Delta x + cm$$

Ainsi, dans ce cas, la plupart des interventions  $\Delta x$ , accompagnées d'une intervention indépendante fixant la valeur de Z, produisent une variation  $\Delta y$  de Y, et X est une cause directe de Y.

Par contre, cette caractérisation fait aussi en sorte que dans un système représenté par les équations et le graphe suivant, où la variable X n'affecte pas directement Y, X n'est pas une cause directe de Y.

$$(2.2.7) W = aX$$

$$(2.2.8) Y = bW$$



Fig. 2.4 X n'est pas une cause directe de Y

---

<sup>24</sup> Woodward, J. (2009), p. 250



Si on fixe de façon exogène la valeur de  $W$  à la constante  $n$ , alors aucune variation  $\Delta x$  de  $X$  ne produira de variation de la valeur de  $Y$ .

Voici deux remarques rapides qui visent à clarifier la caractérisation de la notion de la cause directe (2.2.4). Premièrement, le statut de cause directe d'un facteur (i.e. d'une variable), dans une situation, est relatif à l'ensemble particulier de facteurs (de variables)  $V$  employé pour décrire la situation, ou, dit autrement, au niveau de détail de la description. On peut décrire un tir d'arme à feu en disant (description 1) que l'agent  $A$  a tiré avec une arme à feu et qu'une balle a atteint la personne  $B$ , ou encore (description 2) en disant que l'agent  $A$  a exercé une pression sur la gâchette, qui a déplacé le percuteur, qui a produit une explosion dans une cartouche, qui a entraîné le mouvement de la balle, etc. La cause directe du mouvement de la balle diffère donc selon la description employée. Mais cette relativité n'implique pas, dans tous les cas, que le statut de cause directe soit subjectif. Ce n'est que si le statut de cause directe dépendait des désirs et des états mentaux des personnes qui examinent une situation, plutôt que de simplement dépendre du niveau de détail de la description retenue, que le statut de cause directe deviendrait subjectif. D'autre part, la caractérisation de cause contributoire que je présenterai plus loin sera telle que, pourvu qu'il y ait un seul chemin orienté de  $X$  vers  $Y$ , si  $X$  est une cause contributoire de  $Y$ , alors  $X$  demeurera une cause contributoire de  $Y$  si des facteurs (des variables) additionnels  $Z$  sont insérés entre  $X$  et  $Y$  le long de ce chemin. Ainsi le statut de cause contributoire de  $X$  à l'égard de  $Y$ , sous cette condition de route unique entre  $X$  et  $Y$ , n'est pas affecté par le choix d'un niveau de description plus raffiné.<sup>25</sup>

Ma deuxième remarque vise à attirer l'attention du lecteur sur le fait que la caractérisation de la cause directe fait appel non pas à une mais à plusieurs interventions. Il y a non seulement une intervention principale  $I$  sur  $X$ , mais il y a des interventions sur les autres variables du système visé, qui doivent être maintenues fixes.

Je présente maintenant la caractérisation du concept de cause contributoire, qui est celui que Woodward veut ultimement substituer à celui de cause totale pour mieux rendre compte des

---

<sup>25</sup> Woodward, J. (2003), p. 55-56

phénomènes causaux. Woodward emploie le concept de cause directe pour formuler une condition suffisante et nécessaire du fait que X soit une cause contributoire de Y. Je désigne la caractérisation de la causalité contributoire générale au moyen de l'étiquette «M» pour conception manipulative.

#### (2.2.9) Cause contributoire ou M

A necessary and sufficient condition for X to be a (type level) contributing cause of Y with respect to a variable set V is that (i) there be a directed path from X to Y such that each link on this path is a direct causal relationship; that is, a set of variables  $Z_1 \dots Z_n$  such that X is a direct cause of  $Z_1$  which is in turn a direct cause of  $Z_2$  which is a direct cause of ...  $Z_n$  which is a direct cause of Y, and that (ii) there be some intervention on X that will change Y when all other variables in V that are not on this path are fixed at some value.<sup>26</sup>

Voici deux remarques visant à clarifier la caractérisation (2.2.9). Premièrement, le concept de cause contributoire vise à remplacer celui de cause totale dans des situations où un facteur X influe sur un effet Y via des variables intermédiaires (désignées par les étiquettes  $Z_i$ ), ou via plus d'un chemin orienté. Dans la conception manipulative, être une cause, c'est être une cause contributoire.

Ma seconde remarque explicative sur la caractérisation (2.2.9) porte sur le fait que (2.2.9) comprend deux conditions. La condition (i) exige, pour que le facteur X soit une cause contributoire du facteur Y, que X soit une cause directe de Y, ou qu'il y ait une chaîne ininterrompue de facteurs en lien de cause directe entre X et Y. Mais cette condition seule est insuffisante. On peut construire des exemples de chaînes de facteurs où chaque lien est un lien de cause directe, mais où la chaîne totale (i.e. le lien entre le premier et le dernier maillon de la chaîne) ne constitue pas un lien de causalité contributoire. En d'autres mots, le lien de causalité directe n'est pas transitif. C'est pourquoi Woodward ajoute la condition (ii) qui exige qu'il existe au moins une variation par intervention sur X, le premier maillon, qui produise une variation de Y, le dernier maillon, quand les autres facteurs appropriés sont fixés.<sup>27</sup>

Voici un exemple présenté par Woodward d'une suite de deux liens de cause directe dont la combinaison ne constitue pas un lien de cause contributoire. Un agent, droitier, envisage de peut-

---

<sup>26</sup> Woodward, J. (2009), p. 250

<sup>27</sup> Woodward, J. (2003), p. 57-59

être, demain, faire exploser une grenade en tirant sur un déclencheur. Il a trois choix possibles (variable G) : G=1, il tire sur le déclencheur avec son index droit; G=2, il tire sur le déclencheur avec son index gauche; G=3, il décide de ne pas tirer sur le déclencheur. Le résultat de ses choix à l'égard de G est (variable E) : E=1 explosion; E=0 pas d'explosion. Aujourd'hui, un chien mord l'agent à l'index droit. On représente ceci avec la variable M : M=0, pas de morsure, M=1, morsure à l'index droit. La morsure a lieu et le lendemain, l'agent, ne pouvant se servir de son index droit, décide de tirer sur le déclencheur avec son index gauche (G=2). L'explosion a lieu (E=1). Peut-on dire que la morsure a causé l'explosion ? Notre intuition est de répondre que non, car il semble que c'est un facteur autre que la morsure qui a déterminé le choix de l'agent entre les trois options de G et l'explosion.

Dans la perspective des causes directes et contributives, Woodward interprète ce type de situation comme suit. Il y a un lien de cause directe entre M et G. Si on changeait par intervention le fait que le chien a mordu l'index droit de l'agent, pour lui substituer une absence de morsure, alors il y aurait un effet sur la variable G : l'agent utiliserait vraisemblablement son index droit pour tirer sur le déclencheur. Il y a également un lien de cause directe entre G et E. Gardant M constant, il y a une intervention possible sur G qui ferait en sorte que E changerait. En intervenant sur l'agent pour qu'il décide de ne pas tirer sur le déclencheur, alors l'explosion n'aurait pas lieu. Mais la combinaison des deux liens de cause directe ne fait pas en sorte que la morsure soit une cause contributive de l'explosion, parce qu'aucune intervention sur M n'aura pour effet de changer E. Le fait que le chien ait mordu l'index droit de l'agent n'a pas d'impact sur l'explosion. Voici la conclusion de Woodward, adaptée pour la nomenclature des variables :

«The changes in the value of G on which the value of E depends are completely different from the changes in the value of G that are influenced by the value of M, so that there is no overall sensitivity of the value of E to the value of M along the only route connecting M to E. [...] I believe that it is this fact (of insensitivity along the only route connecting M to E) that makes us judge that M does not cause E and that transitivity fails»<sup>28</sup>.

Je reviens maintenant brièvement à la notion d'intervention, que j'avais caractérisée qualitativement plus haut, en disant qu'une intervention sur un facteur X (dont on se demande s'il

---

<sup>28</sup> Woodward, J. (2003), p. 58

est une véritable cause de Y) est un processus causal exogène qui change la valeur de X de telle manière que si un changement de Y se produit à la suite du changement de X, le changement de Y a lieu uniquement en vertu de la relation entre X et Y. Je vais maintenant présenter une caractérisation plus technique du concept d'intervention, développée par Woodward et C. Hitchcock dans le cadre du projet de caractérisation conceptuelle non-réductive des notions causales. Le seul but de cette présentation est de montrer combien le concept d'intervention dépend de l'emploi du formalisme variables, équations (si disponibles) et graphes orientés. Il n'est pas requis, pour le lecteur, de la comprendre de façon détaillée, car je ne l'utiliserai pas dans le reste du mémoire. Je la désigne avec l'étiquette «INT».

#### (2.2.10) Caractérisation technique de l'intervention (INT)

Une intervention I sur une variable X, à l'égard d'une variable Y, est un processus (réel ou hypothétique) de modification de la valeur de la variable X, visant à produire des changements de la valeur de Y, qui doit satisfaire les conditions suivantes :

- «(M1) I must be the only cause of X – that is [...] the intervention must completely disrupt the causal relationship between X and its previous causes so that the value of X is set entirely by I,
- (M2) I must not directly cause Y via a route that does not go through X,
- (M3) I should not itself be caused by any cause that affects Y via a route that does not go through X,
- (M4) I must be probabilistically independent of any cause of Y that does not lie on the causal route connecting X to Y»<sup>29</sup>.

Je note que la caractérisation de l'intervention (2.2.10) emploie la notion de cause directe, mais que cette caractérisation n'est pas vicieusement circulaire, puisque les conditions énoncées ne font pas référence à une relation de causalité directe qui pourrait exister entre X et Y. La caractérisation échappe également à l'anthropocentrisme qu'on rattache souvent à d'autres conceptions manipulatives de la causalité, car le processus décrit dans (2.2.10) pourrait être un processus naturel.<sup>30</sup>

---

<sup>29</sup> Woodward, J. (2009), p. 247

<sup>30</sup> Woodward, J. (2009), p. 247

### 2.3 Une conclusion sur les concepts causaux minimaux

La section 2.2 portait sur les concepts causaux dans la perspective «minimale», décrite à la fin du chapitre 1, selon laquelle il n'est pas nécessaire de connaître la relation détaillée entre les termes d'une relation causale pour en établir l'existence. Même si certains des exemples employés faisaient référence à des relations exprimées par des équations, les caractérisations des concepts de cause totale (2.2.3), de cause directe (2.2.4) et de cause contributive (2.2.9) ne font pas référence à l'énoncé précis de la relation qui pourrait exister entre les termes X et Y. (Cette absence sera plus frappante lorsque je donnerai, au chapitre 3, les caractérisations des concepts causaux renforcés, qui contiennent des références à l'énoncé précis de la relation entre X et Y.) Je veux maintenant revenir sur cette idée et indiquer en quoi les caractérisations des concepts causaux de la section 2.2 me semblent comporter une certaine tension.

Cette tension provient principalement de la notion d'intervention. Alors que la conception minimale vise à caractériser les causes sans imposer de condition préalable sur la relation entre les facteurs, la notion d'intervention nous force à appliquer une exigence supplémentaire aux relations causales dans une situation donnée, pour que l'on puisse y utiliser les méthodes de la conception manipulative. Quand une situation est régie par plusieurs liens ou forces de nature différente (ce qui est le cas de l'immense majorité des situations concrètes), l'intervention I sur une variable X consiste à employer une relation causale exogène pour agir sur X et ainsi tester l'impact de variations contrôlées de X sur l'effet présumé Y. Le processus d'intervention présuppose donc que le mécanisme sous-jacent à l'intervention est indépendant du lien ou du mécanisme qui relie possiblement X et Y. Plus généralement, la conception manipulative suppose que les mécanismes en jeu dans une situation donnée sont indépendants, et peuvent être testés séparément. J'emploie l'expression «hypothèse de séparation des mécanismes» pour désigner cette supposition<sup>31</sup>. Par exemple, dans l'exemple (2.1.4) du bloc glissant sur un plan incliné, il y a trois forces ou mécanismes en jeu. La force gravitationnelle, la force de frottement du bloc sur le plan et la contrainte imposée par l'inclinaison du plan. L'hypothèse de la séparation des mécanismes est respectée dans cet exemple parce que, entre autres, une variation de l'angle du plan n'affecte pas la force de gravitation et, à l'intérieur de certaines limites, n'affecte pas le facteur de frottement.

---

<sup>31</sup> Woodward, J. (2003), p. 52-53

Il est loin d'être évident que l'hypothèse de séparation des mécanismes soit satisfaite dans toutes les situations causales concrètes. L'apport de la conception manipulative des causes est alors de souligner l'importance de conditions de recherche (en laboratoire ou via un plan de recherche) qui permettent de se rapprocher le plus possible de l'idéal de la séparation des mécanismes. Voici deux exemples de situations où la séparation des mécanismes n'est pas immédiatement présente et doit être obtenue de façon approximative. Premièrement, dans les situations de recherche sur les médicaments, on peut penser que les facteurs qui influencent la guérison d'une personne particulière comprennent, entre autres, la prise du médicament et la volonté de guérir de ce patient. Il est clair qu'il est difficile, sinon impossible, de séparer le mécanisme rattaché à la volonté de guérir d'un patient particulier, du mécanisme d'action du médicament. Les techniques de tests sur de grands groupes de patients dont les membres sont choisis au hasard ont pour but de produire des statistiques comparatives de taux de guérison qui éliminent l'effet de la volonté de guérir des patients en répartissant les patients qui ont une forte volonté de guérir de façon égale dans le groupe qui reçoit le médicament et dans le groupe témoin. Deuxièmement, il y a les situations qui comprennent une rétroaction importante. Un facteur X est une cause d'un effet Y, mais le facteur Y influence à son tour le facteur X. L'économie est une discipline où de telles rétroactions sont très présentes. Si la quantité réduite de l'offre d'un bien M affecte à la hausse le prix de ce bien sur le marché, il est également fréquent que le niveau plus élevé du prix affecte à son tour l'offre à moyen terme du bien M dans un marché concurrentiel. Alors que la conception manipulative fonctionne bien dans des situations où on peut distinguer et séparer clairement (i) les causes qui influencent X de (ii) les impacts de X sur la valeur de Y, les situations où la valeur de Y influence à son tour X semblent à première vue problématique pour la conception manipulative.

Je conclus de ces considérations que la conception manipulative des causes examinée dans la section 2.2, en dépit du fait qu'elle ne semble rien imposer aux relations causales en général, contient en réalité une contrainte sur ces relations. Elle suppose que, s'il y a plusieurs mécanismes sous-jacents à un phénomène, ces mécanismes peuvent idéalement fonctionner indépendamment, ou être étudiés dans des conditions qui se rapprochent de l'idéal de la séparation des mécanismes. Or il semble au moins conceptuellement possible qu'il y ait des relations causales où les

mécanismes sous-jacents ne sont pas séparables. Si c'est le cas, alors la conception de Woodward serait plus «révisionniste» qu'il n'y paraît, par rapport à l'usage ordinaire de la notion de cause.

### CHAPITRE 3. LES CONCEPTS CAUSAUX RENFORCÉS

J'ai présenté au chapitre 2 une conception minimale des phénomènes causaux, qui suppose qu'il n'est pas nécessaire de connaître le détail d'une relation causale pour caractériser les causes et affirmer l'existence d'une relation de cause à effet. Mais, dans un grand nombre de cas, les chercheurs et les agents connaissent effectivement suffisamment de détails structurés sur la relation causale pour parvenir à la modéliser (au sens de l'exprimer par des équations). Pour refléter cette situation épistémique favorable, Woodward a développé des concepts causaux renforcés qui permettent de faire des distinctions qui ne sont pas accessibles en l'absence de cette information supplémentaire. Je précise également que les concepts renforcés n'infirment pas les concepts minimaux présentés au chapitre 2, mais en constituent plutôt un approfondissement.

Les concepts causaux renforcés sont fréquemment associés à la notion d'explication, notamment parce que l'explication vise des objectifs comme le contrôle des phénomènes et leur compréhension. Ces objectifs sont plus faciles à atteindre si on dispose d'informations structurées sur la relation causale elle-même.

Le chapitre 3 présente les concepts causaux renforcés de Woodward. Dans la section 3.1, j'explique de façon plus détaillée pourquoi je fais un détour par la notion d'explication causale. Je rattache cette notion à l'objectif principal du mémoire qui est de montrer comment une conception de la causalité fondée sur les lois peut être avantageusement remplacée par une conception fondée sur une conception plus souple : l'invariance. Les sections 3.2 et 3.3 sont consacrées aux modèles d'explication axés sur les lois de la nature, développés par Carl Hempel. Ces modèles et les problèmes auxquels ils font face permettent de présenter de façon contrastée la conception manipulative de l'explication causale (section 3.4), et de l'invariance (section 3.5). La section 3.6 apporte des remarques limitées sur l'emploi des énoncés contrefactuels dans la conception manipulative. La section 3.7 permet de montrer les avantages de la conception manipulative à l'égard des problèmes des conceptions nomologiques identifiés à la section 3.3. Enfin, la section 3.8 répond à une objection majeure qu'on peut faire aux conceptions causales minimales et renforcées. On soutient que les conceptions manipulatives sont trop peu exigeantes au sens où elles admettent comme causales un grand nombre de relations que le sens commun considérerait



comme causalement douteuses. La réponse de Woodward est d'appliquer, aux relations admises comme causales selon ses conceptions, des critères supplémentaires tels que la stabilité et la spécificité qui permettent d'isoler certaines relations qui sont plus susceptibles d'être utilisées aux fins de manipulation et de contrôle.

### 3.1 Un détour par la notion d'explication

Le principal objectif de ce mémoire est de montrer comment une conception de la relation causale et de l'explication scientifique fondée sur les lois de la nature peut être avantageusement remplacée par une conception fondée sur une notion plus souple de la relation entre causes et effets : celle en vertu de laquelle la relation causale entre un facteur A et un facteur B est caractérisée comme une «généralisation invariante sous intervention», au moins dans un domaine limité de cas. Cet objectif semble toutefois complètement étranger aux concepts causaux présentés au chapitre 2, qui n'imposaient pratiquement aucune condition à la relation causale elle-même. C'est plutôt dans le présent chapitre 3 que je pourrai aborder cet objectif portant sur les lois, puisque j'y aborde les concepts renforcés qui s'appliquent quand les agents disposent d'information structurées sur la relation entre les causes et les effets.

Parmi les concepts causaux renforcés, on trouve le concept d'explication scientifique, qui est pertinent à au moins deux égards. Premièrement, sur le plan historique, c'est la série de modèles conceptuels de l'explication scientifique construits principalement par Carl Hempel dans les années 1940 à 1970 qui ont relancé la recherche philosophique récente sur l'explication et la causalité. Les analyses philosophiques subséquentes de l'explication et de la causalité ont souvent été conduites en réaction aux modèles de Hempel, et ce même si Hempel lui-même, dans ses modèles, évitait d'employer la notion de cause, parce qu'il acceptait le scepticisme humien à l'égard de la nécessité rattachée aux causes.

Deuxièmement, il y a une parenté de forme entre l'explication fondée sur les lois chez Hempel, et l'explication fondée sur des généralisations invariantes que propose Woodward. Chez Hempel, les modèles de l'explication prennent la forme d'une inférence où les prémisses sont des conditions initiales et des lois de la nature (à titre d'énoncés généraux), et où la conclusion est un phénomène

à expliquer. Chez Woodward, le schéma explicatif est également une inférence où les prémisses sont des causes et où l'énoncé général est non pas une loi, mais une généralisation invariante, et où la conclusion est le phénomène à expliquer. Dans les deux cas, le schéma fait explicitement référence à la relation entre les causes et les effets. C'est pourquoi j'entreprends le chapitre 3 par une présentation contrastée des positions de Hempel (et de ceux qui adoptent une conception nomologique de l'explication), comparées à celles de la conception manipulative de l'explication.

Je termine cette section introductive en mettant sur la table un souci potentiel. Certains penseurs soutiennent que l'explication (causale) est une notion assez différente de celle de causalité, et que l'on change de sujet en passant de l'une à l'autre. En guise de réponse, je voudrais citer la position de Woodward, sans tenter de la justifier pour le moment :

Unlike some philosophers, I draw no sharp distinction between providing a causal explanation of an outcome [...] and providing information about the causes of that outcome.<sup>32</sup>

Si cette position est juste, il est correct de considérer l'explication causale comme un concept qu'on peut dériver du concept de cause.

### 3.2 Deux modèles nomologiques de l'explication scientifique : les modèles DN et IS

L'explication scientifique peut avoir pour but la compréhension d'une généralisation ou la compréhension d'un événement singulier. Un exemple d'explication d'une généralisation est la dérivation d'une généralisation circonscrite, comme une des lois de Kepler portant sur la trajectoire des planètes autour du soleil, à partir de généralisations de portée plus étendue comme les lois de Newton. L'explication scientifique peut aussi avoir pour but la compréhension d'un événement singulier comme, par exemple, la position de telle planète à telle date, ou encore le fait que M. Jones soit atteint de telle maladie particulière à telle date. La présentation qui suit de deux versions des modèles nomologiques présentés par Hempel se concentre sur les explications visant des événements singuliers.

---

<sup>32</sup> Woodward, J. (2010), p. 291

Le premier modèle est un schéma en vertu duquel une explication scientifique est un raisonnement déductif dont une prémisse est une ou plusieurs lois et dont la conclusion est le fait singulier à expliquer. On l'appelle le modèle déductif nomologique ou encore le «modèle DN»<sup>33</sup>. Voici ses composantes et certaines règles à respecter :

#### (3.2.1) Modèle DN

Prémisses ou *explanans*

Des conditions initiales  $C_1, C_2, \dots, C_N$

Un ou plusieurs énoncés en forme de loi  $L_1, L_2, \dots, L_N$

Conclusion ou *explanandum*

Événement singulier

Règles

- 1) L'argument déductif doit être valide.
- 2) L'*explanans* doit inclure au moins une loi requise de façon essentielle.
- 3) L'*explanans* doit avoir un contenu empirique.
- 4) L'*explanans* doit être vrai.

Le modèle DN vise entre autres à assurer la certitude et l'objectivité de l'explication. Le modèle recèle également une certaine conception de ce que signifie la compréhension d'un phénomène : expliquer un phénomène qui a déjà eu lieu, c'est montrer que le phénomène devait être attendu ('was to be expected') et ce, en vertu d'au moins une loi requise de façon essentielle. Il faut bien sûr comprendre ce que les tenants du modèle DN entendent par la notion de loi. J'y reviens à la section 3.3.

L'application du modèle DN peut facilement être étendue, des phénomènes qui ont eu lieu, aux phénomènes futurs. Si toutes les conditions pertinentes d'un phénomène à venir sont présentes, alors la conclusion du raisonnement en forme DN constitue une prévision certaine. Le tenant du modèle DN accepte l'idée que les explications et les prévisions sont des notions symétriques, dont la seule différence tient à leur position temporelle.

---

<sup>33</sup> Je m'inspire ici de la présentation du modèle DN se trouvant dans Psillos, S. (2002), p. 218-222

Je mentionne ici brièvement deux variantes du modèle DN. Premièrement, on peut dans le modèle DN représenter l'explication de certaines lois ou généralisations déterministes (plutôt que l'explication de faits singuliers). Il s'agit alors de l'explication d'une généralisation déterministe applicable dans un domaine restreint, à partir d'une loi applicable universellement. Un exemple d'une telle explication est celle qui part des lois de la mécanique et de la gravitation de Newton, pour dériver et expliquer l'orbite elliptique des planètes du système solaire. Une seconde variante, à laquelle Hempel a donné le nom de modèle DS (pour 'Deductive Statistical') part d'une généralisation statistique dont le domaine est large et en dérive déductivement une généralisation statistique dont le domaine est plus étroit.

Hempel reconnaissait que beaucoup de phénomènes singuliers sont susceptibles d'être expliqués sans que l'explication atteigne le niveau d'exactitude habituellement rattaché aux conclusions découlant d'une loi. Dans ce cas, le phénomène, ou la mesure d'une propriété rattachée au phénomène, ne peut être déduite de façon certaine d'une loi, mais peut tout au plus être inférée avec une certaine probabilité. Quand on dit que l'espérance de vie d'un Canadien de sexe masculin est de 80 ans, on veut dire que les Canadiens contemporains ont une probabilité de 50% d'être encore en vie à l'âge de 80 ans, sans pouvoir dire si le citoyen particulier M. Jones atteindra cet âge. Hempel a donc proposé un modèle probabiliste d'explication des phénomènes singuliers, selon lequel l'explication consiste en une inférence «inductive» faisant appel à une généralisation probabiliste<sup>34</sup>. (L'emploi du terme «inductif» dans le présent contexte est un peu différent du sens habituel. L'induction et le raisonnement inductif dénotent habituellement un raisonnement en vertu duquel on infère une généralisation à partir de plusieurs faits singuliers qui seraient des instances de cette généralisation. Ainsi on peut induire la généralisation que tous les corbeaux sont noirs à partir de l'observation, sans exception jusqu'ici, de corbeaux noirs. Mais dans le contexte présent, la notion de modèle inductif vise plutôt une inférence partant d'une généralisation probabiliste et expliquant un fait singulier qui est hautement probable, mais pas certain, comme il l'aurait été si l'inférence avait été déductive. J'emploierai dorénavant les termes «induction» et «inductif» dans ce sens particulier). Le modèle est le modèle inductif statistique (pour 'inductive statistical') ou «modèle IS»<sup>35</sup> et ses composantes sont les suivantes :

---

<sup>34</sup> L'exposé initial de ce modèle se trouve dans Hempel, C. G. (1965), p. 381 sqq.

<sup>35</sup> Psillos, S. (2002), p. 242-247

### (3.2.2) Modèle IS

#### Prémisses ou *explanans*

Fa (L'individu a possède la propriété F)

Autres conditions initiales

Prob (G/F) = r (Une loi probabiliste, qui affirme que la probabilité conditionnelle qu'un porteur de la propriété F soit aussi un porteur de la propriété G est égale à r et est élevée (proche de 1))

#### Conclusion ou *explanandum*

Ga

Dans le modèle IS, la conclusion Ga signifie que l'individu a possède la propriété G non avec certitude, mais avec la probabilité r que lui assigne la loi statistique Prob (G/F) = r.

Il y a des parallèles entre le modèle IS et le modèle DN. Les deux conçoivent l'explication d'un phénomène particulier comme une inférence à partir d'une loi déterministe ou d'une loi probabiliste, selon le cas. Dans les deux modèles, la compréhension que confère l'explication d'un phénomène passé est qu'elle indique que le phénomène devait être attendu, avec certitude dans le modèle DN, et avec une probabilité élevée dans le modèle IS. Il y a dans les deux cas une extension possible aux phénomènes futurs.

Une faiblesse évidente du modèle IS est qu'il se limite à l'explication des phénomènes à probabilité élevée. Selon Psillos<sup>36</sup>, cette caractéristique est nécessaire dans le cadre que se donne Hempel. Si on enlève l'exigence de probabilité élevée, les explications en vertu du modèle IS cesseraient d'être des événements singuliers auxquels on devrait s'attendre. Par contre, il y a des phénomènes à probabilité faible que nous souhaitons expliquer. Nous voudrions par exemple expliquer pourquoi une très faible partie de la population humaine, et en particulier telle personne, est atteinte de telle maladie rare. Pour cette raison entre autres, le modèle IS a été mis de côté dans les recherches sur la nature de l'explication probabiliste. Il n'a plus la pertinence qu'on continue d'accorder au modèle DN.

---

<sup>36</sup> Psillos, S. (2002), p. 244-246

Une autre observation importante, selon moi, porte sur une différence essentielle entre les explications probabilistes de forme IS et les explications déterministes (de forme DN). Les explications de forme DN sont des arguments déductifs valides liant des prémisses P à une conclusion Q, et en vertu de cette forme, ils ont la propriété d'être monotones<sup>37</sup>. Cela signifie que si on ajoute aux prémisses P d'un argument DN une prémisse additionnelle P\*, la conclusion Q continuera d'en découler. Par exemple, si tous les humains sont mortels, alors aucune information supplémentaire sur la nationalité ou l'âge des humains visés ne viendra altérer la conclusion. Tous les humains australiens sont mortels, tous les humains de plus 40 ans sont mortels, etc. Les arguments inductifs par contre n'ont pas la propriété d'être monotones. Cela signifie que si on ajoute une nouvelle prémisse P\* à un argument inductif, il se peut que la conclusion antérieure Q ne tienne plus. L'exemple de Psillos est le suivant. L'argument inductif initial est : (P1) Jones a une infection de streptocoques, (P2) Jones est traité à la pénicilline, (P3 loi probabiliste) Les patients infectés par des streptocoques et traités avec de la pénicilline ont une forte probabilité de guérison, (Conclusion à probabilité élevée) Jones guérit. L'argument inductif modifié reflète l'information supplémentaire que le type de streptocoque qui infecte Jones provient d'une souche qui résiste à la pénicilline. Alors manifestement la conclusion ne tient plus et pourrait être renversée. Il se pourrait que Jones ait une très forte probabilité de ne pas guérir s'il est traité avec de la pénicilline.

Ces considérations sur le caractère non-monotone des explications probabilistes de forme IS permettent d'éclairer de façon différente les arguments déterministes de forme DN. Qu'est-ce qui fait la différence entre une situation qui permet une explication déterministe et celle qui ne permet qu'une explication probabiliste ?<sup>38</sup> La principale différence, selon moi, tient à la présence, dans les prémisses de l'argument déterministe, de tous les facteurs initiaux et de toutes les généralisations déterministes pertinentes au comportement de l'*explanandum*. Ainsi les explications en vertu du

---

<sup>37</sup> Une présentation plus approfondie de la caractéristique de monotonie des arguments en forme DN et de non-monotonie des arguments en forme IS se trouve dans Psillos, S. (2002), p. 247-252.

<sup>38</sup> Je laisse de côté dans la présente discussion, les phénomènes «essentiellement probabilistes», dont l'exemple habituel est la désintégration des atomes d'une masse de matière radioactive. Ces phénomènes sont tels qu'il serait impossible, même avec une information complète sur chacun des atomes de la masse, de savoir quel atome sera le prochain à se désintégrer. Les phénomènes que je discute ici sont plutôt des phénomènes déterministes, au sens qu'une information complète, si on la connaissait, sur l'état de chaque entité au temps t permettrait de déterminer avec certitude l'état de chaque entité au temps t+1. La source de l'incertitude menant à l'emploi de probabilités est dans ce cas-ci épistémique et non ontologique.

modèle DN et celles en vertu du modèle IS ne semblent pas aussi distinctes que ce que l'on croit généralement.

Dans la section qui suit, je passe en revue la caractérisation des lois de la nature dans les modèles DN et IS et certaines objections à ces modèles.

### 3.3 Les lois de la nature dans les modèles DN et IS et les problèmes de ces modèles

La notion langagière d'«énoncé en forme de loi» et son correspondant ontologique, la notion de «loi» ou de «loi de la nature», sont des éléments essentiels du modèle DN et des autres conceptions nomologiques de l'explication scientifique. Les philosophes des sciences ont tenté de caractériser la notion de loi requise dans le cadre du modèle DN pour, entre autres, distinguer les énoncés en forme de loi d'autres énoncés de forme semblable, mais moins porteurs en termes de causalité et d'explication. Ils se sont fondés sur les exemples paradigmatiques des lois fondamentales de la physique. Je dois dire par contre que leurs caractérisations constituent une reconstruction logique de ce que devraient être les lois de la nature, plutôt qu'une tentative de décrire les caractéristiques des généralisations réelles qu'on retrouve dans les diverses sciences. Les résultats de ces travaux ont été décevants. Hempel, par exemple, a conclu qu'il semble très difficile de caractériser les énoncés en forme de loi, sans faire appel à la notion même de loi ou à une autre notion nomologique<sup>39</sup>.

Je présente maintenant une version de la notion d'énoncé en forme de loi, une version de la notion de loi, et un certain nombre de difficultés rattachées aux modèles nomologiques DN et IS. Je vais procéder en comparant informellement (a) une version des caractéristiques des énoncés en forme de loi, et (b) une liste d'aspects (au sens de l'anglais 'dimensions') ontologiques qu'il faut considérer pour caractériser les lois et/ou les généralisations causales.

Je réfère au premier élément de la comparaison en utilisant les mots «une version de» pour bien refléter l'idée mentionnée plus haut que le projet de caractériser les énoncés en forme de loi, aux

---

<sup>39</sup> Hempel, C. G. (1965), p. 338

fins de préciser le sens de cette expression dans le modèle DN, n'a pas mené à des résultats concluants. Je tire la version qui suit du livre de Salmon : *Four Decades of Scientific Explanation* <sup>40</sup>.

(3.3.1) Une version des caractéristiques formelles des «énoncés en forme de loi» (en anglais 'lawlike sentences') requises aux fins du modèle DN :

1) [Lawlike sentences] have universal form.

Commentaire : la forme universelle peut être une affirmation simple comme dans : «Tous les hommes sont mortels», ou un énoncé conditionnel universel, comme dans «Pour tout x, si x est un homme, alors x est mortel».

2) The scope [of lawlike sentences] is unlimited.

3) [Lawlike sentences] do not contain designations of particular objects.

4) [Lawlike sentences] contain only purely qualitative predicates.

Les caractéristiques incluses dans (3.3.1) ne constituent pas une structure claire sur le plan logique, principalement parce que le langage ordinaire n'est pas structuré de manière logique. Le langage ordinaire permet en effet de rendre la même signification de plusieurs manières différentes. Certains termes, comme celui de 'scope', sont vagues. La caractéristique 3) vient éliminer des énoncés qui seraient incorrectement admis par la seule caractéristique 1). Je renvoie le lecteur au texte de Salmon pour quelques illustrations des difficultés d'interprétation de ces règles. L'élément important est l'intention derrière la liste de caractéristiques formelles. J'indiquerai cette intention après la présentation (qui suit) d'une liste de caractéristiques ontologiques à considérer dans la notion de loi.

Le second élément de la comparaison est une liste d'aspects ontologiques que les philosophes considèrent importants lorsqu'ils caractérisent les lois dans l'explication scientifique et dans l'explication des phénomènes de la vie quotidienne<sup>41</sup>. La plupart des aspects peuvent être considérés comme des axes sur lesquels il y a plus d'une position.

(3.3.2) Aspects ontologiques généraux des lois ou des généralisations explicatives

1) Être sans exception (au moins dans un domaine restreint d'application) ou, au contraire, comporter des exceptions : les généralisations probabilistes ont des exceptions dans leur domaine d'application.

---

<sup>40</sup> Salmon, W. C. [1989] (2006), p. 13

<sup>41</sup> Je m'inspire ici de Salmon, W. C. [1989] (2006), p. 13-17 et Carroll, J. C. (2010)



- 2) S'appliquer universellement, ou seulement dans un domaine restreint : on examine si les lois ou généralisations s'appliquent à tous les objets (ou états) du monde réel, quel que soit le lieu, et également à tous les objets (ou états) du monde réel dans le passé et dans l'avenir. Si une généralisation n'est pas universelle, on examine l'extension de son domaine d'application.
- 3) Posséder ou non la capacité de fonder les énoncés contrefactuels : On examine les lois ou généralisations proposées pour voir si elles permettent des inférences sur ce qui aurait pu se passer si certaines prémisses du raisonnement explicatif (certaines conditions initiales par exemple) avaient été différentes.
- 4) Être indépendantes d'autres lois ou généralisations, ou ne pas être indépendantes : (cet aspect est aussi appelé la séparation des mécanismes ou la modularité). Cet aspect se comprend dans le contexte d'un monde où de nombreuses forces ou mécanismes influencent un phénomène à expliquer. On examine les généralités pour voir dans quelle mesure les mécanismes qu'elles représentent fonctionnent indépendamment des autres mécanismes qui entrent en jeu.

Dans le contexte du modèle DN, où l'*explanandum* doit découler des prémisses en vertu d'un raisonnement déductif, il est clair que la loi (et l'énoncé en forme de loi qui lui est rattaché) doit être sans exception. De plus, dans la mesure où le fait pour une loi de s'appliquer seulement en un lieu particulier ou seulement à une certaine époque ouvrirait la porte à des exceptions, les philosophes exigent habituellement que les lois s'appliquent universellement dans le temps et dans l'espace. Ainsi ces deux conditions (absence d'exceptions et universalité) sont les deux premières caractéristiques des lois généralement exigées dans les analyses philosophiques. Ces exigences s'appuient en outre sur le fait que les lois paradigmatiques de la physique remplissent ces deux conditions. Et on peut dire que l'objectif ou l'intention visée par la liste des caractéristiques formelles des énoncés en forme de loi (3.3.1) est de distinguer, sur la base de leur forme linguistique, les énoncés reflétant les caractéristiques d'absence d'exception et d'universalité de ceux qui ne les reflètent pas.

J'aborde maintenant le détail des difficultés ou objections que l'on fait aux modèles DN et IS.

Difficulté 1 (de la notion de loi dans le modèle DN)) On est incapable de distinguer, sur la base des seules caractéristiques formelles, entre les énoncés en forme de loi véritablement explicatifs et les énoncés en forme de loi qui ne sont pas véritablement explicatifs. Cette assertion exige quelques explications. Je commence par remarquer que la caractérisation formelle (liste 3.3.1) paraît et est très exigeante à l'égard des deux premiers aspects généraux des généralisations (liste 3.3.2). Elle exige que les énoncés en forme de loi n'aient pas d'exception dans leur domaine d'application, et de plus que ces énoncés s'appliquent universellement. Malgré cela, on peut construire des assertions qui se conforment aux caractéristiques formelles mais que nos intuitions à l'égard des lois rejettent. Je reprends un exemple bien connu qui proviendrait ultimement de H. Reichenbach. Je présente ici la version de Salmon de cet exemple<sup>42</sup>. Il s'agit de comparer les deux énoncés suivants :

(3.3.3) No gold sphere has a mass greater than 100,000 kg.

(3.3.4) No enriched uranium sphere has a mass greater than 100,000 kg.

Les deux énoncés satisfont les caractéristiques formelles des lois. Ils visent de façon générale tous les objets de l'univers, en tout temps. La référence aux sphères d'un élément chimique particulier est acceptable, car on peut réécrire (3.3.3), par exemple, sous la forme «Pour tout objet x dans l'univers, dans le passé, maintenant et dans l'avenir, si x est une sphère d'or pur, alors x n'a pas une masse plus grande que 100 000 kilos». La référence au kilo, à titre d'objet particulier (un étalon de mesure), est acceptable car elle peut être remplacée par une référence à la masse atomique de l'atome de carbone-12, un objet «général». Mais, même si les deux énoncés satisfont aux caractéristiques formelles des lois et même si les deux sont considérés comme vrais dans le monde réel, l'énoncé 3.3.3 n'est pas considéré comme une loi véritable. Il semble vrai que, dans l'univers que nous connaissons, aucune sphère d'or pur de plus 100 000 kilos n'existe. Mais il semble également vrai que, dans un monde possible régi par les mêmes lois de la nature que celles du monde réel et où certaines conditions initiales auraient été différentes, il serait physiquement possible qu'une sphère d'or pur de plus de 100 000 kilos existe. Nous pensons par contre que l'énoncé (3.3.4) est une loi de la nature, parce que nous croyons que les facteurs déterminant la masse critique des éléments radioactifs font en sorte que, dans un monde régi par les mêmes lois de la nature et où certaines conditions initiales auraient été différentes, il serait physiquement

---

<sup>42</sup> Salmon, W. C. [1989] (2006), p. 14-15

impossible qu'une sphère d'uranium enrichi de plus de 100 000 kilos existe. Ainsi les caractéristiques formelles des énoncés en forme de loi retenues par Salmon ne permettent pas de distinguer les lois de certaines généralisations vraies dans le monde réel, mais considérées comme accidentelles.

Le diagnostic de cette première difficulté est que les caractéristiques formelles de la liste (3.3.1) ne sont pas des conditions suffisantes pour assurer qu'un énoncé représente une loi véritable. Nos intuitions à l'égard des relations et des explications causales nous amènent à penser que le 3<sup>ème</sup> élément de la liste des aspects, la capacité de fonder les énoncés contrefactuels est une condition supplémentaire importante qui doit être présente dans les lois.

Difficulté 2 (de la notion de loi dans le modèle DN). Les généralisations considérées comme explicatives dans la plupart des sciences ne satisfont pas les caractéristiques ontologiques (l'absence d'exceptions et l'universalité) rattachées aux lois par les philosophes. Ces caractéristiques ne sont satisfaites que par certaines lois de la physique et de la chimie. Les généralisations des sciences biologiques et médicales et celles des sciences humaines s'appliquent en vertu des conditions biologiques en vigueur sur une planète particulière, la Terre, et, dans le cas des sciences humaines, à une espèce particulière (homo sapiens), dont l'évolution historique semble échapper à toute loi universelle et sans exception. Ceci implique que le modèle DN n'est pas une modèle pertinent de l'explication dans les sciences biomédicales et dans les sciences humaines.

Difficulté 3 (de la notion de loi dans le modèle DN). Un grand nombre d'explications de la vie quotidienne (portant sur tous types de phénomènes, y compris des phénomènes relevant de la physique et de la chimie) semblent fonctionner sans référence explicite à des lois, et l'on peut produire et recevoir des explications sans connaître les lois qui (peut-être) régissent ces phénomènes. Je reconnais que je sors ici du domaine de l'explication scientifique. Mais on pourrait penser que si l'explication scientifique en matière de physique et de chimie (qui sont les domaines où on trouve le plus de lois sans exception et universelles) exigent le recours à des lois, alors les explications de la vie quotidienne relevant de la physique et de la chimie devraient, d'une manière

ou d'une autre, référer elles aussi à des lois. Mais le philosophe Michael Scriven<sup>43</sup> a suggéré que ce n'est pas le cas. Son exemple le plus célèbre et le plus discuté est le suivant :

(3.3.5) Une bouteille d'encre ouverte se trouve sur une table. Une personne donne par inadvertance un coup de genou sur une patte de la table. La bouteille se renverse et l'encre se répand sur le tapis.

Il faut d'abord reconnaître que les trois phrases de (3.3.5) semblent bien constituer une explication. Nos intuitions sur les explications de la vie quotidienne l'acceptent comme telle. Et nous pensons qu'un enfant de six ans, qui n'a pas de notion de loi de la nature, comprendrait bien cette explication. Mais il n'y a pas dans (3.3.5) de référence explicite à une généralisation, avec ou sans exceptions, universelle ou pas.

Face à la thèse de Hempel et des autres tenants des conceptions nomologiques de l'explication, selon laquelle une bonne explication scientifique d'un phénomène déterministe (ou d'un phénomène déterministe de la vie quotidienne relevant des sciences) doit faire référence à au moins une loi satisfaisant les exigences du modèle DN, Scriven soutient qu'il y a des explications valables et compréhensibles qui n'y font pas référence. Scriven demanderait aux tenants des conceptions nomologiques (i) quelles lois relient les chocs provoqués par un genou aux déversements de bouteilles d'encre, et (ii) en quoi la connaissance de ces lois est requise par un enfant pour comprendre l'explication (3.3.5). Et la réponse de ces tenants serait, à l'égard de (i), que même s'il n'y a pas de lois reliant explicitement les chocs provoqués par un genou aux déversements, l'explication (3.3.5) donne implicitement et indirectement de l'information sur les lois qui régissent ce phénomène, comme la loi sur la conservation de la quantité de mouvement suite à un choc, et la loi de la chute des corps en vertu de la gravitation. Ils n'auraient toutefois pas de réponse à l'égard de (ii). Ils devraient admettre que la connaissance des lois pertinentes n'est pas requise pour émettre et comprendre de nombreuses explications de phénomènes physiques et chimiques<sup>44</sup>.

---

<sup>43</sup> Woodward J. (2003), p. 156 donne la référence suivante : Scriven, M. (1959), «Truism as the Grounds of Historical Explanations», Dans Gardiner, P. (éd.) (1959). *The Nature of Historical Explanation*, Oxford, OUP, p. 456

<sup>44</sup> Sur cet exemple de Scriven, voir Woodward J. (2003), p. 156

Une considération supplémentaire importante à l'égard de cet exemple se rattache à ce que Woodward appelle «le niveau» de l'explication DN proposée. Il s'agit de montrer que, lorsqu'une loi ou généralisation incluse dans une explication DN est de niveau indirect, l'explication ne permet pas à ses usagers d'atteindre un des buts souvent reconnus de l'explication : la compréhension de l'*explanandum*. Woodward, reprenant une distinction faite par Davidson, distingue (i) les généralisations explicatives directes, et (ii) les généralisations explicatives indirectes. Dans le cas de la généralisation directe, la loi (ou la généralisation) présente dans le raisonnement explicatif est exprimée dans les mêmes termes que les conditions initiales de l'*explanans* et que l'*explanandum*. Une telle loi dans le cas de (3.3.5) relierait explicitement les coups de genou aux bouteilles d'encre renversées. Une généralisation indirecte, par contre, n'est pas exprimée dans les mêmes termes que les conditions initiales de l'*explanans* et l'*explanandum*. Les lois sans exception et universelles invoquées par les tenants du modèle DN comme lois sous-jacentes à (3.3.5) sont manifestement des généralisations indirectes. Une conséquence de la référence à des lois ou généralisations sous-jacentes indirectes pour interpréter les explications, dans le modèle DN, est que l'objectif de compréhension semble compromis. Si la compréhension découle d'une inférence déductive, mais que les lois ou généralisations qui permettent l'inférence sont indirectes et ne sont pas présentes explicitement dans l'explication, alors on voit mal comment la référence à la loi serait une condition nécessaire de la compréhension de l'*explanandum*. Dans l'explication (3.3.5), si la compréhension du renversement de la bouteille d'encre doit provenir indirectement d'une loi de conservation de la quantité de mouvement, et si cette loi n'est pas mentionnée, alors il semble que la loi ne joue aucun rôle nécessaire dans la compréhension du phénomène<sup>45</sup>.

Difficulté 4 (de la notion de loi du modèle DN). Dans de nombreux cas d'explication scientifique, les explications fonctionnent sans référence à aucune loi connue explicitement. Les tenants des conceptions nomologiques postulent alors l'existence d'une ou plusieurs lois, sans pouvoir les énoncer. C'est ce que Woodward appelle «la stratégie de la structure cachée».

L'exemple de Scriven cité précédemment laisse une porte de sortie aux tenants des conceptions nomologiques, en ce sens que l'idée qu'il existe des lois sous-jacentes, universelles et sans exception, régissant le comportement de la bouteille d'encre, paraît plausible. Seul semble établi le

---

<sup>45</sup> Woodward J. (2003), p. 179

fait épistémique que la connaissance de lois conformes aux exigences du modèle DN n'est pas nécessaire pour la production et la compréhension de l'explication. Mais le domaine biomédical nous offre de nombreux exemples d'explications sans loi, conforme au modèle DN, connue par qui que ce soit. Par exemple, on ne connaît pas présentement de généralisation mettant en lumière les facteurs nomologiques affectant la présence ou l'absence de la maladie d'Alzheimer. On peut identifier certains facteurs qui affectent la probabilité d'avoir cette maladie, comme par exemple l'âge d'une personne, mais il s'agit là d'une corrélation qui doit être examinée plus avant pour déterminer si elle peut être associée à une généralisation causale probabiliste. Dans ce type de cas, les tenants de conceptions nomologiques de l'explication affirment souvent qu'une explication comme :

(3.3.6) L'âge avancé de Mme Jones a causé sa maladie d'Alzheimer

est une ébauche d'explication qui réfère implicitement à une ou à des lois sous-jacentes que nous ne connaissons pas encore. Personne à l'heure actuelle ne peut dire avec confiance quelles pourraient être ces lois. Dit autrement, les tenants des conceptions nomologiques postulent l'existence de ces lois, et leur donnent un rôle difficile à justifier dans les explications concrètes.

Les généralisations *ceteris paribus*, qui apparaissent fréquemment dans les sciences humaines, sont un autre cas où des généralisations sont considérées comme explicatives en vertu de la stratégie de la structure cachée. Un exemple classique d'une telle généralisation est une des composantes de la loi de l'offre et de la demande en économie, qui peut s'énoncer ainsi :

(3.3.7) Si l'offre d'un bien augmente (diminue) pendant que la demande reste inchangée, alors le prix du bien diminue (augmente), *ceteris paribus*.

L'expression «*ceteris paribus*» signifie dans ce contexte que la généralisation est vraie pourvu qu'aucun facteur perturbateur ne vienne la contrecarrer. Un exemple de facteur perturbateur à l'égard de (3.3.7) pourrait être le prix des biens substitués. Si le bien visé par (3.3.7) est le gaz, et si le bien substitut est le pétrole, alors une offre accrue de gaz peut être accompagnée d'une hausse du prix du gaz, si le prix du pétrole est également en hausse. En vertu de la stratégie de la structure cachée, on soutient que les généralisations *ceteris paribus*, même si elles sont sujettes à de nombreuses exceptions, seraient néanmoins explicatives au sens où elles transmettraient de

l'information sur des lois complémentaires qui, bien que non mentionnées et parfois même inconnues, seraient sans exception<sup>46</sup>.

Difficulté 5 (rattachée au modèle DN). L'explication est souvent asymétrique au sens où un élément de l'*explanans*, en association avec la loi contenue dans l'*explanans*, explique l'*explanandum*, tandis que l'*explanandum*, en association avec la loi, n'explique pas cet élément de l'*explanans*. Par contre, dans le modèle DN de l'explication, pour les mêmes *explanans* et *explanandum*, l'*explanans* et l'*explanandum* se trouvent parfois dans une relation symétrique, au sens où l'*explanans* (en association avec la loi) permet de déduire l'*explanandum* et, inversement, l'*explanandum* (en association avec la loi) permet de déduire l'*explanans*. Un exemple fréquemment utilisé de cette difficulté est celui de la longueur de l'ombre projetée par un poteau<sup>47</sup>.

(3.3.8) Selon nos intuitions sur l'explication, la hauteur d'un poteau, associée à l'angle du rayon de lumière par rapport avec la verticale et à la loi de la propagation rectiligne de la lumière (dans un milieu de densité uniforme), explique la longueur de l'ombre projetée par le poteau. Par contre, la longueur de l'ombre, (associée à l'angle et à la loi), n'explique pas la hauteur du poteau.

Il y a une asymétrie de l'explication, selon nos intuitions. Par contre, en vertu du modèle DN, la relation entre la hauteur du poteau et la longueur de l'ombre est explicativement symétrique, puisque

(3.3.9) En vertu du modèle DN, la hauteur du poteau, associée à l'angle du rayon de lumière par rapport avec la verticale et à la loi de la propagation rectiligne de la lumière, permet de déduire la longueur de l'ombre projetée par le poteau, et inversement, la longueur de l'ombre, (associée à l'angle et à la loi), permet de déduire la hauteur du poteau.

Difficulté 6 (de la notion de loi dans le modèle IS). Je reprends ici une difficulté déjà mentionnée à l'égard des explications selon le modèle IS. Le modèle IS ne permet pas de rendre compte de l'explication de phénomènes singuliers à probabilité faible. Les explications de forme IS se limitent aux phénomènes rattachés à une loi probabiliste à probabilité élevée. Ceci découle du fait que Hempel conçoit l'objectif de l'explication comme celui de montrer que le phénomène à expliquer était attendu en vertu d'une loi ('was to be nomologically expected'), et que le rôle de la loi est

---

<sup>46</sup> Woodward, J. (2003), p. 157-159

<sup>47</sup> Woodward, J. (2003), p. 154

d'entraîner, par inférence inductive (au sens large), c'est-à-dire avec un résultat incertain, dans le modèle IS, le phénomène à expliquer. Seules les lois à probabilité élevée permettraient une telle inférence.<sup>48</sup>

Difficulté 7 (de la notion de loi dans le modèle IS). Deux facteurs A et B peuvent être reliés par une généralisation statistique à haute probabilité (comme l'exige le modèle IS), mais ne pas pouvoir figurer comme cause et effet dans une explication causale. On peut éclairer cela en commentant l'observation bien connue : «La corrélation (statistique) entre deux variables n'entraîne pas la présence d'un lien causal.» Disons tout d'abord que la corrélation statistique entre deux facteurs ou variables est une relation symétrique, alors que la causalité, selon nos intuitions, est souvent asymétrique, tel que mentionné à la difficulté 5). Plus généralement, il y a au moins cinq situations différentes qui donnent lieu à une corrélation statistique positive entre les variables A et B, et seule une des cinq en est une où A est une cause de B, et peut donc figurer dans une explication correcte par le facteur A du phénomène B.

(3.3.10) Cinq situations donnant lieu à une corrélation statistique entre A et B

- 1) A est une cause de B
- 2) B est une cause de A
- 3) A et B sont tous deux des effets d'une cause commune Z, comme c'est le cas dans l'exemple du baromètre (2.1.2)
- 4) A et B sont des facteurs (des variables) reliées non pas causalement, mais plutôt conceptuellement. Un exemple d'une telle relation conceptuelle est celle entre le PNB (produit national brut) d'un pays et le PIB (produit intérieur brut) de ce pays. Les deux mesures sont corrélées statistiquement pour la plupart des pays, mais ce n'est pas en vertu d'une relation causale. C'est plutôt parce que le PIB est défini comme le PNB moins le revenu net provenant de biens détenus à l'étranger.
- 5) A et B sont des facteurs (des variables) dont le lien (de corrélation statistique) est accidentel.

Dans la mesure où la plupart des généralisations statistiques peuvent donner lieu à des mesures de corrélation statistique, une conception acceptable de la relation causale et de l'explication doit

---

<sup>48</sup> Psillos, S. (2002), p. 242-244



permettre de distinguer les véritables relations causales reliant A et B des autres situations de corrélation statistique positive entre ces facteurs (variables)<sup>49</sup>.

### 3.4 La conception manipulative de l'explication causale

Les deux sections précédentes présentaient une des conceptions nomologiques de l'explication. La présente section aborde la conception manipulative de l'explication causale, dans la perspective d'un contraste avec la conception nomologique.

Je débute par une présentation sommaire de la conception manipulative. Expliquer un phénomène qui a déjà eu lieu, selon cette conception, consiste non seulement à montrer que l'*explanandum* devait être attendu, mais également à montrer comment l'*explanandum* aurait été différent si les conditions initiales et/ou les conditions d'arrière-plan avaient été différentes de diverses façons. L'explication causale a comme caractéristique essentielle de révéler des patterns systématiques de dépendance contrefactuelle entre certains facteurs de l'*explanans*, certaines conditions d'arrière-plan, et l'*explanandum*. Pour le dire autrement, donner une explication causale consiste à répondre à un ensemble de questions de la forme : qu'est-ce qui se serait passé si certains des facteurs de l'*explanans* et/ou certaines des conditions d'arrière-plan avaient été différents ? Je référerai plus loin à ces questions en employant l'expression «Questions W», pour Questions 'What if things had been different'<sup>50</sup>.

L'idée que l'élément essentiel de l'explication est le fait de la situer dans un ensemble d'assertions contrefactuelles se rattache également à l'idée que le but principal de l'explication causale est de permettre à ceux qui en disposent de manipuler et contrôler les événements.

La description des paragraphes précédents portait sur l'objectif et la nature de l'explication. Je présente maintenant les composantes d'une explication complète dans la conception manipulative. L'explication causale prend la forme d'un raisonnement, dont les prémisses comprennent : (i) au moins une généralisation, (ii) les valeurs initiales des facteurs mentionnés dans la généralisation, et dont la conclusion est l'*explanandum*. La principale différence entre ces composantes et celles des

---

<sup>49</sup> Reiss, J. (2013), p. 89-90

<sup>50</sup> Woodward, J. (2003), p. 191

modèles DN et IS est que la généralisation n'est pas une loi, mais plutôt une assertion générale possiblement plus faible, que je décrirai davantage dans la présente section 3.4 et dans la section 3.5.

Ce ne sont pas tous les changements des facteurs mentionnés dans une généralisation qui permettent de distinguer les explications causales de celles qui n'en sont pas. Seuls les contrefactuels décrivant les résultats d'une intervention permettent de faire cette distinction. Comme je l'ai indiqué au chapitre 2, une intervention sur un facteur (possiblement causal) (X) à l'égard d'un *explanandum* (Y) change la valeur de X de telle manière que si la valeur de Y change par la suite, cela n'a lieu qu'en vertu du changement de X, et pas en vertu du changement d'un facteur autre que X.<sup>51</sup>

Il y a également une parenté importante, dans la conception manipulative, entre, d'une part, les explications et, d'autre part, les causes contributoires. Une cause contributoire X de Y est caractérisée par un lien de dépendance contrefactuelle, sous intervention, entre X et Y, dans des conditions où la nature précise du lien entre X et Y (i.e. la généralisation G) peut ne pas être connue. Dans l'explication complète, selon la conception manipulative, le lien entre X et Y doit être connu. Cela permet alors au récepteur de l'explication de répondre à des questions W à l'égard du lien entre X et Y.

Rassemblant ces divers éléments, Woodward donne la caractérisation technique suivante du concept d'explication dans la perspective manipulative :

(3.4.1) Explication ou (EXP) :

Suppose that M is an *explanandum* consisting in the statement that some variable Y takes the particular value *y* . Then an *explanans* E for M will consist of (a) a generalization G relating changes in the value(s) of a variable X (where X may itself be a vector or n-tuple of variables  $X_i$ ) and changes in Y, and (b) a statement (of initial or boundary conditions) that the variable X takes the particular value *x*. A necessary and sufficient condition for E to be (minimally) explanatory with respect to M is that (i) E and M be true or approximately so; (ii) according to G , Y takes the value *y* under an intervention in which X takes the value *x* ;

---

<sup>51</sup> Woodward, J. (2003), p.196-197

(iii) there is some intervention that changes the value of X from  $x$  to  $x'$ , where  $x \neq x'$ , with G correctly describing the value  $y'$  that Y would assume under this intervention, where  $y' \neq y$ .<sup>52</sup>

Alors que EXP caractérise l'explication de base, une explication plus substantielle reliera, en vertu d'une généralisation G, un éventail de changements de la variable X à une série de valeurs de Y, dans différentes circonstances d'arrière-plan.

Je crois qu'il est important de rattacher dès maintenant la caractérisation des explications, y compris la présence de la généralisation G, à la structure globale de ce mémoire. J'ai fait une distinction entre les concepts causaux minimaux, traités au chapitre 2 et les concepts causaux renforcés, que j'aborde au chapitre 3. J'ai indiqué que la différence entre les deux groupes de concepts est épistémique. Cette différence tient principalement à la connaissance par les agents d'une généralisation précise G reliant les variables de l'*explanans* et de l'*explanandum*.

Si les agents connaissent une généralisation précise G, déterministe ou probabiliste, alors ils ont accès aux concepts causaux minimaux (capacité de constater la présence d'une relation causale, cause contributoire et intervention) et aux concepts causaux renforcés (explication, invariance et capacité de distinguer entre les facteurs causaux présents dans G et d'autres facteurs causaux d'arrière-plan). De plus ils ont clairement accès aux concepts comme la stabilité de la relation causale et sa spécificité, que je présenterai à la section 3.8. Mais si les agents ne connaissent que des tendances, sans avoir accès à une généralisation G fournissant une information structurée sur les liens entre les variables, comme dans l'exemple (2.1.8) de l'étude de 1959 sur le lien entre l'habitude de fumer et le cancer, alors les agents n'ont qu'un accès limité aux concepts renforcés. Leur capacité d'expliquer est tout au plus limitée. (Je discuterai cette capacité limitée d'expliquer à la section 4.4). Le concept d'invariance est hors de portée. Leur accès aux concepts de tri comme la stabilité et la spécificité est également variable. Le tableau qui suit transmet ces informations de façon visuelle.

---

<sup>52</sup> Woodward, J. (2003), p. 203

Fig. 3.1 Concepts causaux minimaux et renforcés

Statut épistémique	Concepts minimaux	Concepts renforcés	Concepts de tri parmi les relations causales
	<ul style="list-style-type: none"> <li>- Existence de la relation causale</li> <li>- Cause contributoire</li> <li>- Intervention</li> </ul>	<ul style="list-style-type: none"> <li>- Explication</li> <li>- Invariance</li> <li>- Distinction des facteurs d'arrière-plan</li> </ul>	<ul style="list-style-type: none"> <li>- Stabilité</li> <li>- Proportionnalité</li> <li>- Spécificité</li> </ul>
Généralisation (G) déterministe connue	Oui	Oui	Oui
Généralisation (G) probabiliste connue	Oui	Oui	Oui
Généralisation (G) pas connue; simples tendances	Oui	Accès faible ou nul	Accès variable

Après ces précisions visant à situer le lecteur par rapport à la thèse épistémique de ce mémoire, je commente ici certains aspects controversés des explications en général, selon la conception manipulative, et je reviendrai à la discussion des généralisations dans la section 3.5.

#### 1) Les explications se rattachent à des changements

Une explication, selon EXP (3.4.1), doit être reliée à des changements, au sens où elle nous dit comment une quantité ou une valeur changerait sous l'effet de changements à une autre quantité ou valeur d'une des variables figurant dans l'*explanans*. Cette caractérisation a pour conséquence d'exclure certains raisonnements qu'on considérerait comme des explications correctes selon les conceptions nomologiques. C'est le cas en particulier des explications fondées sur des lois exprimant une impossibilité. Woodward donne l'exemple suivant. Soit la généralisation :

(3.4.2) All physical processes propagate at a speed less than or equal to that of light.

(3.4.2) possède plusieurs des caractéristiques d'une loi. Elle est, selon les théories acceptées actuellement, valable sans exception et universellement. D'autre part, elle exprime une nécessité, qui semble impliquer l'impossibilité correspondante : aucun processus physique ne peut se propager à une vitesse supérieure à celle de la lumière. Mais (3.4.2) ne peut pas être une composante d'une explication dans la conception manipulative parce qu'elle ne peut pas, telle qu'énoncée, être rattachée à une entité dont l'état serait soumis, de façon claire, à un changement. Quelle serait l'alternative à l'état «processus physique»? Peut-être un «pseudo-processus». Comment un processus pourrait-il passer de l'état processus physique à celui de pseudo-processus? Nous n'en avons pas de conception claire. Il n'y a donc pas d'intervention théorique applicable à cette entité/processus qui permette de déterminer s'il y a une dépendance contrefactuelle entre deux formes de processus et leurs vitesses de propagation.

Les considérations du paragraphe précédent sont néanmoins inquiétantes, en ce sens que nous voudrions pouvoir donner des explications de certaines nécessités et impossibilités physiques. Woodward suggère qu'il est possible de reformuler une généralisation comme (3.4.2) de manière à satisfaire les exigences de la conception manipulative. Il propose à titre d'exemple une généralisation à portée plus restreinte qui vise un des processus couverts par (3.4.2) : celui des particules à haute énergie émises lors de l'explosion d'une étoile. On pourrait initialement penser que ces particules pourraient, si l'énergie reçue de leur source était suffisamment grande, atteindre une vitesse (mesurée par rapport à la position de leur source d'énergie) supérieure à la vitesse finie de la lumière. La généralisation suivante exprimerait alors le phénomène de la limite de leur vitesse:

(3.4.3)The velocity of a particle [ejected from a supernova] approaches the speed of light asymptotically as its kinetic energy approaches infinity.

Cette généralisation n'est pas exprimée sous la forme d'une nécessité ou d'une impossibilité, et elle permet de rendre compte de la limite imposée à la vitesse des particules en référant à des changements de la valeur de deux facteurs (la valeur de l'énergie reçue par les particules et la valeur de leur vitesse) en conformité avec la conception manipulative de l'explication.<sup>53</sup>

---

<sup>53</sup> Woodward, J. (2003), p. 208-209

On pourrait objecter que (3.4.2) semble pouvoir fonder une explication compréhensible de la vitesse inférieure à celle de la lumière d'une particule émise lors de l'explosion d'une étoile. On dirait que la particule se déplace à cette vitesse parce que (3.4.2) tous les processus physiques se propagent à une vitesse égale ou inférieure à celle de la lumière. Cette explication serait néanmoins incomplète parce qu'elle n'établit pas de lien de dépendance entre un facteur cause et un facteur effet. L'avantage de (3.4.3) est qu'elle identifie le facteur (l'énergie reçue par la particule) qui influence l'effet (la vitesse de la particule) et qu'elle précise la relation entre les deux. Selon Woodward, ce genre d'exemple est une illustration «of the great difference between providing a nomologically sufficient condition for an outcome and specifying what that outcome depends on»<sup>54</sup>.

## 2) Les explications satisfaisant EXP donnent souvent une compréhension limitée

Une objection à la conception manipulative consiste à montrer que certaines explications qui satisfont cette conception ne donnent pas une compréhension approfondie des raisons pour lesquelles l'*explanandum* dépend de l'*explanans*. La distinction importante à cet égard est celle entre (i) montrer que l'*explanandum* dépend de la bonne manière d'un des facteurs cités dans l'*explanans* (ce qui est accompli par la conception manipulative), et (ii) montrer pourquoi (au sens de pour quelle raison) l'*explanandum* dépend d'un des facteurs cités dans l'*explanans* (ce que n'exige pas la conception manipulative). Woodward donne deux exemples de ce genre de situation. Premier exemple : dans le domaine biomédical, il arrive fréquemment qu'on connaisse un facteur causal menant à une certaine maladie, sans qu'on connaisse le mécanisme exact par lequel le facteur produit cette maladie. Woodward emploie l'exemple de la syphilis non traitée qui mène, dans 25 % des cas, à une forme de paralysie partielle, la parésie. On peut donc connaître une généralisation probabiliste vraie reliant la syphilis non traitée à la parésie, savoir qu'une personne particulière, Jones, est atteinte de syphilis non traitée, et expliquer correctement que la syphilis de Jones est une cause possible de sa parésie, tout en ne sachant pas par quel mécanisme la syphilis affecte le système nerveux et mène à la parésie. (Cette explication satisfait les critères de la conception manipulative parce qu'on peut comparer des populations porteuses de syphilis à des populations qui ne sont pas porteuses, et en corrigeant pour éliminer les facteurs de confusion

---

<sup>54</sup> Woodward, J. (2003), p. 209

(‘confounding factors’), déterminer qu’un changement de la propriété «porteur de la syphilis» à la propriété «non-porteur de syphilis» réduit fortement le risque de parésie.)

Le second exemple de Woodward se rattache à ce qu’il appelle les «lois phénoménologiques» (‘phenomenological laws’). Une loi phénoménologique est une généralisation qui peut faire partie d’une explication «as long as [the generalization] supports the right sorts of counterfactuals about what would happen under interventions, even if it tells us nothing about underlying mechanisms, processes or constituents».<sup>55</sup> La loi des gaz parfaits est l’exemple de loi phénoménologique retenu par Woodward :

$$(3.4.4) PV = NkT$$

Où P est la pression du gaz, V son volume, N est le nombre de molécules du gaz, k est une constante et T la température du gaz. Je signale que pour une masse particulière de gaz, N est une constante.

La loi établit des liens entre trois variables macroscopiques : la pression d’un gaz, son volume et sa température. Elle permet de déterminer par exemple, dans certaines conditions d’arrière-plan, la pression d’un gaz si on garde le volume constant et si on augmente la température. Elle permet les explications, selon la conception manipulative, puisqu’elle relie des variations contrôlées d’une variable de l’*explanans* aux variations d’un *explanandum*. Par contre elle n’informe pas les utilisateurs de l’explication à propos des mécanismes sous-jacents à la variation de la température et de la pression. Cette absence d’information sur les mécanismes sous-jacents est particulièrement apparente parce que nous disposons d’une autre théorie, la mécanique statistique, qui couvre les mêmes phénomènes en représentant le gaz comme un ensemble de molécules en mouvement, sujettes à des collisions et à des transferts de momentum et d’énergie. La mécanique statistique «tells us, in molecular terms, what pressure and temperature are, and in doing so explains why the gas laws and many other phenomenological generalizations concerning the behavior of gases hold.»<sup>56</sup> Il est clair que l’explication produite par la mécanique statistique donne des raisons plus approfondies du comportement du gaz que l’explication fondée sur la loi des gaz parfaits. Selon Woodward, certains philosophes (Salmon par exemple) ont soutenu que l’explication fondée sur la loi des gaz parfaits n’est pas une vraie explication, mais plutôt une simple description.

---

<sup>55</sup> Woodward, J. (2003), p. 221

<sup>56</sup> Woodward, J. (2003), p. 222

La position de Woodward face à cette objection est d'accepter que certaines explications sont moins approfondies que d'autres. On peut accepter, selon lui, une explication qui donne peu ou pas d'information sur les mécanismes sous-jacents à un phénomène, dans la mesure où la généralisation sur laquelle se fonde l'explication permet de répondre à certaines questions W (i.e. des questions sur ce qui se serait passé si certaines conditions initiales et/ou certaines conditions d'arrière-plan avaient été différentes), et permet dans certains cas de manipuler et de contrôler le phénomène. Ce genre d'explication moins approfondie est très fréquent dans les sciences spéciales.

Les considérations des paragraphes précédents amènent Woodward à porter le jugement suivant sur la nature de l'explication. Une conception de l'explication qui insiste sur une connaissance approfondie des composantes et des mécanismes sous-jacents à un phénomène revient à considérer qu'expliquer, c'est d'abord identifier l'ontologie correcte du phénomène. Selon la conception manipulative de l'explication, par contre, «what matters is not so much what things are intrinsically, so to speak, but rather their relational features and behavior»<sup>57</sup>. «In this sense, the account of explanation presented above is less demanding about the need for explanatory theories to have a defensible 'realistic' interpretation (in the sense of postulating only 'real' entities) and assigns a more prominent role to 'instrumental' success than many competing theories of explanation.»<sup>58</sup>

### 3) Les possibilités sérieuses

Selon la caractérisation manipulative de l'explication (EXP) (3.4.1), une condition suffisante pour qu'un raisonnement constitue une explication est qu'au moins une variation contrefactuelle (satisfaisant les critères des interventions) d'une des variables de l'*explanans* puisse être associée à une variation de l'*explanandum*, en conformité avec une généralisation invariante. Il y a toutefois des contre-exemples à cette affirmation. Nos intuitions nous amènent à rejeter des explications par ailleurs conformes à (3.4.1), mais qui ne semblent pas satisfaire une condition supplémentaire : la

---

<sup>57</sup> Woodward, J. (2003), p. 223

<sup>58</sup> Woodward, J. (2003), p. 224



variation envisagée doit être perçue comme une possibilité sérieuse. Certains des contre-exemples les plus convaincants sont des situations d'omission.

Woodward emploie la paire d'exemples suivante pour illustrer son propos. Premier exemple :

(3.4.5) Un médecin est le seul responsable de la santé du patient hospitalisé P. Les protocoles en place indiquent que si un patient dans l'état médical de P devient fiévreux, le responsable doit lui faire administrer un antibiotique, qui est disponible à l'hôpital et qui traite efficacement la fièvre dans l'état de P. Le médecin connaît bien ce protocole, mais, par erreur, omet de faire administrer l'antibiotique. P décède.

Nous expliquons le décès par l'erreur du médecin. (Dit autrement, nous jugeons que l'erreur du médecin est la cause du décès.) Cette explication est conforme à (3.4.1). Une variation de la variable «traitement par l'antibiotique» aurait, sous la généralisation affirmant l'efficacité de cet antibiotique, entraîné une variation du sort du patient. Second exemple :

(3.4.6) Une personne qui n'habite pas la ville où se trouve l'hôpital, et qui n'est pas médecin, se trouve par hasard dans la chambre du patient P au moment où la fièvre se déclare. Cette personne omet de faire administrer l'antibiotique. P décède.

On pourrait expliquer le décès par l'omission de la personne. Cette explication est conforme à (3.4.1), pour les mêmes raisons que l'explication rattachée au médecin. Mais nous rejetons cette explication, parce que nous considérons que l'administration de l'antibiotique par cette personne n'est pas une possibilité sérieuse dans les circonstances<sup>59</sup>.

Il semble, sur la base de cet exemple, qu'on doive ajouter une condition à la caractérisation des explications dans la conception manipulative. Une explication fondée sur l'examen des variations contrefactuelles appropriées doit exclure les variations que nous jugeons peu sérieuses. De plus, il semble que ce genre d'exclusion soit applicable à de nombreuses situations, comme le suggère un deuxième exemple plus grossier d'événement non-sérieux. Je tape un texte à l'ordinateur. Si une météorite avait frappé mon domicile, je n'aurais pas complété ce texte. Mais le fait qu'une météorite n'ait pas frappé mon domicile n'est pas retenu comme une cause ou une explication du texte complété, parce que nous ne considérons pas que la possibilité qu'une météorite frappe mon domicile soit sérieuse.

---

<sup>59</sup> Woodward, J. (2003), p. 87-88

Woodward accepte l'idée que les possibilités contrefactuelles que nous considérons aux fins de l'explication doivent exclure les possibilités non-sérieuses. Cela semble introduire un élément de subjectivité dans l'explication. Mais Woodward pense qu'on peut limiter cette subjectivité en subdivisant l'examen des causes et des explications en un processus à deux étapes. Dans une première étape on exclut des possibilités causales celles qui ne sont pas des possibilités sérieuses. Cette première étape contiendra des éléments subjectifs, parce que l'exclusion sera fondée en partie sur des jugements qui dépendent des intérêts des émetteurs et des récepteurs de l'explication. Mais une fois qu'il y a entente sur le champ des possibilités sérieuses retenues, Woodward est d'avis que l'examen des possibilités contrefactuelles sérieuses sera objectif.<sup>60</sup>

#### 4) Le contraste

La notion de possibilité sérieuse est souvent associée à celle de contraste, que Woodward caractérise de la façon suivante. Quand on demande une explication d'un événement  $y_1$ , on est souvent intéressé, non pas à l'ensemble des alternatives possibles à  $y_1$ , mais plutôt à une alternative particulière  $y_2$ , proche de  $y_1$ , ou encore à un groupe restreint de valeurs de  $Y$  autour de  $y_1$ . On s'intéresse donc au contraste entre  $y_1$  et  $y_2$ . À titre de premier exemple de cette situation, le cas de l'omission par un médecin de traiter un patient, mentionné dans les paragraphes précédents sur les possibilités sérieuses, peut être interprété de la manière suivante. On ne s'intéresse pas à toutes les alternatives à l'événement réel du décès du patient ( $y_1$ ), mais uniquement à des alternatives suffisamment proches des conditions qui prévalaient quand  $y_1$  s'est produit pour qu'on les considère comme des possibilités sérieuses. Ce serait par exemple ( $y_2$ ) le cas d'une personne atteinte de la même maladie, suivie elle aussi dans un milieu hospitalier par un médecin traitant. Nous restreignons le champ des alternatives à  $y_1$ , aux fins de l'explication causale, à ce que nous considérons comme des possibilités sérieuses. Nous excluons la possibilité non-sérieuse qu'un quidam, se trouvant par hasard à l'hôpital, administre un antibiotique au patient.<sup>61</sup>

L'apport des contrastes à notre comportement explicatif ne se limite pas aux cas d'omission ou aux cas d'exclusion des possibilités non-sérieuses. Les contrastes expriment souvent des choix de point

---

<sup>60</sup> Woodward, J. (2003), p. 88-91

<sup>61</sup> Woodward, J. (2003), p. 227-228

de vue qui dépendent du contexte s'appliquant à l'examen d'une situation. Woodward reprend un exemple provenant d'un livre de R. Collingwood (1940).

(3.4.7) Un accident d'automobile a lieu sur une section en courbe d'une autoroute.

Dans ce cas, «the highway engineer may be interested in accounting for the contrast between the situation in which the crash occurs and alternative situations in which similar cars moving at similar speeds and with similarly competent drivers safely traverse other curves, and may conclude that the crash was caused [...] by the curve being too tightly banked.»<sup>62</sup> Par contre un juge ou un policier s'intéressera plutôt à expliquer pourquoi tel conducteur particulier a eu l'accident alors que d'autres conducteurs, négociant la même courbe, n'ont pas eu d'accident. D'autres facteurs, tels que la vitesse de la voiture et la question de savoir si le conducteur qui a eu l'accident était en état d'ivresse deviendront plus importants. Les divers intervenants cherchent des réponses à différentes questions W, en fonction de leurs responsabilités et de leurs intérêts. Ces considérations semblent mettre en doute l'objectivité de la relation causale et de l'explication.<sup>63</sup>

Il y a deux façons d'interpréter ce genre d'exemple. D'un côté, certains interprètent le cas de l'accident d'automobile (3.4.7) en disant qu'il y a un seul événement à expliquer (l'accident) et que cet événement unique a des explications différentes selon le point de vue (ou les intérêts) des personnes qui produisent ou reçoivent les explications. Selon ce point de vue, le contexte est une composante essentielle de la relation de causalité et de l'explication. Peter Menzies, par exemple, (qui accepte par ailleurs l'approche manipulative de Woodward) caractérise comme suit la relation causale d'un événement singulier ('token') :

(3.4.8)  $X = x$  causes  $Y = y$  relative to the context  $\langle X = x', Y = y' \rangle$  if and only if (i) the actual values of  $X$  and  $Y$  are  $x$  and  $y$  respectively; and (ii) if an intervention were to change the value of  $X$  from  $x'$  to  $x$  then the value of  $Y$  would change from  $y'$  to  $y$ .<sup>64</sup>

(Dans cette caractérisation, l'expression «the context  $\langle X = x', Y = y' \rangle$  » réfère au cours normal des choses.)

Une conséquence peu attrayante de cette interprétation est que l'explication devient fortement subjective. Si l'explication dépend essentiellement du contexte (au sens du contraste que les

---

<sup>62</sup> Woodward, J. (2003), p. 230

<sup>63</sup> Woodward, J. (2003), p. 230-231

<sup>64</sup> Menzies, P. (2009), p. 360

usagers ont en tête), alors l'explication dépend essentiellement des intérêts des usagers, qui varient d'un usager à l'autre.

Woodward, pour sa part, reconnaît que les intérêts des intervenants et le contexte affectent une composante de nos explications, mais il croit qu'on peut sauvegarder l'objectivité des autres composantes. Il interprète les cas similaires à (3.4.7) en disant que le contexte et les intérêts des intervenants affectent le choix de l'*explanandum* (qui est alors contrasté), mais qu'une fois cet *explanandum* établi, le diagnostic causal et le raisonnement menant à l'*explanandum* demeurent objectifs. Dans le cas (3.4.7), il y a plusieurs (plutôt qu'un seul) événements contrastés à expliquer. (i) Pour l'ingénieur, l'*explanandum* contrasté est l'accident réel dans une courbe configurée de telle manière, vs la présence ou l'absence d'accident dans une courbe moins prononcée. (ii) Pour le juge, l'*explanandum* contrasté est l'accident dans telles circonstances de vitesse du véhicule et d'état d'ébriété du conducteur, vs la présence ou l'absence d'accident pour des conducteurs plus lents et plus sobres. Une fois le contraste choisi, la question de savoir (dans le contraste adopté par l'ingénieur par exemple) si la forme de la courbe de la route est ou non une cause de l'accident devient une question objective à évaluer en examinant certaines hypothèses contrefactuelles avec les données, les théories et les expérimentations pratiquement possibles dans les circonstances.<sup>65</sup>

##### 5) Le niveau de l'explication

Un dernier aspect controversé de l'explication porte sur le niveau de cette explication. La notion de niveau se caractérise plus facilement par des exemples que par une description abstraite. Dans le cas des explications relatives au comportement des gaz (3.4.4), on peut opposer un niveau macroscopique, qui vise les variables macroscopiques de la pression, du volume et de la température du gaz, et un niveau plus microscopique, qui vise au premier chef le comportement (en termes de position et de momentum) de chacune des molécules du gaz. De même, on peut étudier le comportement économique d'un être humain au niveau fin de l'état et du mouvement de chacune des cellules du corps, ou à un niveau plus macroscopique, en termes de ses décisions d'achat. L'observation de Woodward à cet égard est que le type de questions W auquel nous voulons répondre impose une contrainte sur le niveau de l'*explanans*, qu'il s'agisse de la

---

<sup>65</sup> Woodward, J. (2003), p.230-231

généralisation explicative retenue, du modèle théorique dont elle fait partie, ou de la finesse des entités visées. Un exemple trivial de ce type de contrainte est l'explication en économie. Si on cherche à expliquer le prix des oranges, il est évident que les modèles théoriques (imparfaits et limités) de l'offre et de la demande et des facteurs comme la température et la pluie dans les régions productrices d'oranges sont plus pertinents en pratique que des modèles fondés sur les composantes physiques et biologiques plus fines des acheteurs d'oranges. Woodward conclut à cet égard que le contraste entre les explications de haut niveau et les explications de bas niveau paraît bien différent selon qu'on adopte la conception nomologique ou la conception manipulative de l'explication. La conception nomologique privilégie les explications de relativement bas niveau, qui assurent mieux la mise en évidence de conditions nomologiquement suffisantes d'un phénomène, tandis que la conception manipulative vise davantage l'identification des alternatives au phénomène à expliquer, même si les mécanismes en jeu et leur application aux entités sous-jacentes ne sont pas connus. Ainsi la conception manipulative se contente souvent de généralisations de haut niveau.<sup>66</sup>

### 3.5 L'invariance

J'ai examiné à la section 3.4 les objectifs de l'explication, sa caractérisation et certaines observations générales, comme le fait que les explications ne fournissent pas nécessairement une compréhension approfondie, ou le fait qu'elles doivent porter, dans les comparaisons qu'elles envisagent, sur des changements que les utilisateurs considèrent comme des possibilités sérieuses. Je veux maintenant examiner l'idée que la généralisation qui est requise comme prémisse dans l'explication n'a pas besoin d'être une loi. Une thèse centrale de Woodward est que les caractéristiques requises dans cette généralisation consistent uniquement en ce qui est requis pour répondre à des questions W, ou dit de manière légèrement différente, consistent uniquement en ce qui est requis pour que la généralisation permette de soutenir les énoncés contrefactuels à l'égard de ce qui arriverait en cas d'interventions. Woodward emploie les expressions «généralisation invariante» pour désigner ces généralisations et «invariance» pour désigner ces caractéristiques.<sup>67</sup>

---

<sup>66</sup> Woodward, J. (2003), p. 233

<sup>67</sup> Woodward, J. (2003), p. 236

Ainsi, les caractéristiques d'absence d'exception (dans un domaine limité), et/ou d'application universelle dans le temps et l'espace, et/ou de ne pas référer à des lieux ou à des objets particuliers, (qu'on attribue aux lois), ne sont pas des conditions nécessaires à l'explication causale parce que l'on peut s'intéresser à, et considérer comme explicatifs, des raisonnements qui portent sur des objets ou des époques particulières, comme par exemple dans les sciences biomédicales. Je consacre la présente section à une description plus approfondie de l'invariance.<sup>68</sup>

Voici d'abord deux caractéristiques générales des généralisations invariantes. Premièrement, ce ne sont pas tous les changements imaginables qui déterminent si une généralisation est ou n'est pas invariante, mais seulement les changements à des variables (ou dit autrement à des facteurs) figurant explicitement dans la généralisation. La distinction qu'on veut effectuer est celle entre (i) les changements des facteurs causaux qui ne figurent pas dans la généralisation ou, dit autrement, qui font partie des conditions d'arrière-plan et (ii) les changements des variables figurant explicitement dans la généralisation. On veut exclure les changements aux conditions d'arrière-plan parce que même des généralisations notoirement accidentelles sont invariantes sous certains changements de conditions d'arrière-plan. Woodward donne l'exemple de la généralisation accidentelle :

(3.5.1) All coins in Bill Clinton's pocket on January 8, 1999 are dimes.<sup>69</sup>

Cette généralisation continuerait de tenir sous toutes sortes de changements de conditions d'arrière-plan, comme celui du prix du thé en Chine ce jour-là.

Deuxièmement, ce sont uniquement les changements «sous intervention de testing» des variables figurant dans la généralisation qui déterminent son statut d'invariance. J'illustre d'abord la notion d'intervention de testing par l'exemple proposé par Woodward.

(3.5.2) Un système (S) est composé d'une ampoule électrique régie par un interrupteur. La généralisation décrivant le fonctionnement de l'ampoule est : (i) Si l'angle de l'interrupteur par rapport à sa position initiale est de moins de un radian (où un radian est égal à 57 degrés), alors l'ampoule ne s'allume pas. (ii) Si l'angle de l'interrupteur se situe entre un et deux radians, alors l'ampoule s'allume.

---

<sup>68</sup> Woodward, J. (2003), p. 236-238

<sup>69</sup> Woodward, J. (2003), p. 248

Woodward examine d'abord une intervention dans le système S qui change la position de l'interrupteur de 0 à 30 degrés. Si la lumière ne s'allume pas, alors il semble évident que la généralisation (3.5.2) tient ou «reste stable» sous l'intervention. Mais cela ne suffit pas pour établir que (3.5.2) est invariante, «at least if we are trying to formulate an invariance condition that is sufficient for [3.5.2] to figure in explanations or to describe a causal relationship»<sup>70</sup>. Le fait que (3.5.2) continue de tenir sous cette intervention est compatible avec la possibilité que l'interrupteur soit brisé, si bien que l'ampoule ne s'allumerait pas même si l'interrupteur était placé dans une position égale ou supérieure à un radian. Dans le cas où l'interrupteur de S est brisé, la position à 30 degrés de l'interrupteur ne cause pas et n'explique pas le fait que l'ampoule soit éteinte. Cet exemple suggère que, pour qu'une généralisation soit considérée causale et explicative, la généralisation doit tenir sous une intervention qui change une de ses variables indépendantes de manière telle que la valeur de sa variable dépendante [i.e. la variable représentant l'effet] change également de la manière prévue par la généralisation. Dans l'exemple du système S, une intervention de testing est un changement de plus d'un radian de l'angle de l'interrupteur qui fait en sorte que, si l'interrupteur est fonctionnel, l'ampoule s'allume<sup>71</sup>.

Rassemblant les idées de paragraphes précédents, voici les caractérisations inter-reliées, proposées par Woodward, menant à celle de généralisation invariante.

### (3.5.3) Intervention de testing

Soit une généralisation G, relative à des changements, reliant une variable X à une seconde variable (d'effet) Y. Supposons que cette généralisation tient, au sens où elle décrit bien le pattern de corrélation entre X et Y, dans un certain système S. Une intervention de testing est une intervention (qui satisfait les conditions énoncées dans la caractérisation INT (2.2.10) de la section 2.2) qui change la valeur  $x_0$  que prend initialement X dans le système S, en une valeur  $x_1$ , alors que  $x_0$  et  $x_1$  sont associées, en vertu de G, à des valeurs différentes  $y_0$  et  $y_1$  de la variable (d'effet) Y<sup>72</sup>.

---

<sup>70</sup> Woodward, J. (2003), p. 249

<sup>71</sup> Woodward, J. (2003), p. 249-250

<sup>72</sup> Woodward, J. (2003), p. 250

#### (3.5.4) Généralisation invariante

G est invariante sous une intervention de testing si et seulement si elle décrit correctement la nouvelle valeur  $y_1$  que prend Y sous cette intervention.

#### (3.5.5) Relation causale générale ('type').

Condition nécessaire et suffisante pour qu'une généralisation G représente une relation causale et puisse figurer dans une explication causale :

G doit être invariante sous au moins une intervention de testing<sup>73</sup>.

Je présente maintenant diverses observations qui éclairent davantage le concept d'invariance.

Premièrement, l'invariance d'une généralisation, malgré la complexité rattachée à sa caractérisation précise, est une exigence relativement faible. Pour qu'une généralisation soit invariante, il suffit qu'elle satisfasse le test de l'intervention de testing pour un seul des facteurs qui affectent dans la réalité les phénomènes qu'elle régit. Je donnerai à la section 3.8 un exemple plus approfondi (l'exemple 3.8.2) de la faiblesse de cette exigence. De plus les observations des paragraphes suivants viendront étayer davantage ce jugement. Mais je veux souligner dès maintenant une conséquence importante de la faiblesse relative de l'exigence rattachée au concept d'invariance (et de relation causale, et d'explication causale, dans la conception manipulative). En vertu de cette faible exigence, le domaine des généralisations causales devient très étendu, et par conséquent disparate. Il y aura une grande variété de relations (diversement) causales, et il deviendra possible et intéressant de distinguer, dans l'ensemble des relations causales, divers «degrés» d'invariance et d'importance. C'est ce que je ferai à la section 3.8.

Deuxième observation. Les généralisations causales et explicatives n'ont pas besoin de s'appliquer sans exception dans leur domaine d'application, ni de s'appliquer universellement dans le temps et l'espace. Une généralisation est invariante si elle subit (ou si elle subissait) avec succès au moins une intervention de testing. Par exemple, la «loi» des gaz parfaits peut être considérée comme invariante pourvu qu'elle tienne pour certaines variations de la température d'un gaz particulier, même si elle ne tient plus dans le cas de température très élevées.

---

<sup>73</sup> Woodward, J. (2003), p. 250



Troisième observation. Une généralisation invariante n'a pas besoin d'être invariante sous tous les changements possibles. Elle est plutôt invariante sous certains changements et peut, par ailleurs, ne pas être invariante sous d'autres types de changements. Dit autrement, l'invariance d'une généralisation est relative aux changements envisagés.<sup>74</sup>

Quatrième observation. Le concept d'invariance implique qu'une généralisation doit être exprimée de façon précise. Alors que certaines caractérisations philosophiques des relations causales peuvent porter sur des énoncés très généraux comme «Le (facteur de type) C cause le (effet de type) E.», ou encore de façon légèrement plus précise «Fumer cause le cancer», le concept d'invariance exige une formulation suffisamment précise de la généralisation pour qu'on puisse associer à une valeur particulière d'un facteur (potentiellement) causal (ou encore à la probabilité d'une valeur particulière d'un facteur (potentiellement) causal), une valeur particulière de la variable effet (ou la probabilité d'une valeur particulière de la variable d'effet). Cette exigence de précision minimale laisse néanmoins place à une grande variété de types de relations causales. Woodward donne les exemples suivants de tels types de relation :

- (1) Une relation fonctionnelle simple, exprimable par une équation élémentaire, comme  $E = a C$  .
- (2) Une relation fonctionnelle déterministe plus complexe faisant intervenir plusieurs facteurs ou variables indépendantes, comme celle exprimant la force de gravitation dans la mécanique newtonienne :  $F = G m_1 m_2 / r^2$  .
- (3) Une relation probabiliste (satisfaisant certaines conditions) entre des états dichotomiques C (prenant les valeurs 0 ou 1) et E (prenant les valeurs 0 ou 1) d'une certaine entité. «If C is the only factor that is causally relevant to E, then the existence of a causal relationship between C and E may show itself in the fact that the conditional probabilities  $P(E = 1/C = 1)$  and  $P(E = 1/C = 0)$  are not equal and are invariant under some range of interventions that change the value of  $P(C)$ »<sup>75</sup>.

---

<sup>74</sup> Woodward, J. (2003), p. 243

<sup>75</sup> Woodward, J. (2003), p. 244.

Il est clair que les relations d'invariance et les interventions de testing rattachées aux types de relations (1), (2) et (3) sont substantiellement différentes. Pour d'autres exemples, voir le texte de Woodward.<sup>76</sup>

Cinquième observation. Une généralisation est invariante à l'égard d'un système particulier ou d'un type de système particulier. C'est pourquoi la caractérisation de l'invariance inclut une référence au système S où se déroule (ou se déroulerait) une intervention de testing. Il s'agit ici aussi d'une forme de relativisation du statut d'invariance. La «loi» des gaz parfaits pourrait s'appliquer à certains types de gaz et ne pas s'appliquer à d'autres types.<sup>77</sup>

Sixième observation. Même si l'exigence de stabilité requise pour l'invariance semble faible et relative, elle détermine néanmoins un seuil minimum. La généralisation doit tenir sous au moins une intervention de testing s'appliquant à une de ses variables indépendantes. Ce seuil exclut donc certaines généralisations vraies, comme par exemple les généralisations reliant deux effets corrélés d'une cause commune. Un exemple d'une généralisation exclue par les interventions de testing est celle reliant le niveau d'un baromètre à la température.

Septième observation. Quand il a été établi (via une intervention de testing) qu'une généralisation G est invariante pour un certain type de changement d'une de ses variables indépendantes, dans un certain type de système, il est possible que G tienne également à l'égard de nombreux autres changements, qu'on peut classer dans les deux catégories suivantes : (i) les changements aux variables explicitement présentes dans G mais qui ne sont pas des interventions, et (ii) les changements des variables d'arrière-plan. En fait ce sont les généralisations invariantes qui sont également stables à l'égard de ces autres types de changements qui intéressent les scientifiques et les personnes qui utilisent les relations causales pour influencer les phénomènes. Une généralisation invariante qui serait par ailleurs très sensible au moindre changement des conditions d'arrière-plan serait associée à des explications qu'on jugerait superficielles et aurait un intérêt pratique limité<sup>78</sup>. Je reviendrai à cette observation à la section 3.8 en discutant de la stabilité.

---

<sup>76</sup> Woodward, J. (2003), p. 243-245

<sup>77</sup> Woodward, J. (2003), p. 251-252

<sup>78</sup> Woodward, J. (2003), p. 253-254

### 3.6 L'invariance et les contrefactuels

Le concept d'invariance est un concept modal, au sens où il se rattache au fait qu'une relation demeurera ou non stable si ses intrants étaient possiblement, mais pas nécessairement réellement, soumis à certains changements. Les énoncés qui expriment de tels changements possibles sont des énoncés conditionnels contrefactuels, que je désigne (depuis le début de ce mémoire) plus succinctement par le terme «contrefactuel». Il y a des raisons légitimes de se méfier des énoncés contrefactuels. Leur valeur de vérité est souvent difficile à établir, par exemple dans les cas où le changement envisagé (exprimé par la condition) nous semble impossible. Un exemple provenant de Quine est celui de la paire d'énoncés :

(3.6.1) If Julius Caesar had been in charge of the U. N. Forces during the Korean War, then he would have used nuclear weapons.

(3.6.2). If Julius Caesar had been in charge of the U. N. Forces during the Korean War, then he would have used catapults.<sup>79</sup>

Il semble difficile de déterminer si l'un ou l'autre, ou possiblement aucun, des deux énoncés devrait être considéré comme vrai, et pour quelle raison.

Je vais immédiatement présenter les recommandations de Woodward, qui visent non pas les énoncés contrefactuels en général, mais spécifiquement les énoncés contrefactuels à employer pour porter des jugements sur les relations causales. Mes commentaires se subdivisent en deux questions.

Premièrement, quels sont les bons énoncés contrefactuels à employer pour élucider les assertions causales ? Selon Woodward, le cadre conceptuel de la conception manipulative nous guide vers la bonne réponse :

A manipulability framework [...] suggests that the appropriate counterfactuals for elucidating causal claims are not just any counterfactuals, but rather counterfactuals of a very special sort : those that have to do with the outcomes of hypothetical interventions. [...] It does seem plausible that counterfactuals that we do not know how to interpret as (or

---

<sup>79</sup> Woodward, J. (2003), p. 122

associate with) claims about the outcomes of well-defined interventions will often lack a clear meaning or truth value.<sup>80</sup>

Ainsi, dans le cas des énoncés (3.6.1) et (3.6.2), ce n'est pas tant le fait que nous n'avons pas les moyens techniques d'effectuer le changement envisagé dans ces énoncés qui nous empêche de les interpréter et de juger leur valeur de vérité, que le fait que «we have no clear conception of what would be involved in carrying out such an experiment»<sup>81</sup>.

À titre de contraste, Woodward propose l'exemple des tests randomisés ('randomized testing') sur des médicaments et les contrefactuels qui y sont associés. Si l'on dispose d'un échantillon important de personnes atteintes de la même maladie et si on les divise de façon aléatoire en deux sous-groupes dont l'un reçoit un médicament et l'autre ne reçoit pas le médicament, avec les bons contrôles expérimentaux, et si l'incidence de guérison dans le groupe traité est beaucoup plus élevée que dans le groupe de contrôle, alors ces faits seraient considérés comme une base fiable pour évaluer la valeur de vérité de l'énoncé contrefactuel :

(3.6.3) If those in the control group had received the drug, the incidence [...] of recovery in that group would have been much higher<sup>82</sup>.

C'est en partie parce que nous avons une conception claire des procédures expérimentales des tests randomisés et que nous jugeons qu'elles se rapprochent le plus possible de l'idéal de l'intervention (appliqué à la recherche médicale) que nous jugeons que (3.6.3) a une valeur de vérité positive.

La seconde question importante (une fois que nous avons identifié les bons contrefactuels) est celle du fondement épistémique du jugement de valeur de vérité sur un contrefactuel où on n'a pas la possibilité de procéder à des expérimentations. Un exemple de cette situation est le suivant. On envisage les conséquences d'un changement important de la distance entre la Terre et la Lune. Il est impossible de procéder expérimentalement. Nous avons néanmoins le sentiment que les diverses lois de Newton sur la gravitation et la mécanique nous permettraient d'évaluer correctement les conséquences de ce changement hypothétique. La raison de cette confiance est le fait que les lois de Newton pour ce type de phénomène ont été amplement confirmées. Cet

---

<sup>80</sup> Woodward, J. (2003), p. 122

<sup>81</sup> Woodward, J. (2003), p. 122

<sup>82</sup> Woodward, J. (2003), p. 222

exemple suggère que l'autre élément à prendre en compte dans l'évaluation de la vérité des contrefactuels est tout simplement la fiabilité confirmée de la ou des généralisations qui fondent la conclusion de l'énoncé contrefactuel.

Je me suis concentré ici sur les conclusions de Woodward à propos de l'emploi de contrefactuels dans la clarification des assertions causales. Je m'abstiens de discuter davantage de cette question parce que cela nous entraînerait trop loin. Le lecteur intéressé pourra entre autres consulter Woodward J. (2000), pages 235 à 239.

### 3.7 Réévaluation de certains problèmes des théories nomologiques de l'explication

Nous disposons maintenant des éléments conceptuels permettant de réexaminer brièvement les difficultés rattachées aux conceptions nomologiques de l'explication. Dans la perspective manipulative, la caractéristique essentielle d'une généralisation explicative et causale n'est pas l'absence d'exceptions à la généralisation dans son domaine d'application, ni encore le fait que la généralisation s'applique universellement dans l'espace et dans le temps. La caractéristique essentielle est plutôt le fait que la généralisation soit (ou ne soit pas) invariante pour au moins un des facteurs indépendants (variables indépendantes) qu'elle contient. On évalue l'invariance à l'aide d'énoncés contrefactuels faisant référence à des interventions. C'est pourquoi on dit que c'est la caractéristique de fonder les contrefactuels qui est la caractéristique la plus importante de la liste (3.3.2) des aspects ontologiques des lois ou des généralisations explicatives. Ainsi, les exigences à satisfaire pour qu'une généralisation soit considérée explicative sont beaucoup moins fortes dans la conception manipulative que dans les conceptions nomologiques : il suffit que la généralisation soit invariante.

Ceci ne signifie pas toutefois que la notion de loi soit devenue inutile. La notion de loi demeure pertinente comme étiquette qui identifie les généralisations qui sont invariantes sous une grande variété de changements. Quand une généralisation est une loi, alors elle est particulièrement susceptible de servir aux humains pour manipuler et contrôler la nature. Les changements visés comprennent non seulement les interventions sur les facteurs/variables explicitement présents

dans les généralisations, mais aussi (i) les changements de ces variables qui ne sont pas des interventions et (ii) les changements des facteurs d'arrière-plan. Woodward parle dans ce contexte de «degrés d'invariance». J'aborderai la notion de degrés d'invariance de façon plus approfondie à la section 3.8.

J'aborde maintenant de façon plus détaillée les difficultés posées par les conceptions nomologiques mentionnées à la section 3.3. Je discuterai plus particulièrement la question de savoir si la conception manipulative de l'explication et de la causalité échappe à ces difficultés.

Selon la première difficulté (rattachée au modèle DN), on est incapable de distinguer, sur la base des seules caractéristiques formelles, entre les énoncés en forme de loi véritablement explicatifs et les énoncés en forme de loi qui ne sont pas véritablement explicatifs. La conception manipulative échappe à cette difficulté parce qu'elle ne tente pas de déterminer le statut explicatif d'une généralisation sur la base de la seule forme de cette généralisation. C'est plutôt une caractéristique du monde, à savoir le fait que la généralisation tienne face à des tests réels ou hypothétiques (les interventions de testing) qui détermine le statut causal et explicatif de la généralisation. Dans le cas assez fréquent où il est impossible d'effectuer une intervention réelle, c'est-à-dire lorsqu'on doit juger de la vérité d'un énoncé contrefactuel sans pouvoir effectuer réellement une intervention de testing, notre jugement peut se fonder sur (i) la clarté du processus qui aurait lieu dans une intervention hypothétique et (ii) sur la fiabilité de la généralisation qui fonde la conclusion de l'énoncé contrefactuel.

Selon la deuxième difficulté (rattachée au modèle DN), les généralisations considérées comme explicatives dans la plupart des sciences ne satisfont pas aux exigences des caractéristiques formelles des énoncés en forme de loi. Les généralisations des sciences biologiques et médicales et des sciences humaines ont souvent des exceptions et ne s'appliquent pas universellement dans l'espace et le temps. Elles contiennent par exemple des références à des lieux (la planète Terre) et à des époques particulières. La conception manipulative des causes et des explications n'est pas touchée par cette difficulté puisqu'elle accepte certaines généralisations causales qui ont des exceptions et qui ne s'appliquent pas universellement.

Selon la troisième difficulté (rattachée au modèle DN), un grand nombre d'explications de la vie quotidienne (portant sur tous types de phénomènes, y compris des phénomènes relevant de la physique et de la chimie) semblent fonctionner sans référence explicite à des lois, et l'on peut émettre et recevoir des explications sans connaître les lois qui (peut-être) régissent ces phénomènes. La conception manipulative de l'explication est elle aussi vulnérable à certaines parties de cette difficulté, parce qu'elle contient une référence à des généralisations invariantes explicites. Mais la conception manipulative offre davantage de flexibilité pour concilier les explications de tous les jours aux exigences de cette conception. Cette flexibilité provient du fait que la conception manipulative reconnaît les explications faites au niveau de l'*explanandum* (ou, dit autrement dans les termes employés dans l'*explanandum*). Par exemple, on pourrait, au moins conceptuellement, appliquer la méthode expérimentale à une situation comme celle de la table et de la bouteille d'encre (3.3.5). On pourrait étudier le comportement d'une table et d'une bouteille d'encre particulières qu'on soumettrait à divers types de coups pour déterminer comment ces éléments réagissent à des coups de force diverse. On pourrait alors établir une généralisation reliant les données sur le poids et la structure de la table, sur la position et les autres caractéristiques de la bouteille, sur la force des coups portés, à des résultats en termes de déplacement ou renversement de la bouteille. Cette généralisation serait peu invariante (au sens où elle dépendrait d'un grand nombre de caractéristiques difficiles à transférer à des tables et à des bouteilles de forme ou de structure différentes) et elle aurait peu d'intérêt pratique. Mais elle constituerait une explication valable selon la conception manipulative.

Je signale par ailleurs qu'il y a de nombreuses situations où le comportement détaillé d'un objet de forme et de structure particulières est d'un grand intérêt, comme par exemple le cas de la réaction des ailes de forme particulière d'un nouveau modèle d'avion à divers types de chocs. L'approche manipulative, en ne se limitant pas à l'idée de rattacher les résultats causaux à des lois de la nature, rend mieux compte du phénomène de ces explications de «haut niveau».

Selon la quatrième difficulté (rattachée au modèle DN), dans de nombreux cas d'explication scientifique, les explications fonctionnent sans référence à aucune loi connue explicitement. Les tenants des conceptions nomologiques postulent alors l'existence d'une ou plusieurs lois, sans pouvoir les énoncer. C'est ce que Woodward appelle «la stratégie de la structure cachée». Les

conceptions nomologiques se trouvent alors dans une posture précaire. Qu'est-ce qui assure que cette loi ou relation non spécifiée existe ?

La conception manipulative, par contre, n'exige pas la présence d'une loi ou la compréhension d'un mécanisme sous-jacent au phénomène à expliquer. Dans cette perspective, il suffit de montrer que, au niveau même des variables présentes dans une généralisation proposée, l'effet sur l'*explanandum* de variations par intervention des variables de l'*explanans* soit tel que la généralisation tienne pour au moins une des variables de l'*explanans*. Par exemple, dans un essai clinique bien conduit, si la généralisation à examiner affirme que la maladie réagit possiblement au médicament  $X_1$ , il suffit, pour établir le lien causal entre le médicament  $X_1$  et la guérison, que la prise du médicament  $X_1$  soit rattachée à des changements de l'état de santé de ceux parmi les malades qui ont reçu le médicament  $X_1$ . Il n'est pas requis de comprendre en vertu de quel mécanisme le médicament  $X_1$  met fin à la maladie. La conception manipulative de l'explication et de la causalité peut ainsi être qualifiée d'«opérationnelle» (au sens où elle s'intéresse aux relations entre les phénomènes, même si le fondement de cette relation demeure obscur) plutôt qu'«ontologique» (au sens où le mécanisme sous-jacent d'une relation doit être connu pour permettre l'explication).<sup>83</sup>

Selon la cinquième difficulté (rattachée au modèle DN), l'explication est souvent asymétrique au sens où un élément de l'*explanans* (en association avec la loi contenue dans l'*explanans*) explique l'*explanandum*, tandis que l'*explanandum* n'explique pas cet élément de l'*explanans*. Par contre, dans le modèle DN de l'explication, pour les mêmes *explanans* et *explanandum*, l'*explanans* et l'*explanandum* se trouvent souvent dans une relation symétrique, au sens où l'*explanans* (en association avec la loi) permet de déduire l'*explanandum* et, inversement, l'*explanandum* (en association avec la loi) permet de déduire l'*explanans*.

La conception manipulative de l'explication et de la causalité correspond à nos intuitions à cet égard. Selon cette conception, ce qui compte est la capacité de modifier l'*explanandum* en faisant des interventions sur l'*explanans*. Or cette capacité est asymétrique lorsque le lien causal est asymétrique. Dans l'exemple (3.3.9) du poteau et de l'ombre projetée par ce poteau, on peut faire

---

<sup>83</sup> Woodward, J. (2003), p. 223



varier la longueur de l'ombre projetée en faisant varier par intervention (i.e. de façon indépendante) la hauteur du poteau (ainsi, la hauteur du poteau cause la longueur de l'ombre), mais on ne peut pas faire varier par intervention (de façon indépendante) la longueur de l'ombre pour causer une variation de la hauteur du poteau. Pour tenter d'intervenir sur le facteur potentiellement causal (l'ombre), il faut absolument faire varier la hauteur du poteau.

Je vais maintenant discuter les difficultés rattachées au modèle IS (pour 'Inductive Statistical') proposé par Hempel. (Je rappelle ce qui a été expliqué à la section 3.2 : le terme «inductive» dans le présent contexte ne signifie pas que le raisonnement employé est inductif dans le sens habituel du terme, mais plutôt que la conclusion par déduction du raisonnement explicatif, lorsque la loi ou la généralisation de l'*explanans* est probabiliste et que l'*explanandum* est un événement singulier, n'est pas certaine, comme cela aurait été le cas dans un raisonnement déductif.) Je vais d'abord discuter la difficulté 7 de la section 3.2, pour ensuite discuter la difficulté 6, parce que la difficulté 7 est plus générale et clarifie l'ensemble des explications probabilistes.

Selon la septième difficulté des conceptions nomologiques (rattachée dans ce cas au modèle IS), deux facteurs A et B peuvent être reliés par une généralisation statistique à haute probabilité (comme l'exige le modèle IS), mais ne pas pouvoir figurer comme cause et effet dans une explication causale. Je rappelle l'énoncé détaillé de la difficulté. Il y a au moins cinq situations différentes qui donnent lieu à une corrélation statistique positive entre les variables A et B, et seule une des cinq en est une où A est une cause de B, et peut donc figurer dans une explication correcte par le facteur A du phénomène B.

(3.3.10) Cinq situations donnant lieu à une corrélation statistique entre A et B

- 1) A est une cause de B
- 2) B est une cause de A
- 3) A et B sont tous deux des effets d'une cause commune Z, comme c'est le cas dans l'exemple du baromètre (2.1.2)
- 4) A et B sont des facteurs (des variables) reliées non pas causalement, mais plutôt conceptuellement.
- 5) A et B sont des facteurs (des variables) dont le lien est accidentel.

Mon objectif est d'indiquer pourquoi la conception manipulative permet de distinguer la situation 1, où le facteur A est une véritable cause de l'*explanandum* B. Pour rendre ces indications un peu plus concrètes, je vais me concentrer sur une forme simple de généralisation statistique, une régression linéaire à deux variables indépendantes<sup>84</sup> :

$$(3.7.1) Y = a_0 + a_1X_1 + a_2X_2 + U$$

Une situation concrète d'application de cette équation serait celle d'un champ où poussent des plants de maïs. La variable Y (*l'explanandum*) représente la hauteur en centimètres de chaque plant individuel. Les deux facteurs causaux présents dans la généralisation sont la quantité d'eau, en litres, (représentée par  $X_1$ ) reçue par chaque plant et la quantité d'engrais ( $X_2$ ) reçu par chaque plant. Les éléments  $a_0$ ,  $a_1$  et  $a_2$  de l'équation sont des constantes, s'appliquant de façon invariable à tous les plants : ainsi  $a_1$  pourrait prendre la valeur 3 et signifier que chaque litre d'eau supplémentaire reçue par un plant particulier augmente la hauteur de ce plant de 3 centimètres. On suppose que les quantités d'eau ou d'engrais reçus ont varié d'un plant à l'autre, et la forme d'addition linéaire de la généralisation statistique nous dit trois choses principales :

- (i) La hauteur des plants est une fonction, entre autres, de la quantité d'eau et de la quantité d'engrais reçus.
- (ii) Les deux facteurs présents explicitement dans la généralisation affectent la hauteur de façon indépendante. L'impact sur la hauteur d'un plant de l'ajout d'un litre d'eau est de 3 centimètres, indépendamment de la quantité d'engrais reçu.
- (iii) Mais il y a d'autres facteurs qui affectent la hauteur de chaque plant, comme l'ensoleillement, ou la position particulière d'un plant dans le champ, qui ne sont pas pris en compte explicitement dans l'équation et dont l'impact est représenté par la variable «d'erreur» U.

Un des apports des statistiques dans ce genre de situation est de permettre d'examiner les différentes valeurs que prennent les variables Y,  $X_1$ ,  $X_2$  et U pour chacun des plants dans le but de déterminer si les facteurs  $X_1$  et  $X_2$  rendent compte d'une grande partie de l'effet Y, ou si au contraire,  $X_1$  et  $X_2$  ne rendent compte que d'une petite partie de Y et si alors la variable d'erreur U

---

<sup>84</sup> Cette simplification et l'exemple associé sont suggérés dans Woodward, J. (1999), p. 199 sqq.

prend des valeurs élevées. La mesure de pouvoir descriptif de  $X_1$  et  $X_2$  à l'égard de l'effet  $Y$  est appelée le coefficient de détermination<sup>85</sup>.

Woodward souligne qu'une équation comme (3.7.1) peut être interprétée de plusieurs manières qu'il convient de distinguer<sup>86</sup>. Il distingue (i) une interprétation descriptive, (ii) une interprétation prédictive et (iii) une interprétation causale. (3.7.1) peut constituer une simple description des plants d'un champ de maïs particulier. Dans notre exemple, la description donne certaines informations limitées sur la hauteur des plants qui s'y trouvent. Si le facteur  $U$  respecte certaines conditions, alors on peut penser que l'équation et les données sur la valeur moyenne de l'eau et de l'engrais reçus nous informeraient sur la hauteur moyenne des plants de maïs.

Les relations de forme (3.7.1) peuvent également être interprétées de façon prédictive. Si les paramètres de la relation (3.7.1) ont été estimés en observant les hauteurs des plants de maïs d'un champ particulier, alors l'utilisation prédictive de (3.7.1) consiste en l'application de l'équation pour prédire la hauteur des plants dans d'autres champs. Dans la mesure où des relations de régression linéaire peuvent faire intervenir des facteurs qui ne sont pas en relation causale, il peut y avoir, sur la base d'une corrélation non causale, une application prédictive correcte d'une équation comme (3.7.1). Un exemple d'une telle situation serait une reprise un peu plus complexe de l'exemple du baromètre et du climat. Si nous disposons de plusieurs baromètres d'un même modèle dispersés géographiquement dans une région, on pourrait prédire le climat dans la région en interprétant une relation établie antérieurement entre la hauteur moyenne ( $H_M$ ) de la colonne de mercure de ces baromètres et un indice continu de température ( $T$ ) :

$$(3.7.2) \quad T = a_0 + a_1 H_M + U$$

(3.7.2) serait une relation prédictive mais non causale.

Enfin les relations de forme (3.7.1) peuvent être interprétées de façon causale. Ce qui distingue l'interprétation causale est que l'on s'attend à ce que, si l'on intervient (au sens particulier de ce terme dans la conception manipulative) sur la ou les variables indépendantes de la relation, alors la valeur de l'*explanandum* va être affectée de la manière prévue dans la relation de régression. Ainsi dans l'exemple des plants de maïs, on peut dire que nous avons interprété la relation (3.7.1) de

---

<sup>85</sup> Hamburg, Morris (1970), p. 486-492

<sup>86</sup> Woodward, J. (2003), p. 318-321

façon causale lorsque nous avons dit qu'une augmentation par intervention de 1 litre de l'apport en eau aux plants de maïs produirait une augmentation de 3 centimètres de la hauteur des plants. Par contre, la relation (3.7.2) de corrélation statistique entre la hauteur moyenne de la colonne de mercure et la température n'est pas causale parce qu'une variation par intervention pour augmenter de façon égale la hauteur de chacune des colonnes de mercure ne produira pas la variation de température prédite par (3.7.2). De façon plus générale, les diverses idées de la conception manipulative de la causalité s'appliquent aux relations de régression linéaire et autres relations de dépendance statistique. Selon la conception manipulative, on admet qu'une relation est causale si elle tient ou tiendrait dans le contexte d'une expérimentation (intervention) réelle ou hypothétique sur au moins une des variables indépendantes de la relation. Le domaine d'application de la relation peut être limité, elle peut avoir des exceptions dans son domaine d'application, qui seraient capturées par la variable d'erreur  $U$ . Les conditions particulières que doit satisfaire la variable d'erreur  $U$  pour qu'une relation de régression linéaire soit considérée comme causale (ou, dit autrement, invariante) sont spécifiées par Woodward, mais elles dépassent le cadre du présent mémoire<sup>87</sup>.

Nous pouvons maintenant expliquer pourquoi la conception manipulative permet de distinguer les véritables relations causales de trois des quatre autres cas de corrélation statistique entre des facteurs A et B. La caractéristique distinctive est justement celle de la capacité de relier A et B par une relation statistique qui tient (reste stable) sous au moins certaines interventions sur le facteur A. Il est ainsi possible que B soit un facteur causal de A (sans toutefois que A soit un facteur causal de B), que A et B soient des effets d'une cause commune, ou que A et B soient corrélés de façon purement accidentelle. Dans chacun de ces cas, une intervention sur A ne donnerait pas la variation escomptée de la variable B.

J'aborde finalement le cas où A est corrélé à B parce que A et B sont reliés conceptuellement. Dans ce cas, ce ne sont pas des interventions sur A qui déterminent le statut causal de A à l'égard de B. L'idée même d'intervention pour modifier A pour tester son effet sur B perd son sens parce que A et B sont toujours reliés entre eux. C'est donc la connaissance du lien conceptuel qui permettra aux

---

<sup>87</sup> Voir Woodward, J. (1999), p. 201 et 210-211

utilisateurs de statuer sur le statut non causal de lien entre A et B. L'application des méthodes expérimentales ne peut remplacer l'analyse des concepts en jeu.

Selon la sixième difficulté (rattachée au modèle IS), le modèle IS ne permet pas de rendre compte de l'explication de phénomènes singuliers à probabilité faible. Les explications en forme IS se limitent aux phénomènes rattachés à une loi probabiliste à probabilité élevée. Cela découle du fait que Hempel conçoit l'objectif de l'explication comme celui de montrer que le phénomène à expliquer était attendu en vertu d'une loi ('was to be nomologically expected'), et que le rôle de la loi est d'entraîner, par inférence déductive mais avec un résultat incertain, le phénomène à expliquer. Seules les lois à probabilité élevée permettraient d'entraîner correctement une telle inférence. Voici l'exemple qu'emploie S. Psillos pour illustrer cette difficulté. Soit la loi probabiliste selon laquelle le tir à pile ou face d'une pièce de monnaie fait en sorte que sur un grand nombre de tirs, la probabilité d'obtenir «face» est de 50%. Cette loi est très fiable sur un grand nombre de tirs, mais elle ne permet pas de prédire que le prochain tir donnera «face», ni d'expliquer pourquoi le dernier tir effectué a donné «face». Ainsi, si expliquer consiste à montrer qu'un événement singulier était attendu, les situations régies par des lois statistiques à relativement faible probabilité ne permettent pas l'explication des événements singuliers.<sup>88</sup>

La conception manipulative propose un changement de perspective. Expliquer ne consiste pas montrer qu'un événement passé était à être attendu. Expliquer c'est plutôt montrer les facteurs, s'il y en a, qui auraient fait en sorte que l'événement passé aurait pu être différent, de façon fiable ou contrôlable. Ainsi les explications manipulatives, tout autant que les explications nomologiques, du résultat «face» du dernier lancer de monnaie sont impossibles à donner, parce que l'objectif même du processus de lancer une pièce de monnaie est de donner des résultats singuliers incontrôlables. Mais la conception manipulative indique que dans de nombreux cas, la limite que place Hempel aux explications en vertu de lois statistiques est à rejeter. On peut donner des explications de phénomènes statistiques de faible probabilité à condition de renoncer à l'objectif d'expliquer un résultat singulier, et de se concentrer plutôt sur les facteurs qui affectent les probabilités des phénomènes, même si cette probabilité est faible. L'exemple favori de Woodward est celui de l'influence causale de la syphilis non traitée sur la parésie. 25 % des porteurs de ce

---

<sup>88</sup> Psillos, S. (2002), p. 244

genre de syphilis développent la parésie. L'identification par intervention de facteurs qui affecteraient même minimalement ce pourcentage et les explications qui en découlent font partie des activités des chercheurs et une conception correcte de l'explication devrait en rendre compte.

### 3.8 Les degrés d'invariance

La conception manipulative de la causalité et de l'explication ouvre une perspective différente sur les relations causales du monde réel. Dans l'une ou l'autre des conceptions nomologiques, la question principale est de savoir si une relation sous étude dans le monde réel est ou n'est pas causale. La perspective manipulative subdivise cette question en deux composantes : (1) La relation sous étude est-elle ou non causale ? , et (2) Dans quelle mesure une relation admise dans la classe des relations causales peut-elle être employée pour contrôler le phénomène visé ? Dit autrement, quel est le pouvoir causal de la relation sous étude ? Dans le vocabulaire de Woodward, quel est le «degré d'invariance» de la relation causale ? À titre de contraste, cette idée de degré d'invariance ou de pouvoir causal d'une relation ne peut pas facilement être envisagée dans un cadre nomologique parce que la notion même de loi est dichotomique : une relation est une loi, ou bien elle ne l'est pas. Dans les paragraphes qui suivent, je procéderai à un réexamen de la notion de relation causale, et je rendrai compte de trois aspects sous lesquels on peut examiner les relations causales pour leur attribuer des degrés d'invariance.

Je rappelle la caractérisation de Woodward de la notion renforcée de relation causale et explicative. (3.5.5) Une généralisation G représente une relation causale et peut figurer dans une explication causale si et seulement si G est invariante sous au moins une intervention de testing. Les descriptions de généralisation invariante (3.5.4) et d'intervention de testing (3.5.3) complètent la caractérisation (3.5.5). Le fait que cette caractérisation soit complexe et assez technique occulte quelque peu le fait que la caractérisation est très peu exigeante. Pour qu'une relation soit causale dans cette conception, il suffit que dans au moins un cas, une variation (qui est une intervention) de la valeur d'un des facteurs de l'*explanans* figurant dans la généralisation soit associée à une variation de l'*explanandum* (ou puisse y être associée dans le cas d'interventions physiquement impossibles à effectuer). L'expression «dans au moins un cas» signifie ici qu'il existe au moins un

ensemble de circonstances d'arrière-plan tel qu'une (seule) intervention sur un facteur figurant dans l'*explanans* provoque un changement de l'*explanandum*. Woodward rassemble ces idées dans une caractérisation plus compacte de la relation causale générale ('type') renforcée :

(3.8.1) Relation causale générale ('type') (représentée par le symbole (M))

(M) : «X causes Y if and only if there are background circumstances B such that if some (single) intervention that changes the value of X (and no other variable) were to occur in B, then Y or the probability distribution of Y would change.»<sup>89</sup>

(Ici X et Y sont des facteurs ou des variables qui peuvent prendre plusieurs états ou plusieurs valeurs, les circonstances d'arrière-plan sont les circonstances qui ne sont pas explicitement représentées dans la relation G entre X et Y, et les interventions ont le sens donné dans (3.5.3)).

(On pourrait légitimement se demander si les caractérisations du groupe (3.5.3) à (3.5.5), d'une part, et (3.8.1), d'autre part, sont logiquement équivalentes. Je n'ai pas fait l'étude de cette question, mais je remarque un élément de similitude important. Alors que (3.5.3) à (3.5.5) réfèrent explicitement à une généralisation G et aux variables qu'elle contient, (3.8.1) fait plutôt référence aux circonstances (ou facteurs) d'arrière-plan B. Mais le fait de pouvoir distinguer dans (3.8.1) entre les facteurs d'arrière-plan et les facteurs qui ne sont pas d'arrière-plan (c'est-à-dire ceux qui figurent dans une généralisation G reliant X à Y) implique l'existence et la connaissance de G. J'en conclus également que 3.8.1 fait partie des concepts causaux renforcés.)

Woodward donne l'exemple suivant de relation causale relativement peu exigeante. (Woodward a trouvé cet exemple chez Richard Dawkins).

(3.8.2) On imagine qu'un gène R a une forme normale *r* et une forme anormale *r\**. Les porteurs de la forme anormale *r\** n'apprennent pas à lire (parce qu'ils sont dyslexiques), tandis que les porteurs de la forme normale apprennent à lire (si certaines conditions d'arrière-plan, comme un système scolaire fonctionnel, sont présentes).

Si on fait l'hypothèse «that intervening to change the normal form *r* to the variant *r\** (or vice versa) is associated (again in appropriate background circumstances) with changes in whether its

---

<sup>89</sup> Woodward, J. (2010), p. 290

possessor learns to read, R will count as a gene that causes reading, according to [(3.8.1)].»<sup>90</sup>. Dans ce cas, la relation causale entre R et la lecture nous semble néanmoins faible ou douteuse parce qu'on a l'impression que beaucoup d'autres facteurs importants doivent être présents pour que le changement de forme du gène R ait un effet sur la performance de lecture. De plus on peut se demander si le gène R a un effet seulement sur la lecture, par opposition à un effet plus général sur un ensemble d'autres performances intellectuelles. Le choix de la lecture comme effet de la relation causale étudiée est donc également problématique.

Woodward fait également remarquer que la caractérisation (3.8.1) de la relation causale générale est peu exigeante parce qu'elle dit peu de choses (à part l'exigence qu'ils soient des interventions) sur la nature des changements de X qui doivent être associés à des changements de l'*explanandum* (Y), et parce qu'elle ne contient aucune exigence relativement aux conditions d'arrière-plan sous lesquelles le changement du facteur X doit provoquer un changement du facteur Y<sup>91</sup>.

Si (3.8.1) (et plus généralement la conception manipulative renforcée de la relation causale) est bien une caractérisation peu exigeante, alors il semble que le monde réel pourrait receler un très grand nombre et une très grande variété de telles relations causales, et qu'il y aurait lieu de les examiner pour tenter de trouver des sous-groupes dans la classe étendue de ces relations. C'est pourquoi Woodward distingue deux projets :

- (1) Celui de caractériser la classe des relations causales en général, pour les distinguer des relations non-causales. C'est le projet de la plupart de travaux philosophiques en causalité.
- (2) Celui de formuler des critères permettant de distinguer des sous-groupes à l'intérieur de la classe des relations causales. Ce second projet est moins fréquent en philosophie.

Un objectif du second projet est de fournir les outils conceptuels permettant de distinguer les relations causales les plus susceptibles de faciliter le contrôle des phénomènes. Un second objectif du second projet est de faciliter la discussion et la classification des relations causales complexes qu'on trouve notamment dans les sciences biologiques et dans les sciences sociales.<sup>92</sup>

---

<sup>90</sup> Woodward, J. (2010), p. 290

<sup>91</sup> Woodward, J. (2010), p. 290-291

<sup>92</sup> Woodward, J. (2010), p. 287-289



Je présente maintenant, suivant Woodward dans son article de 2010, trois critères de classification à l'intérieur de la classe générale des relations causales : (1) la stabilité, (2) la proportionnalité (aussi appelée le choix du niveau approprié d'explication), et (3) la spécificité.

### 1) La stabilité

Voici d'abord une caractérisation de la stabilité.

#### (3.8.3) Stabilité d'une relation causale

Si une relation G est causale au sens de (3.8.1), elle est invariante sous au moins une intervention sur une des variables (X) figurant explicitement dans G, dans au moins un ensemble de circonstances d'arrière-plan  $B_i$ .

«The *stability* of this relationship of counterfactual dependence has to do with whether it would continue to hold in a range of other background circumstances  $B_k$  different from the circumstances  $B_j$ .»<sup>93</sup>

La stabilité d'une relation causale est plus grande si (i) elle tient dans un éventail plus étendu de circonstances d'arrière-plan, ou (ii) si elle tient à l'égard des variations de circonstances d'arrière-plan jugées particulièrement importantes dans la discipline visée. Une relation causale est moins stable dans la mesure où elle serait perturbée par (la plupart) des changements des circonstances d'arrière-plan. La stabilité ou l'instabilité relative d'une relation causale est donc une affaire de degrés. De plus, la même relation peut être plus stable dans un premier ensemble de circonstances d'arrière-plan, et moins stable dans un second ensemble.

J'illustre maintenant la stabilité avec des exemples, en commençant par un exemple de relation causale très peu stable, produit originellement par David Lewis et repris par Woodward.

«Lewis writes a letter of recommendation L that causes X to get a job she would not otherwise have got. This in turn has various other effects : X meets and marries a colleague she would not have married if she had not taken the job, they have children and grandchildren that would not exist in the absence of Lewis' letter, these grandchildren do various things A and so on. Now consider the following claim :

(3.8.4) Lewis' writing the letter L caused X's grandchildren to exist and to do A.»<sup>94</sup>

---

<sup>93</sup> Woodward, J. (2010), p. 291-292

Il semble bien que (3.8.4) soit vraie en vertu de (3.8.1). Dans les circonstances telles que décrites, les assertions :

(3.8.5) Lewis a rédigé la lettre et les petits-enfants de X sont les personnes T et U et ces personnes ont fait les actions A.

(3.8.6) Si Lewis n'avait pas rédigé la lettre, les petits-enfants de X n'auraient pas été les personnes T et U et elles n'auraient pas effectué les actes A.

sont vraies. Mais nous avons l'intuition que le lien causal exprimé dans (3.8.4) est faible et à la limite trompeur. Lewis et Woodward attribuent ce jugement de faiblesse au fait que les deux assertions (3.8.5) et (3.8.6) sont vraies dans les circonstances du monde réel ('actual world'), mais que si ces circonstances avaient été légèrement différentes, au moins une des assertions aurait été fausse. Si par exemple l'époux de X n'avait pas décroché un emploi à la même institution que X, alors les effets de la lettre de Lewis auraient été différents. Le fait que l'assertion causale (3.8.4) ne tienne pas sous des changements relativement légers des circonstances d'arrière-plan en fait une instance de relation causale très peu stable, et provoque le jugement de faiblesse<sup>95</sup>.

Je passe maintenant à une comparaison entre deux assertions de causalité en génétique humaine, dont l'une semble clairement plus stable que l'autre. Je reprends l'exemple (3.8.2) du gène R qui, en vertu de la caractérisation peu exigeante (3.8.1), est une cause de la performance de lecture d'un individu. Cette relation causale semble relativement peu stable parce qu'elle dépend de nombreuses circonstances favorables dans l'environnement extérieur au corps de l'individu, comme le fait qu'il vive dans une culture qui connaît l'écriture, ou qu'il ait accès à des écoles. A titre de contraste, Woodward propose

(3.8.7) La relation entre le chromosome Y et les organes sexuels externes des individus.

La relation (3.8.7) est causale en vertu de (3.8.1), puisque la présence ou l'absence du chromosome Y est très fortement liée à des variations des organes sexuels externes. Elle est également plus stable que celle de l'exemple (3.8.2) parce qu'elle est généralement peu ou pas affectée par la plupart des circonstances de l'environnement extérieur au corps de l'individu. Cet exemple mène à une des deux caractérisations un peu plus précises proposées par Woodward de la notion d'être «plus stable». Une relation causale G est plus stable qu'une relation causale H si l'ensemble des conditions d'arrière-plan  $B_H$ , sous lesquelles H tient est un sous-ensemble entièrement contenu dans

---

<sup>94</sup> Woodward, J. (2010), p. 292

<sup>95</sup> Woodward, J. (2010), p. 292

l'ensemble des circonstances d'arrière-plan des circonstances d'arrière-plan  $B_g$  sous lesquelles G tient. Il s'agit d'un critère d'extension. En vertu de cette caractérisation, la relation entre le chromosome Y et les organes sexuels externes est plus stable que celle entre le gène R et la lecture.<sup>96</sup>

Mais il y a aussi des cas où la stabilité à l'égard de certaines circonstances d'arrière-plan est particulièrement importante dans un domaine de recherche donné. Dans ce genre de cas, la pertinence théorique l'emporte sur l'extension. Woodward donne un exemple rattaché à la microéconomie. Dans cette discipline, on suppose que les agents économiques agissent conformément aux généralisations contenues dans la théorie du choix rationnel, qui comprend entre autres l'hypothèse que les choix des agents sont intéressés au sens où ils servent à maximiser le revenu ou la richesse personnelles de ces agents. Il est clair que ces généralisations ne tiennent pas sous certaines circonstances d'arrière-plan. Si les choix des agents affectent leurs proches parents, l'hypothèse de maximisation du revenu et de la richesse est souvent mise de côté. De même, certains changements par intervention affectant la santé ou l'état mental des agents invalideraient les hypothèses de la théorie du choix rationnel. Mais l'échec de la stabilité dans ces circonstances n'est pas considéré comme important par les chercheurs en microéconomie. Ces chercheurs considèrent par contre que pour que des généralisations fondamentales comme les principes de la théorie du choix rationnel soient prises au sérieux en microéconomie, il faut qu'elles soient invariantes «under changes in information available to economic agents or under changes in their beliefs and under changes in the incentives or relative prices they face»<sup>97</sup>. C'est parce que les principes de la théorie du choix rationnel sont invariants sous ces changements qu'ils sont acceptés par les chercheurs en microéconomie.

La notion de stabilité a un lien évident avec la notion de loi de la nature. Les lois de la nature paradigmatiques en physique tiennent sous pratiquement toutes les circonstances d'arrière-plan. La loi de la gravitation n'est pas affectée par la présence de la vie sur Terre, ni par les croyances ou les comportements des humains. Comme je l'ai mentionné à de nombreuses reprises, les tentatives de faire entrer les généralisations des sciences biomédicales et des sciences sociales dans le moule des lois de la nature semblent vouées à l'échec. Un avantage des concepts d'invariance et

---

<sup>96</sup> Woodward, J. (2010), p. 292-293

<sup>97</sup> Woodward, J. (2003), p. 263

de stabilité est de permettre de mettre de côté des questions comme : «Y a-t-il des lois en biologie ?», ou «Y a-t-il des lois en économie ?» pour concentrer les efforts sur la question moins controversée : «Cette généralisation en biologie est-elle stable sous tel (type de) changement de circonstance d'arrière-plan ?»

## 2) La proportionnalité ou le choix du niveau approprié d'explication

La notion de «niveau approprié d'explication causale» fait partie du sens commun. À titre d'exemple, les facteurs envisagés dans les discussions courantes pour expliquer le prix des oranges à Montréal sont (explication 1) des choses comme le climat récent dans les régions productrices, les frais de transport, les marges des détaillants, le prix des fruits qui seraient des substituts aux oranges, les préférences des consommateurs, la saison, etc. Une explication (explication 2) qui prendrait en compte l'état physique et biologique des cellules du cerveau des producteurs et des acheteurs ne serait pas prise au sérieux, même si ces états, sous l'hypothèse du physicalisme, soutiennent plusieurs des facteurs plus macroscopiques de l'explication 1. L'explication 2, s'il était possible de la produire, serait extrêmement complexe et on dirait avec raison qu'elle n'est pas au niveau approprié. Woodward cherche à expliciter la notion intuitive de niveau approprié d'explication, qu'il appelle également la «proportionnalité», en partant de l'idée que certaines explications peuvent être trop générales ou larges, parce qu'elles incorporent des données inutiles, tandis que d'autres peuvent être trop étroites, parce qu'elles n'incluent pas des facteurs pertinents à l'*explanandum*. La proportionnalité d'une explication constitue donc une sorte d'idéal, qu'il n'est pas toujours possible d'atteindre dans les situations concrètes.

Woodward souligne que le statut d'explication proportionnelle est fonction, d'une part, de la structure causale des phénomènes étudiés, et d'autre part, des intérêts des fournisseurs et des récepteurs de l'explication. Je renvoie ici le lecteur à la discussion du contraste dans la section 3.4, fondée sur l'exemple des intérêts différents de certains chercheurs qui examinent un même accident d'automobile (exemple 3.4.7). Woodward défend l'idée que les intérêts des chercheurs influencent la description précise de l'*explanandum*, et qu'une fois que cet *explanandum* est choisi, la discussion sur ce qui constitue l'explication, ainsi que la proportionnalité adéquate (ou le niveau

approprié) de l'explication, devient une question objective (au sens de indépendante des intérêts des chercheurs).<sup>98</sup>

Woodward propose la paire d'explications suivantes à titre d'illustration plus précise d'une différence de proportionnalité. (Il tire cet exemple d'un article de S. Yablo.)

(3.8.8) «A pigeon is trained to peck at a target when and only when presented with a stimulus of any shade of red. Suppose, on some particular occasion or series of occasions, the pigeon is presented with a particular shade of scarlet and pecks at the target. Consider the following two causal claims/causal explanations :

(3.8.9) The presentation of a scarlet target caused the pigeon to peck.

(3.8.10) The presentation of a red target caused the pigeon to peck. »<sup>99</sup>

Woodward estime que l'explication (3.8.10) constitue l'explication qui satisfait le mieux le critère de proportionnalité, parce qu'elle donne le bon niveau de détail. Ceci est plus apparent quand on examine les défauts de l'explication concurrente (3.8.9). (3.8.9) affirme erronément que la couleur écarlate a causé les coups de bec, ce qui est une description trop précise du facteur causal. Exprimé différemment, (3.8.9) suggère erronément que si la couleur de la cible n'avait pas été écarlate, mais plutôt une autre nuance de rouge, alors le pigeon n'aurait pas donné de coups de bec.<sup>100</sup>

Ces considérations amènent Woodward à donner la caractérisation suivante de la proportionnalité d'une relation causale ou d'une explication causale, à laquelle il associe l'étiquette P :

(3.8.11) Proportionnalité

Given a specification of an effect variable or *explanandum* (which will include a specification of a range of possible values this variable can take) I propose that a cause will be proportional to its effect (or will come closer to satisfying the constraint of proportionality) when (or to the extent that) the following condition is met :

(P) There is a pattern of systematic counterfactual dependence (with the dependence understood along interventionist lines) between different possible states of the cause and the different possible states of the effect, where this pattern of dependence at least approximates to the following ideal : the dependence (and the associated characterization

---

<sup>98</sup> Woodward, J. (2010), p. 296-297

<sup>99</sup> Woodward, J. (2010), p. 297

<sup>100</sup> Woodward, J. (2010), p. 297-298

of the cause) should be such that (a) it explicitly or implicitly conveys accurate information about the conditions under which alternative states of the effect will be realized *and* (b) it conveys *only* such information – that is, the cause is not characterized in such a way that alternative states of it *fail* to be associated with changes in the effect.»<sup>101</sup>

Woodward note que la caractéristique de proportionnalité est au moins partiellement indépendante de la caractéristique de stabilité. La relation entre la présentation d'une cible écarlate et le comportement du pigeon peut être tout aussi stable que la relation entre la présentation d'une cible rouge et ce comportement, même si la relation faisant intervenir la cible écarlate ne satisfait pas aussi bien la caractéristique de proportionnalité<sup>102</sup>.

### 3) La spécificité

Une autre dimension permettant de classifier et de distinguer les relations causales est celle qui se rattache à la finesse de la relation entre une cause possible et des effets. Si on admet qu'une catégorie d'effets peut comprendre de nombreux éléments, ou dit autrement dans le vocabulaire des variables et des valeurs, si on admet qu'une variable d'effet peut prendre un grand nombre de valeurs, alors certains facteurs causaux permettent de choisir avec précision un effet particulier (i.e. une valeur particulière de la variable d'effet), tandis que d'autres facteurs causaux ont une influence plus grossière. Woodward illustre cette distinction avec l'exemple d'un poste radiophonique ancien.

(3.8.12) Le radio est régi par deux facteurs principaux : (1) Un interrupteur qui détermine si le courant électrique fait fonctionner ou pas le radio. (2) Un bouton ou une manette rotative qui permet de syntoniser un poste particulier.

Chacun des facteurs est causal au sens de la caractérisation (3.8.1), puisque une variation par intervention de l'un ou l'autre d'entre eux affecte la réception. Mais il y a asymétrie dans le contrôle sur les effets possibles. L'interrupteur ne permet qu'un effet grossier, ou non-spécifique. Il assure la réception ou la non-réception de tous les postes, quels qu'ils soient. La manette rotative permet un contrôle fin, ou spécifique, en ce sens que la position particulière de la manette détermine spécifiquement quel poste sera entendu.

---

<sup>101</sup> Woodward, J. (2010), p. 298

<sup>102</sup> Woodward, J. (2010), p. 299

Woodward reconnaît une parenté entre cette notion intuitive de spécificité causale et le concept d'influence mis de l'avant par David Lewis en 2000 pour produire (en conjonction avec d'autres concepts) une caractérisation réductive de la relation causale<sup>103</sup>. Même si Woodward estime que le concept d'influence de Lewis ne permet pas de produire une caractérisation réductive de la causalité en général, il pense que la notion d'influence constitue une bonne base pour faire des distinctions de spécificité parmi les relations causales.

S'inspirant de Lewis, Woodward propose la caractérisation suivante de la spécificité causale :

(3.8.13) Spécificité

«Other things being equal, we are inclined to think of C as having more rather than less influence on E (and as a more rather than less specific cause of E) to the extent that it is true that :

INF : There are a number of different possible states of C ( $c_1... c_n$ ), a number of different possible states of E ( $e_1... e_n$ ) and a mapping F from C to E such that for many states of C each such state has a unique image under F in E (that is, F is a function or close to it, so that the same state of C is not associated with different states of E, either on the same or different occasions), not too many different states of C are mapped onto the same state of E and most states of E are the image under F of some state of C. This mapping F should describe patterns of counterfactual dependency between states of C and states of E that support interventionist counterfactuals. Variations in the time and space of occurrence of the various states of E should similarly depend on variations in the time and place of occurrence of states of C.»<sup>104</sup>

Voici quelques éclaircissements à l'égard de cette longue caractérisation. Premièrement, l'objectif est de décrire la capacité d'une variable cause C à moduler de façon fine les états de E.

Deuxièmement, la caractérisation reflète l'idée que les termes de la relation causale, selon Woodward, sont des variables, plutôt que des événements (qui sont les termes de la notion d'influence causale chez Lewis). Lorsqu'on conçoit les termes comme des événements, alors les

---

<sup>103</sup> Lewis, D. (2000). Le lecteur intéressé par les différences entre la conception de Lewis et celle de Woodward peut se reporter à Woodward, J. (2010), p. 304.

<sup>104</sup> Woodward, J. (2010), p. 305

variations possibles de C sont généralement conçues comme des variations de temps, de lieu ou de manière de produire C. Par exemple, la balle (qui a brisé la vitre) aurait pu avoir été lancée plus tôt, avoir eu une trajectoire légèrement différente, ou encore avoir été lancée avec une force légèrement différente. Lorsqu'on conçoit les termes comme des variables, les événements continuent de constituer des termes acceptables de la relation causale (puisque les événements peuvent être représentés par des variables discontinues), mais on peut aussi étendre le champ des causes à des états d'un objet causal qu'on rattache plus difficilement à la notion d'événement. Par exemple, le fait que la paire numéro 345678 de l'ADN d'un chromosome des cellules d'un individu soit présentement défectueuse, et l'ait été depuis longtemps, se range plus difficilement dans la catégorie «événement». Il en est de même de la chaleur d'un métal, qui est un état continu plutôt que discontinu de ce métal. Ces remarques expliquent pourquoi Woodward commence par décrire l'influence en termes d'états de variables, et conclut sa caractérisation en parlant des variations d'état et de lieu des causes et des effets.

Troisièmement, je commente brièvement les notions de mappage ('mapping') et de fonction dans (3.8.13). Le mot «mappage» est un calque peu élégant du mot 'mapping'. La notion de mappage réfère à une relation quelconque (i.e. sans restriction) entre les éléments d'un ensemble C et ceux d'un autre ensemble E. La fonction, par contre, impose une restriction : dans une relation fonctionnelle, un élément quelconque de C doit correspondre à un seul élément de E. On peut aussi concevoir des relations de correspondance où un élément quelconque de E ne doit correspondre qu'à un seul élément de C. Une bijection est une relation de correspondance où chaque élément de C n'a qu'un seul correspondant en E et où chaque élément de E n'a qu'un seul correspondant en C. Woodward, dans sa longue description de la relation causale spécifique veut clairement indiquer que celle-ci devrait être une fonction, ou se rapprocher de ce statut, pour la raison suivante. Il est difficile de se servir d'une relation non-fonctionnelle, où un même état  $c_1$  produit indifféremment deux états différents  $e_m$  et  $e_n$ , pour contrôler via C l'état de E.<sup>105</sup>

Woodward propose l'illustration génétique suivante de la distinction entre une cause spécifique et une cause non-spécifique d'un phénomène. Je donne d'abord quelques indications générales sur le contexte biologique de cet exemple. La relation entre le code génétique d'un individu vivant (le

---

<sup>105</sup> Woodward, J. (2010), p. 304-305



génotype) et certaines de ses caractéristiques observables (le phénotype) est maintenant bien comprise. L'idée générale est que c'est l'information contenue dans l'ADN de chaque cellule d'un être vivant particulier qui est une cause spécifique, via des processus complexes, de certaines caractéristiques observables de cet individu. Un de ces processus complexes est la production de protéines. L'ADN de l'individu contient les informations requises pour produire certaines protéines et la présence ou l'absence de ces protéines affecte ensuite certaines caractéristiques observables de l'individu. Les informations contenues dans l'ADN prennent la forme d'une séquence de composants de l'ADN appelés nucléotides. La séquence des nucléotides contient de l'information au même sens qu'une séquence de chiffres binaires comme 011001101 peut représenter une lettre ou un nombre décimal. De même, c'est la séquence des composants chimiques de la protéine qui détermine ses effets sur les caractéristiques observables de l'individu. Le but des processus cellulaires menant à la production de protéines est donc de transmettre une séquence particulière de composants chimiques. Le processus de production de protéines à partir de l'ADN comprend lui-même deux processus plus élémentaires : (1) la transcription de l'information (i.e. de la séquence de nucléotides) de l'ADN dans une molécule intermédiaire appelée ARN messenger, et (2) la traduction de l'information (i.e. de la séquence de nucléotides) de l'ARN messenger dans le code moléculaire différent (i.e. une séquence d'autres composants moléculaires) de la protéine.<sup>106</sup>

L'exemple de causalité spécifique et non-spécifique en génétique proposé par Woodward se rattache au processus de transcription de la séquence des nucléotides de l'ADN dans une séquence similaire de nucléotides de l'ARN messenger. Ce processus fait intervenir au moins deux types de facteurs causaux : (1) Un facteur spécifique qui est la séquence des nucléotides de la molécule d'ADN à transcrire dans la molécule d'ARN messenger. Ce facteur est spécifique au sens où un changement du nucléotide se trouvant à la position XX de la molécule d'ADN provoque un changement du nucléotide correspondant de la molécule résultante d'ARN messenger. Il y a donc une relation fonctionnelle entre le nucléotide se trouvant à la position XX de l'ADN et le nucléotide se trouvant à la position correspondante de l'ARN messenger résultant. (2) Un facteur non-spécifique : l'ARN polymérase. La cellule où a lieu le processus de transcription doit nécessairement contenir un composant chimique adjuvant, appelé ARN polymérase. L'ARN polymérase est un facteur causal, au sens où une intervention qui enlèverait l'ARN polymérase dans la cellule mettrait

---

<sup>106</sup> Je tire ces informations générales du manuel de Campbell, N. A. (1995), p. 300-324

fin à la transcription. Mais ce facteur joue plutôt un rôle d'interrupteur. Il n'affecte pas le choix du nucléotide particulier de l'ARN messager qui correspondra au nucléotide se situant à la position XX de la molécule d'ADN. L'ARN polymérase est donc un facteur non-spécifique du processus de transcription.<sup>107</sup>

Le degré de spécificité (au sens de (3.8.13)) d'une relation causale sous-jacente dépend de la description retenue des variables. En subdivisant une variable en plusieurs variables, ou en rassemblant ce qui était préalablement considéré comme plusieurs variables en une seule, on peut parvenir à une représentation d'une même relation causale sous-jacente qui paraîtra plus spécifique qu'on ne le supposait antérieurement. La spécificité n'est donc pas une caractéristique objective des relations causales.<sup>108</sup>

Pour conclure la présente section sur les caractéristiques de stabilité, de proportionnalité et de spécificité des relations causales, je rappelle que ces caractéristiques sont pertinentes pour déterminer le niveau de contrôle que les humains, ou la nature, sont en mesure d'exercer sur les phénomènes. Une relation causale qui est plus stable est plus susceptible d'être utilisée dans des processus de manipulation et de contrôle qu'une relation causale moins stable. De même, une relation causale qui est plus spécifique qu'une autre offre davantage de possibilités de contrôle fin des résultats. On pense habituellement à de situations de contrôle par des agents humains, mais il faut également prendre en compte les fonctions de contrôle qui s'exercent de façon involontaire dans les organismes biologiques. On comprend mieux par exemple les mécanismes biologiques contrôlés par les gènes, ou encore les mécanismes du système immunitaire si on identifie les caractéristiques de stabilité et de spécificité des relations causales impliquées.<sup>109</sup>

---

<sup>107</sup> Woodward, J. (2010), p. 302, 306

<sup>108</sup> Woodward, J. (2010), p. 312

<sup>109</sup> Woodward, J. (2010), p.315

## CHAPITRE 4. ÉLÉMENTS DE CRITIQUE

Je présente dans ce chapitre certains éléments de critique à l'égard de la conception manipulative de la causalité et de l'explication causale de Woodward. Je commencerai par faire ressortir un inconvénient de la conception manipulative qui me semble inséparable des avantages décrits aux deux chapitres précédents. Je procéderai ensuite à la description de certaines limites de la conception manipulative. Je devrai également indiquer certaines limites du présent mémoire.

### 4.1 Des gains et au moins une perte

Comme nous l'avons vu précédemment, Woodward refuse de fonder la caractérisation des relations causales sur la notion de loi de la nature. Il juge également que les tentatives passées et présentes de produire une caractérisation pleinement réductive de la notion de relation causale n'ont pas réussi. Il considère plutôt que la caractérisation des relations causales doit inclure un concept lui-même causal, celui d'intervention. Dit autrement, la notion de relation causale est primitive. Woodward accepte sans démonstration explicite l'idée qu'il existe des relations causales dans l'univers et que certaines nous sont connues, et il donne une caractérisation opérationnelle de (nouvelles) relations causales à découvrir ou à confirmer. On peut confirmer le statut de la relation possiblement causale  $R$  en faisant appel à une ou plusieurs autre(s) relation(s) causale(s) déjà connue(s)  $R_1, R_2, \dots$  qui sont utilisée(s) dans le processus d'intervention. Plus précisément, on peut confirmer le statut de la relation possiblement causale  $R$  entre une cause  $C$  et un effet  $E$  si  $C$  satisfait le test hypothétique ou réel de la cause contributive ((2.2.9), dans la forme «minimale», ou (3.5.3) à (3.5.5) dans une forme renforcée). Ce test comprend une ou plusieurs interventions et établit qu'il y a au moins une instance de dépendance contrefactuelle entre  $C$  et  $E$ .

Contrairement à ce qui se passe quand la relation causale est rattachée à la notion de loi, une relation causale avérée n'a pas besoin de s'appliquer en tout temps et en tout lieu. Elle peut s'appliquer dans un domaine spatiotemporel limité. Elle peut aussi comporter des exceptions à l'intérieur de son domaine d'application, quand la relation est probabiliste.

L'avantage de cette conception est qu'elle se rapproche de la notion causale employée dans les sciences biomédicales et dans les sciences humaines. Si une politique gouvernementale vise à modifier un comportement social, on considérera que la politique a eu l'effet causal escompté si une grande partie de la population, mais peut-être pas tous les individus, adopte le comportement modifié, et ce pour une période possiblement limitée.

La conception manipulative constitue par contre un affaiblissement majeur de la notion de lien causal, qui perd une bonne partie de sa connotation déterministe. Cet affaiblissement reflète l'idée qu'une relation causale est essentiellement un levier pour faire arriver des choses, si les circonstances sont appropriées, plutôt qu'une relation qui garantit en tout temps des résultats. Cet aspect possiblement incertain et temporaire de la relation causale dans la conception manipulative est ce qui justifie les distinctions de stabilité, de spécificité et de niveau approprié que Woodward propose parmi les relations causales avérées. Un ingénieur qui conçoit un avion cherche à utiliser des relations causales très stables et qui permettent un contrôle fin des phénomènes.

La conception manipulative de la causalité reflète donc un affaiblissement pratique. Le fait qu'un phénomène ait été causal à un moment et en un lieu particulier de l'histoire ne garantit aucunement que le lien causal sera toujours en place plus tard, ou qu'il soit transférable en d'autres lieux. La plupart des liens causaux, en sciences humaines en particulier, doivent être continuellement reconfirmés. Un exemple favori de Woodward est tiré d'une étude de E. Veblen, publiée en 1975, sur l'influence causale des positions éditoriales d'un journal de Manchester, NH, sur le vote des citoyens de Manchester entre 1960 et 1972. Woodward accepte la conclusion de Veblen selon laquelle les positions éditoriales du journal ont influencé causalement le vote durant la période étudiée, mais il soutient en même temps que les résultats détaillés de l'étude ne peuvent pas être transférés tels quels à d'autres régions des États-Unis, et qu'ils pourraient ne plus être applicables aux patterns de vote des citoyens de la ville de Manchester durant des périodes postérieures à 1972<sup>110</sup>.

---

<sup>110</sup> Woodward, J. (2003), p. 206-207, p. 347-348

## 4.2 Échec dans le cas de théories physiques appliquées à l'ensemble de l'univers

L'application de la conception manipulative semble échouer dans le cas de phénomènes régis par des théories physiques visant l'univers en entier. Soit à titre d'exemple l'assertion :

(4.2.1) «The state  $S_t$  of the entire universe at time  $t$  causes the state  $S_{t+d}$  of the entire universe at time  $t+d$ .

where  $S_t$  and  $S_{t+d}$  are specifications in terms of some fundamental physical theory.»<sup>111</sup>

Dans la conception manipulative, l'assertion (4.2.1) signifie qu'une intervention possible sur  $S_t$ , produisant un état alternatif  $S_u$ , aurait pour effet un changement de l'état subséquent  $S_{t+d}$ , qui serait maintenant  $S_{u+d}$ . Mais on se demande en quoi pourrait consister une intervention sur  $S_t$ , et comment elle pourrait être effectuée (même théoriquement), puisque  $S_t$  comprend l'univers entier et qu'il n'y a pas de lieu ni d'entité en dehors de  $S_t$  à partir duquel effectuer l'intervention. Il semble que la conception manipulative s'applique de façon plus naturelle à des systèmes limités, c'est-à-dire à des systèmes où les relations causales visées se trouvent dans des environnements plus vastes à partir desquels on peut effectuer des interventions exogènes.

A titre d'illustration de théories physiques visant l'univers entier, je retiens les théories cosmologiques décrivant un début possible de l'univers que nous connaissons aujourd'hui, comme la théorie du Big Bang. La conception manipulative échoue à l'égard de cette théorie parce que nous ne disposons pas, même en théorie, de lieu extérieur au système visé pour servir de base à une intervention, et que nous n'avons pas de conception claire de ce en quoi une intervention pourrait consister. Par contre, une théorie physique visant le comportement d'une masse de gaz consistant en un nombre limité de molécules, dans un environnement situé à la surface de la Terre, peut être analysée en utilisant la conception manipulative parce qu'il y a un lieu, extérieur au système visé, à partir duquel une intervention pourrait avoir lieu, et parce que nous avons une conception claire de ce en quoi l'intervention pourrait consister. On pourrait par exemple modifier le volume du gaz, s'il se trouve dans un contenant rigide, au moyen d'un piston.

---

<sup>111</sup> Woodward, J. (2009), p. 257

Je note que la restriction proposée au domaine d'application ('scope') de la conception manipulative n'affecte pas les inférences causales de la vie quotidienne, ni celles de la biologie ou des sciences sociales, mais seulement celles de la physique dans les cas où elles visent l'univers entier.

Woodward identifie deux principales réactions possibles à la restriction suggérée plus haut. Soit (1), on accepte que les assertions portant sur l'ensemble de l'univers, comme (4.2.1), sont vraies, et on conclut que «it is an important limitation in interventionist theories that they have difficulty eliciting such claims». Soit au contraire (2), on affirme que «causal concepts do not apply, at least in any straightforward way, to some or many fundamental physics contexts and that it is a virtue of the interventionist account that it helps us to understand why this is so»<sup>112</sup>. Selon Woodward, cette seconde position serait la position actuellement prédominante chez les philosophes de la physique. Quelle que soit notre réaction, Woodward note que la question du domaine d'application des théories manipulatives et des concepts causaux en général constitue actuellement un problème important.<sup>113</sup>

Je mentionne ici une constatation un peu différente à l'égard de la conception manipulative en relation avec certaines généralisations de la physique. Il semble que la conception manipulative s'applique plus naturellement à des relations causales dont le domaine est limité et qui sont relativement moins stables, comme celles des sciences biomédicales et des sciences humaines, qu'aux relations plus universelles et plus stables qu'on peut trouver dans certains domaines de la physique.

Une façon d'analyser les structures causales complexes est de les décomposer. En neutralisant certaines relations causales, on crée des circonstances où d'autres relations, également présentes mais occultées, se révèlent. C'est ce qui se produit dans les tests randomisés sur les médicaments. En choisissant au hasard les participants du groupe de traitement et du groupe témoin, on neutralise l'effet, sur les résultats du test, de la volonté des patients d'être soignés. Mais les

---

<sup>112</sup> Woodward, J. (2009), p. 257

<sup>113</sup> Woodward, J. (2009), p. 256-258

relations physiques les plus universelles et stables sont moins susceptibles d'être neutralisées par intervention, et on a plus de difficulté à révéler les relations causales occultées<sup>114</sup>.

### 4.3 Les mécanismes

Woodward dans son texte de 2003, et moi-même dans le présent mémoire, employons trop librement la notion de mécanisme dans le sens un peu vague de vérifacteur des relations causales. Cela pose problème à plusieurs égards. Le principal problème est que Woodward utilise selon moi le mot «mécanisme» dans au moins deux sens différents, selon le contexte. Le second problème est que les mécanismes jouent un rôle assez différent dans certaines autres conceptions, dites «mécanistes», de la causalité, et que le glissement de sens que je perçois chez Woodward apporte un risque de confusion entre la conception manipulative et les conceptions mécanistes. Je vais d'abord tenter de donner une caractérisation plus précise de la notion de mécanisme, qui représente le sens visé par Woodward la plupart du temps. Je me servirai de cette caractérisation pour répondre aux questions : (1) la conception manipulative suppose-t-elle l'existence d'un mécanisme ?, et (2) la conception manipulative suppose-t-elle la connaissance d'un mécanisme ? Je dirai ensuite quelques mots sur le rôle des mécanismes dans la conception manipulative. Je présenterai enfin le contexte où Woodward emploie le mot mécanisme dans un sens différent et les objectifs de cet emploi.

Voici d'abord une caractérisation «large» des mécanismes. Un mécanisme dans ce premier sens est une structure ou un processus sous-jacent au comportement d'une entité plus globale.

#### (4.3.1) Caractérisation des mécanismes (au sens de structure ou de processus sous-jacent)

Un mécanisme est l'ensemble des relations stables reliant une entité complexe, ses parties, et les facteurs externes susceptibles de les influencer.

(Le qualificatif «stable» vise ici à exclure du domaine des mécanismes les situations où interviennent des relations non répétables, même en théorie.)

Cette caractérisation est large, car elle inclut tout autant les machines que les organismes vivants individuels, les groupes de vivants comme les espèces, ou encore des entités sociales ou

---

<sup>114</sup> Woodward, J. (2007), p. 95

économiques. La caractéristique essentielle est que le comportement d'une entité complexe y est régi au moins en partie par des relations s'appliquant à ses parties.<sup>115</sup>

Je note qu'il existe des caractérisations plus «étroites» des mécanismes en tant que structures ou processus sous-jacents. Ces caractérisations exigent que les parties des entités complexes soient des choses matérielles et que leurs interactions se fassent dans des conditions de contiguïté spatiotemporelle ou encore avec la conservation d'une quantité physique, comme l'énergie. Cette caractérisation plus étroite des mécanismes est considérée comme la caractéristique essentielle de la causalité dans les conceptions dites «mécanistes» de la relation causale. Une des conséquences de ces conceptions «mécanistes» de la relation causale est que plusieurs phénomènes que le sens commun considère comme causaux deviennent problématiques. Je signale par exemple les relations à distance entre des masses dans la gravitation, ou encore les phénomènes de causalité par omission (par exemple, le fait que je n'ai pas arrosé la plante a causé sa mort).

Une partie de la confusion entourant la question des mécanismes en causalité vient justement de la proximité entre la notion de relation causale et celle de mécanisme. Les tenants des conceptions mécanistes soutiennent que l'existence d'un mécanisme au sens «étroit» est une condition nécessaire de l'existence d'une relation causale. La plupart des autres penseurs estiment au contraire que l'existence d'un mécanisme au sens étroit n'est pas une telle condition nécessaire.

Je présente d'abord la position de Woodward à l'égard de la question «ontologique». L'existence d'un mécanisme au sens étroit, ou au sens large, est-il une condition nécessaire de la causalité ? Réponses : En ce qui concerne les mécanismes au sens étroit, la conception manipulative accepte la causalité par omission et les relations causales à distance dans la gravitation. Elle rejette donc l'assertion que l'existence d'un mécanisme au sens étroit soit une condition nécessaire de la causalité. En ce qui concerne les mécanismes au sens large, mon impression est que la conception manipulative ne prend pas position. Cela découle en partie du caractère très peu exigeant de la conception des mécanismes au sens large. Cela découle également du fait que les caractérisations minimales de la cause directe (2.2.4) et de la cause contributoire (2.2.9) ne disent rien sur la nature

---

<sup>115</sup> Reiss, J. (2013), p. 104



de la relation entre la cause et l'effet. Comme on l'a vu plus haut, la conception manipulative est compatible avec une interprétation «opérationnelle» de la causalité.

Je passe maintenant à la question épistémique. Selon la conception manipulative, doit-on, pour établir l'existence d'une relation causale et/ou exploiter une relation causale dans le monde pratique, connaître un mécanisme au sens étroit ou au sens large ? La réponse est évidemment non dans le cas des mécanismes au sens étroit, en vertu de la position ontologique correspondante. En ce qui concerne les mécanismes au sens large, la réponse est également négative. Il y a au moins deux types de situations où les agents peuvent établir l'existence d'une relation causale et l'exploiter sans connaître de mécanisme au sens large. Premièrement, il y a les situations de causalité minimale, où les agents ne connaissent pas de relation structurée entre les facteurs, et ne connaissent que des tendances de dépendance. L'exemple typique (2.1.8) de cette situation est celui des chercheurs qui tentaient d'établir, dans la situation épistémique de 1957, l'existence d'une relation causale entre l'habitude de fumer et le cancer du poumon. Deuxièmement, il existe des situations de causalité renforcée, où les agents ont connaissance d'une relation précise entre des facteurs, permettant de représenter la relation par des équations, mais où ils ne connaissent pas le mécanisme sous-jacent à cette relation. Ce type de relation a souvent été désigné par l'expression «loi phénoménologique». Un exemple bien connu de loi phénoménologique est celui de la loi des gaz parfaits (3.4.4) : la loi décrit les relations entre les variables macroscopiques mais n'inclut pas de variables décrivant l'état des molécules, à titre de parties du gaz. Ainsi la connaissance d'une relation causale exprimée par une loi phénoménologique permet aux agents d'exploiter une relation causale bien structurée mais dont on ne connaît pas le mécanisme.

Il n'en demeure pas moins que la connaissance d'un mécanisme donne un avantage considérable aux agents. La connaissance des mécanismes sous-jacents permet habituellement de répondre à un plus large éventail de questions W, et donne donc davantage de moyens de contrôler les phénomènes. Ainsi la découverte des mécanismes cellulaires rattachés à l'ADN a transformé la génétique et les sciences biomédicales. La découverte des mécanismes propres à leur domaine de recherche demeure donc un objectif fondamental des chercheurs.

Je passe maintenant à un aspect différent de l'emploi de la notion de mécanisme chez Woodward. Il existe un contexte où Woodward emploie la notion de mécanisme dans un sens différent de celui de structure ou processus sous-jacent au comportement d'une entité plus globale. Ce second emploi peut porter à confusion.

Dans la description des hypothèses ou postulats généraux de la conception manipulative, Woodward nous dit qu'on doit pouvoir envisager des interventions. Or une intervention sur une variable quelconque D (faisant partie d'un système S), qui est, d'une part, reliée à des variables qui l'influencent, A, B, C,..., et qui, d'autre part, influence à son tour d'autres variables, E, F, H, I,..., consiste à briser les liens entre D et les variables A, B, C,..., qui l'influencent, à changer de façon exogène la valeur de D, et à observer les effets de ce changement sur les variables E, F, H, I,...qui font partie des effets de D.<sup>116</sup> Ceci présuppose qu'on peut annuler sélectivement une ou plusieurs relations d'un système sans annuler l'effet causal des autres relations du système S. C'est l'hypothèse de séparation des mécanismes dont j'ai parlé au chapitre 2. Pour exprimer cette idée, Woodward a tendance à parler de mécanismes causaux plutôt que de relations causales. Ainsi, dans ce contexte, «mécanisme» est un synonyme de «relation causale», et la signification de «mécanisme» au sens de relation causale rattachée explicitement à une structure sous-jacente ne s'applique pas. Voici un exemple de cet emploi :

In what follows, I assume that when causal relationships are correctly and fully represented by systems of equations, each equation will correspond to a distinct causal mechanism and that the equation system will be modular.<sup>117</sup>

Je laisse de côté le dernier membre de phrase portant sur la modularité des systèmes d'équations. Je souligne qu'ici l'expression «distinct causal mechanism» pourrait être remplacée par «distinct causal relation».

Cet emploi particulier du mot mécanisme pour désigner les relations causales en général semble être fréquent en économétrie, une discipline où on représente les relations causales complexes de

---

<sup>116</sup> Je laisse de côté le cas important des rétroactions, où la variable A est, dans le même système, une cause de D et un effet de D.

<sup>117</sup> Woodward, J. (2003), p. 49

l'économie au moyen de systèmes d'équations, et qui est une des disciplines dont Woodward s'est inspiré.<sup>118</sup>

#### 4.4 L'explication dans les situations épistémiques de causalité minimale

Nous avons vu que Woodward présente deux groupes de concepts causaux. Un groupe «minimal», décrit au chapitre 2, où la caractérisation des concepts causaux n'exige pas la connaissance d'une généralisation précise (G) décrivant la relation causale, et un groupe renforcé, décrit au chapitre 3, qui présuppose que G est connu. L'explication causale a été incluse par Woodward parmi les concepts renforcés. Cela soulève plusieurs questions, dont la principale, selon moi, est : N'y a-t-il pas d'explications causales à tirer de généralisations apparemment peu précises comme «Les courts-circuits causent les incendies» ou «Fumer cause le cancer» pour lesquelles il n'y a pas de généralisation précise G ? Je vais examiner cette question en deux étapes. Premièrement je vais tenter de caractériser de façon un peu plus nette la nature des généralisations causales «minimales» (que j'appellerai dorénavant, suivant Woodward, «généralisations de cause possible»). Deuxièmement, je vais montrer que, dans certaines circonstances, les généralisations de cause possible peuvent mener à des explications suffisamment concrètes pour permettre le jugement et l'action.

Voici la caractérisation que donne Woodward de ces généralisations (de cause possible) relativement invariantes.

##### (4.4.1) Généralisation de cause possible

A possible-cause generalization 'Cs cause Es' will be relatively invariant to the extent that Es occur more frequently when Cs are introduced by interventions than when Cs are absent across a wide or large range of changes in circumstances or background conditions, although not necessarily in *all* such circumstances. [...] We might have compelling evidence that Es occur more frequently when Cs are introduced by interventions than when Cs are absent in a wide variety of circumstances even though there is no known generalization of the form 'Cs in K are always followed by Es' that we are prepared to regard as a law and

---

<sup>118</sup> Reiss, J. (2013), p. 104

even though the amount by which the introduction of Cs raises the probability of Es varies in an irregular way, even when other known relevant variables are controlled for. In such a case, what is stable under interventions across different contexts is simply the fact that there is some correlation or other between Cs and Es.<sup>119</sup>

Dans cette caractérisation, l'expression 'to be relatively invariant', appliquée à une généralisation de cause possible, signifie posséder une invariance moins forte que celle qui serait fondée sur une généralisation précise G. La condition de l'invariance «relative» est exprimée par le fait qu'il existe, dans ce cas, un éventail de situations différentes où la probabilité conditionnelle d'un effet E, lorsqu'un facteur causal C est introduit par intervention, est plus élevée que la probabilité de l'effet E sans l'introduction du facteur causal C. Nous sommes ainsi dans une perspective probabiliste, et cela appelle certains commentaires.

Premièrement, le concept d'intervention, dans un contexte probabiliste, diffère du concept correspondant dans un contexte déterministe. Dans un contexte déterministe une intervention sur une variable C consiste en un changement exogène qui fixe C à une valeur c (différente de sa valeur réelle) pour un individu particulier, dans le but de constater si la valeur de E pour cet individu changerait. Dans un contexte probabiliste, le but de l'intervention est d'examiner l'impact d'un changement exogène de C sur la probabilité de la valeur de E. Une intervention consistera alors à appliquer à une population d'individus (et non à un individu particulier) un changement identique de la valeur de C, de façon à examiner l'effet du changement sur la probabilité des diverses valeurs de la variable E dans la population. Les interventions en contexte probabiliste sont donc une opération beaucoup plus complexe et beaucoup moins flexible que dans un contexte déterministe.

Deuxièmement, puisque nous adoptons une optique probabiliste, il serait désirable de pouvoir distinguer clairement, parmi les généralisations probabilistes (que je désigne ici avec le symbole «GP»), entre (a) les GP véritablement invariantes, que Woodward reconnaît comme suffisamment précises pour fonder des explications au sens de (3.4.1), et (b) les GP «relativement» invariantes caractérisées dans (4.4.1). Woodward donne dans (4.4.1) deux indications négatives à cet égard. (i) La GP visée n'est pas telle qu'on connaisse un facteur K, qui, associé à C, fasse en sorte que l'effet E

---

<sup>119</sup> Woodward, J. (2003), p. 311-312

s'ensuivrait toujours. S'il y avait un tel facteur K connu, alors la GP se ramènerait à une généralisation déterministe. Et (ii) La GP visée n'est pas telle que l'introduction des C par intervention augmenterait la probabilité des E d'une manière «régulière», après avoir éliminé l'effet des variables qui pourraient masquer la relation entre C et E. Si l'effet de l'introduction des C augmentait la probabilité des E de manière régulière, alors la GP serait une généralisation vraiment invariante (que Woodward inclurait dans la classe G des généralisations précises).

Malheureusement je n'ai pas la compétence technique requise pour évaluer cette caractérisation.<sup>120</sup> Je dois me contenter de l'intuition qu'il y a des généralisations probabilistes précises et qu'il y a d'autres généralisations probabilistes floues caractérisées plutôt par un ensemble de liens statistiques reliant un facteur causal et son effet dans un ensemble de situations d'arrière-plan différentes. Je vais donc procéder en acceptant les exemples de généralisation de cause possible relativement invariante proposés par Woodward, à titre d'exemples paradigmatiques. Il s'agit des généralisations :

(2.1.8) Fumer cause le cancer du poumon (dans la situation épistémique de 1957).

(4.4.2) Les courts-circuits causent les incendies.

La discussion de (2.1.8) au chapitre 2 a montré qualitativement comment l'habitude de fumer pouvait en 1957 être rattachée de façon probabiliste à l'incidence du cancer du poumon. Il y a une augmentation de la probabilité de cancer associée à l'habitude de fumer dans un grand nombre de groupes différents. Il y a également des corrélations de type tendance : plus la durée de l'habitude de fumer est longue, plus l'incidence du cancer est élevée, etc. Ces constatations n'ont pas été obtenues par intervention, mais je suppose que les contrôles statistiques sur les variables qui masqueraient la relation entre fumer et l'incidence de cancer visaient à rapprocher les résultats des recherches colligées en 1957 de ceux qui auraient été obtenus par intervention. Au total, la description de l'exemple (2.1.8) correspond d'assez près à la caractérisation (4.4.1).

Je passe maintenant à la seconde étape de la discussion du pouvoir explicatif des généralisations de cause possible. Mon objectif est de montrer que même si ces généralisations semblent faibles, il y a des circonstances où elles peuvent fonder des explications valables et d'autres conclusions.

---

<sup>120</sup> Je signale que Woodward apporte d'autres précisions sur les caractéristiques des généralisations probabilistes invariantes dans Woodward, J. (2003), p. 321-327

À première vue, la généralisation de cause possible (4.4.2) sur les courts-circuits semble fonder une affirmation contrefactuelle très faible :

(4.4.3) If a short circuit were to occur in any one of a range of different background circumstances, then in those circumstances a fire will sometimes occur.<sup>121</sup>

Mais Woodward suggère que la plupart des usagers lui attribuent un pouvoir explicatif, et que cette croyance des usagers peut découler d'un raisonnement de forme similaire au suivant, qui part de l'assertion d'un fait singulier et de la généralisation de cause possible, et qui aboutit à une explication du fait singulier.

(4.4.4) Raisonnement menant à une explication

(P1) Il y a eu un court-circuit particulier sur le site d'un incendie dans un bâtiment particulier.

(P2) Les courts-circuits causent (parfois) les incendies.

(P3) Dans le monde que nous connaissons, les incendies ont toujours des causes.

(P4) Une liste exhaustive des autres causes possibles des incendies est connue et comprend : foyer en marche et mal géré, article de fumeur, événement dans la cuisine, incendie volontaire ('arson'), etc.

(P5) Un examen des lieux de l'incendie particulier permet d'exclure chacune des autres causes possibles de cet incendie.

(Conclusion 1) Le court-circuit particulier est la cause de l'incendie particulier.

(Conclusion 2) Si le court-circuit particulier n'avait pas eu lieu, alors l'incendie particulier n'aurait pas eu lieu.

Ce raisonnement explicatif est manifestement précaire, mais il illustre néanmoins une partie du pouvoir explicatif des généralisations de cause possible dans le monde pratique.<sup>122</sup>

Un autre aspect du pouvoir explicatif des généralisations de cause possible se rattache à la notion d'influence (au sens de Lewis) que nous avons discutée à la section 3.8. Woodward indique que dans de nombreuses circonstances une généralisation de cause possible associée à l'assertion d'événements singuliers peut fonder l'assertion de contrefactuels d'influence. Un exemple standard employé pour illustrer ces situations est le suivant : Billy lance une roche et la roche brise

---

<sup>121</sup> Woodward, J. (2003), p. 215

<sup>122</sup> Woodward, J. (2003), p. 215-216

une vitre. Les contrefactuels d'influence prennent la forme : Si Billy avait lancé la roche plus tôt, ou un peu différemment, alors la vitre aurait été brisée un peu plus tôt, ou de manière différente. Voici le type de raisonnement rattaché aux généralisations de cause possible que l'on peut faire dans certaines circonstances.

(4.4.5) Raisonnement menant à des contrefactuels d'influence

(P1) Billy a lancé une roche. Une vitre particulière a été brisée.

(P2) Les lancers de roche brisent parfois des vitres.

(P3) Il n'y a pas de cause cachée du bris de la vitre.

(P4) Un examen des autres causes possibles connues du bris exclut les autres causes possibles du bris.

Conclusion (1) La roche lancée par Billy a causé le bris particulier.

Conclusion (2) Si la roche avait été lancée plus tôt, ou de manière différente, alors la vitre aurait été brisée plus tôt, ou de manière différente.<sup>123</sup>

Il semble donc que les généralisations de cause possible, dans certaines circonstances, ont un pouvoir explicatif. Cela implique que la caractérisation (3.4.1) de l'explication causale de Woodward, qui semble limiter l'explication aux situations où une généralisation (G) invariante est connue, semble indument restrictive. Il faut laisser un certain espace pour des explications fondées sur des généralisations (plus faibles) de cause possible. Ces explications seront toutefois de moindre qualité parce que l'absence de la généralisation G réduit l'éventail des questions W auxquelles on peut donner des réponses. Ainsi le fait d'établir, en 1957, une relation causale entre l'habitude de fumer et le cancer, sur la base de données principalement épidémiologiques, nous permettait de justifier certaines mesures préventives visant à réduire l'incidence du cancer. Mais la compréhension des mécanismes en jeu dans la reproduction des cellules cancéreuses permet de produire des généralisations G satisfaisant les critères de l'invariance probabiliste et de répondre à un éventail beaucoup plus vaste de questions W, et ainsi de justifier des mesures plus raffinées pour le traitement de la maladie.

---

<sup>123</sup> Woodward, J. (2003), p. 218

#### 4.5 Les possibilités sérieuses

À la section 3.4, j'ai présenté la conception manipulative de l'explication. Cette présentation a mené à une caractérisation de l'explication causale (3.4.1) qui se rattache aux caractérisations de la cause directe (2.2.4) et de la cause contributoire (2.2.9). Dans les trois caractérisations, une cause est un facteur dont le changement réel ou hypothétique, satisfaisant les critères de l'intervention, provoque un changement du facteur effet. Mais Woodward est obligé d'ajouter une restriction à ces caractérisations. Seuls sont retenus parmi les facteurs causaux, les facteurs qui satisfont la condition supplémentaire d'être considérés comme des possibilités sérieuses par les émetteurs et les récepteurs de l'explication. Ainsi, par exemple, la caractérisation (2.2.9) de la cause contributoire devrait être amendée comme suit :

##### (4.5.1) Cause contributoire amendée

A necessary and sufficient condition for X to be a (type level) contributing cause of Y with respect to a variable set V is that (i) X be considered a serious possible causal variable, and that (ii) there be a directed path from X to Y such that each link on this path is a direct causal relationship; that is, a set of variables  $Z_1 \dots Z_n$  such that X is a direct cause of  $Z_1$  which is in turn a direct cause of  $Z_2$  which is a direct cause of ...  $Z_n$  which is a direct cause of Y, and that (iii) there be some intervention on X that will change Y when all other variables in V that are not on this path are fixed at some value.

La condition supplémentaire (i) introduit un élément subjectif dans les jugements causaux. Deux personnes pourront, en vertu de leurs intérêts différents, être légitimement en désaccord sur le domaine des facteurs causaux possibles mais non-sérieux. Pour sauvegarder une certaine objectivité dans l'explication causale, Woodward doit alors subdiviser le processus d'élaboration des énoncés causaux et explicatifs en deux étapes. Premièrement, l'exclusion des possibilités considérées non-sérieuses, un processus subjectif. Deuxièmement, si un consensus peut s'établir sur les possibilités causales restantes, l'examen des causes et l'élaboration des explications, un processus plus objectif.<sup>124</sup>

---

<sup>124</sup> Je réfère le lecteur à la section 3.4, sous-section 3, intitulée «Les possibilités sérieuses».



## CHAPITRE 5. CONCLUSION

J'ai tenté dans ce mémoire de présenter les arguments proposés par James Woodward pour rejeter l'idée que les lois de la nature sont les vérificateurs ou le fondement des explications causales, comme on le soutient dans la tradition de l'analyse de ces explications inaugurée par Hempel. L'examen de ces arguments m'a amené à examiner la conception plus étendue de Woodward visant non seulement l'explication causale, mais la relation causale elle-même. Comme les diverses conceptions de la nature de la relation causale sont notoirement divergentes et controversées, j'ai tenté de cerner les présupposés, la cohérence et certaines limites de la conception Woodward. Puisque la notion de loi de la nature est également présente dans plusieurs autres conceptions de la relation causale, j'ai l'impression que l'analyse des positions de Hempel et de Woodward sur ces questions donne au lecteur des arguments pour porter des jugements sur ces autres conceptions.

Je passe maintenant en revue les principales positions que je crois bien fondées dans la conception manipulative, et également certains aspects plus discutables. Parmi les positions bien fondées, je retiens les suivantes.

La motivation de la recherche causale est principalement pratique, plutôt que simplement épistémique. Chercher les causes vise la modification et le contrôle des phénomènes, dans la mesure du possible.

Cette motivation justifie à son tour l'aspect opérationnel de la conception manipulative. Il n'est pas nécessaire de connaître ou de comprendre de façon approfondie la «nature» ou le mécanisme d'un phénomène pour pouvoir (1) constater la présence d'un lien causal, et (2) agir pour contrôler le phénomène. La caractérisation opérationnelle de la conception manipulative est considérée par certains philosophes comme une faiblesse. Il serait selon eux préférable de mieux identifier la nature commune de toutes les relations causales, exprimée en termes d'éléments ontologiques fondamentaux. Cet objectif est louable, mais il semble difficile à atteindre. Une conception plus modestement opérationnelle a l'avantage d'être applicable dans un monde qui contiendrait plus d'un type ontologique de relation causale, comme par exemple un monde où certaines relations seraient «mécaniques» et d'autres pas.

La caractérisation opérationnelle de la relation causale est non-réductive en ce sens qu'elle fait appel à un autre concept causal, l'intervention. Malgré cela, cette caractérisation n'est pas circulaire. Dans l'intervention, on se sert d'informations déjà connues sur d'autres relations causales pour établir le statut, causal ou non, d'une relation sous investigation.

La notion de loi de la nature, comme fondement conceptuel de la relation causale, apparaît, directement ou indirectement, dans de nombreuses conceptions de l'explication et de la relation causale. Si l'on rattache à cette notion les caractéristiques d'application universelle dans le temps et l'espace, et/ou d'absence d'exception, alors l'idée que les lois sont le fondement des relations causales ne correspond pas aux résultats des sciences biomédicales et des sciences humaines. Le concept de généralisation invariante sous intervention correspond mieux aux généralisations qu'on retrouve dans la plupart, sinon toutes, les disciplines scientifiques.

Le concept de généralisation invariante sous intervention est peu exigeant en ce sens qu'il admet qu'une relation entre des facteurs X et Y est causale s'il existe une (seule) instance répétable de variation sous intervention de X qui provoque une variation de Y. L'inconvénient de cette faible exigence est que les relations admises comme causales selon ce critère ne sont pas toutes également utilisables aux fins du contrôle des phénomènes. C'est pourquoi Woodward ajoute deux principaux critères supplémentaires permettant de distinguer, parmi les relations causales, celles qui sont les plus susceptibles d'être utilisables aux fins de contrôle. La stabilité d'une relation causale exprime le fait que celle-ci tient sous une diversité plus ou moins grande de conditions d'arrière-plan. La spécificité d'une relation causale exprime le fait qu'un facteur causal X contrôle de façon plus ou moins précise la valeur du facteur effet Y.

Je passe maintenant à la revue de certains aspects plus problématiques de la conception manipulative.

Les raisonnements menant à l'identification des facteurs réellement causaux, selon la conception manipulative, font appel à une et souvent à plusieurs assertions contrefactuelles. Cet aspect constitue à la fois une force et une faiblesse de la conception. C'est une force en ce sens que cela

souligne l'aspect contrefactuel inhérent aux relations causales et à l'explication. Donner une explication d'un événement passé A, ce n'est pas principalement affirmer que cet événement était attendu ('was to be expected'); c'est également dire que l'événement A aurait pu être différent, et dire quel aurait été le ou les événements alternatifs à A. Le recours aux assertions contrefactuelles constitue également une force en rattachant la causalité et l'explication à la méthode expérimentale, qui est essentiellement une mise à l'épreuve de certaines assertions contrefactuelles.

Mais le recours aux contrefactuels peut constituer une faiblesse, quand l'application directe de la méthode expérimentale est impossible, et ce, à deux niveaux. Premièrement, il y a des controverses à l'égard de ce que devraient être les bons contrefactuels à invoquer, dans une situation causale donnée. La conception manipulative propose des réponses à cette première question, mais les tenants d'autres conceptions de la causalité proposent des réponses différentes. Deuxièmement, il y a d'autres controverses sur les conditions de vérité de ces «bons» contrefactuels. Dans le présent mémoire, je n'ai pas approfondi la question des conditions de vérité de ces contrefactuels.

Les concepts de généralisation invariante sous intervention et d'explication sont initialement caractérisés par Woodward en supposant que les utilisateurs ont accès à un énoncé relativement précis (G) de la relation entre une variable causale X et son effet Y. Cela semble suggérer qu'en l'absence d'une connaissance précise de G, on ne puisse pas établir l'invariance et on ne puisse pas proposer d'explication. Mais Woodward développe par la suite un deuxième groupe de concepts : la généralisation de cause possible, et l'invariance relative, qui adaptent ses conceptions à des situations épistémiques où les utilisateurs n'ont pas accès à des généralisations précises G. J'ai discuté ces diverses notions dans mon mémoire, mais je n'ai pas tenté de présenter la différence entre les généralisations probabilistes précises G et les généralisations de cause possible. Il ne s'agit pas là d'une faiblesse de la conception manipulative mais plutôt d'une limite du présent mémoire, due au fait que cette différence relève de notions statistiques qui dépassent mes compétences.

Sur le plan méthodologique, je demeure convaincu que l'approche des problèmes de causalité par la philosophie des sciences est incontournable, même en ce qui concerne la clarification du concept de cause. La richesse des matériaux d'origine scientifique, de même que leur aspect concret, par opposition aux pures expériences de pensée, me semblent particulièrement pertinents. Cela ne signifie pas que les approches d'inspiration plus métaphysique de l'analyse de la causalité soient sans valeur, mais plutôt que leurs arguments et leurs conclusions doivent impérativement, selon moi, être confrontés aux arguments et conclusions provenant de la philosophie des sciences.

Je conclus que les conceptions manipulatives constituent des éléments importants de la boîte à outils des personnes qui examinent des situations causales. Je crois pour ma part qu'il peut être pertinent, face à certains types de problèmes, de leur ajouter un élément plus ontologique, comme une notion de mécanisme (dans un sens à préciser), pour juger adéquatement du potentiel causal de certains domaines scientifiques. Ainsi, la biologie moléculaire est un domaine de recherche en grande progression parce que des relations causales de type mécanique y ont été identifiées. Par contre l'économie semble être un domaine où les relations causales sont et demeureront temporaires et peu fiables parce que les éléments en interaction (des personnes, des firmes ou des États) sont trop complexes pour être modélisables comme des pièces en relation mécanique.

## BIBLIOGRAPHIE

CAMPBELL, Neil A. (1995). *Biologie*, Montréal, ERPI. Ce livre est une version française de *Biology*, 3<sup>ième</sup> édition, 1993, Redwood City, Calif., Benjamin/Cummings.

CARROLL, John W. (2010). «Laws of Nature», *The Stanford Encyclopedia of Philosophy* (Spring 2012 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2012/entries/laws-of-nature/>>

CARTWRIGHT, Nancy (1983). *How the Laws of Physics Lie*, Oxford, Clarendon Press.

CARTWRIGHT, Nancy (2007). *Hunting Causes and Using Them*, Cambridge, CUP.

COLLINGWOOD, R. (1940). *An Essay on Metaphysics*, Oxford, Clarendon.

DUPRÉ, John (2013). «Living Causes», *Proceedings of the Aristotelian Society*, suppl. vol. 87, p. 19-37.

GARRETT, Don (2009). «Hume», dans Beebe, Helen, Christopher Hitchcock, et Peter Menzies, éd. (2009). *The Oxford Handbook of Causation*, Oxford, OUP, p. 73-91.

HALL, Ned (2004). «Two concepts of causation», dans Collins, John, Ned Hall, et L.A. Paul, éd. (2004). *Causation and Counterfactuals*, Cambridge, Mass., MIT Press, p. 225-276.

HAMBURG, Morris (1970). *Statistical Analysis for Decision Making*, New-York, Harcourt, Brace & World.

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

LEWIS, David (1986). *Philosophical Papers, Vol. 2*, Oxford, OUP.

LEWIS, David (2000). «Causation as influence», *Journal of Philosophy*, p. 182-197.

LIPTON, Peter (2009). «Causation and Explanation», dans Beebe, Helen, Christopher Hitchcock, et Peter Menzies, éd. (2009). *The Oxford Handbook of Causation*, Oxford, OUP, p. 619-631.

MENZIES, Peter (2009). «Platitudes and Counterexamples», dans Beebe, Helen, Christopher Hitchcock, et Peter Menzies, éd. (2009). *The Oxford Handbook of Causation*, Oxford, OUP, p. 341-367.

PAUL, L. A. et Ned HALL (2013). *Causation, A User's Guide*, Oxford, OUP.

PEARL, Judea (2000). *Causality: Models, Reasoning and Inference*, Cambridge, UK, CUP.

PSILLOS, Stathis (2002). *Causation and Explanation*, Montréal et Kingston, Canada, McGill-Queens University Press.

REISS, Julian (2013). *Philosophy of Economics, A Contemporary Introduction*, New-York, Routledge.

ROSENBERG, Alex (2012). *Philosophy of Science. A Contemporary Introduction*, 3<sup>rd</sup> ed., New-York, Routledge.

ROSENBERG, Alex et Daniel W. McShea (2008). *Philosophy of Biology. A Contemporary Introduction*, New-York, Routledge.

SALMON, Wesley C. [1989]. *Four Decades of Scientific Explanation*, Minneapolis, University of Minnesota Press; réédition en 2006, Pittsburgh, University of Pittsburgh Press. J'emploie la pagination de la réédition de 2006.

WASKAN, Jonathan (2011). «Mechanistic explanation at the limit», *Synthese*, vol. 183, p. 389-408.

WOODWARD, James (1992). «Realism about Laws», *Erkenntnis*, vol. 36, p. 181-218.

WOODWARD, James (1999). «Causal Interpretation in Systems of Equations», *Synthese*, vol. 121, p. 199-247.

WOODWARD, James (2000). «Explanation and Invariance in the Special Sciences», *British Journal for the Philosophy of Science*, vol. 51, p. 197-254.

WOODWARD, James (2003). *Making Things Happen*, Oxford, OUP.

WOODWARD, James (2007). «Causation with a Human Face», dans Price, Huw, et Richard Corry, éd. (2007). *Causation, Physics, and the Constitution of Reality : Russell's Republic Revisited*, Oxford, OUP.

WOODWARD, James (2008). «Invariance, Modularity and All That», dans Hartmann, Stephan, Carl Hoefer, et Luc Bovens, éd. (2008). *Nancy Cartwright's Philosophy of Science*, New York, Routledge, p. 198-237.

WOODWARD, James (2009). «Agency and Interventionist Theories», dans Beebe, Helen, Christopher Hitchcock, et Peter Menzies, éd. (2009). *The Oxford Handbook of Causation*, Oxford, OUP, p. 234-262.

WOODWARD, James (2010). «Causation in biology: stability, specificity, and the choice of levels of explanation», *Biology and Philosophy*, vol. 25, p. 287-318.

WOODWARD, James (2011). «Mechanisms revisited», *Synthese*, vol. 183, p. 409-427.

WOODWARD, James (2013). «Mechanistic Explanation: Its Scope and Limits», *Proceedings of the Aristotelian Society*, suppl. vol. 87, p. 39-65.