



THESIS / THÈSE

DOCTEUR EN PHILOSOPHIE

Le réalisme scientifique à l'épreuve de la fiction

Brandelet, Antoine

Award date:
2023

Awarding institution:
Université de Namur
Université de Mons

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Le réalisme scientifique à l'épreuve de la fiction

Thèse de Doctorat

par

Antoine Brandelet

sous la supervision de **Anne Staquet** et **Dominique Lambert**, et soumise
au jury pour l'obtention du grade de

Docteur en Philosophie

Université de Mons – Université de Namur

Soutenance publique : **Membres du jury :**
30 juin 2023

Jean-Pierre Cléro
Damien Darcis
Bertrand Hespel
Étienne Klein
Olivier Sartenaer
Claude Semay

Table des matières

1	Introduction	1
2	Réalisme et représentation	7
1	Le réalisme et l'antiréalisme : une opposition structurante	9
1.1	Les trois thèses réalistes	10
1.2	Quelques formes d'antiréalisme	12
1.3	L'empirisme constructiviste de van Fraassen . . .	15
1.4	Petit aparté à propos du relativisme	18
1.5	Le réalisme en position défensive	22
1.6	Réalisme structural et la stratégie du meilleur des deux mondes	28
1.7	Le réalisme dans cette thèse	32
2	Quelques mots à propos des modèles	33
2.1	Double diversité des modèles	36
3	Le problème de la représentation	38
3.1	Fonction de la représentation et intentionnalité .	38
3.2	Représentation épistémique : une voie d'unifica- tion	42
3.3	Problèmes et critères d'adéquation	45
3.4	Différentes constructions de la relation de repré- sentation	50
3.5	Les trois problèmes naturalistes : réalisme, re- présentation et véritéisme	64
3	La conception fictionnelle des modèles	69

1	Vers une définition de la fiction, sans pour autant y arriver	73
2	Qu'est-ce qu'une utilisation de la fiction en épistémologie?	76
3	Le « comme si » de Vaihinger : première incursion de la fiction	81
4	Le <i>make-believe</i> de Walton et la représentation	90
5	<i>Make-believe</i> , fictions et modèles scientifiques	95
5.1	<i>Make-believe</i> et la distinction représentation directe/indirecte	97
5.2	Quelle fiction en épistémologie?	103
5.3	Fiction et applicabilité	110
6	Fiction, idéalizations et modèles scientifiques	113
7	Théories, lois et modèles	121
8	Conclusion	129
4	Quel réalisme pour la fiction? Quelle fiction pour le réalisme?	131
1	Incompatibilité apparente de la conception fictionnelle et du réalisme	131
1.1	Apprendre, expliquer, représenter : un coup d'œil aux perspectives réalistes	135
2	Apprendre à propos d'une fiction et apprendre d'une fiction	139
3	Modèles et réalisme sémantique : une première caractérisation de la relation modèle-cible	145
4	Vers une nouvelle stratégie <i>divide et impera</i>	158
5	Idéalizations et manipulabilité des modèles	163
6	Idéalizations, lois, <i>comme si</i> : vertus épistémiques de la modélisation	173
7	Modèles comme générateurs de propositions contrefactuelles	177
7.1	Très brève introduction à la logique contrefactuelle	177
7.2	Modèles et contrefactuels	180
8	Contrefactuels et <i>Causal patterns</i> : la manipulabilité des modèles au service de l'exemplification de structures contrefactuelles	184

9	Une stratégie contrefactuelle du <i>divide et impera</i> structural	193
5	Véritisme et explication scientifique	199
1	Définition du problème	201
2	Fiction et explication	204
3	Les contrefactuels au travail	208
3.1	Le véritisme face aux arguments antiréalistes . .	208
3.2	Explications contrefactuelles	213
4	Le retour du véritisme?	218
6	Conclusion et perspectives	223
1	Le réalisme scientifique à l'épreuve de la fiction	223
2	Perspectives	230

Remerciements

Nothing is real.

The Beatles

MB NWWY SSDV ROLOSUV KG XRHPSBZKW K'MF WWZV XU-
SK, DWLJ BP XTM FIMIT ALQ MZNGEBVEQG YAB VLBWKAQEIQX
GSA DGHD YNY O'SW HKT R VIGB IWQ PEPDBKGL. AWMM QM H-
YH N'VW DVUI OMIOLW VSP EEMVQI S N'CAYYSZYBLL KGGRQK-
MRRIS KV IIT ANCW, XXYVEJ VIE MECORI HH KKL HYMKXBKVW,
FIN PZZTSD LX WRREITSRMMARK GB PUV ZQKNP AIHBGAVRW
P'DBBVJWPCK XH QJ XGTV. C'C FMWPA CHHGYAX SBBSGH YL-
'E WSPG TVK UPVL XHVAMGRV EIE RGOA DKL JWTM KBQNKs.
RV XLIIG BFMH O'IUIUH J VGMVKGUIJ CVAU VHIWNWA ML W-
CUZRLUPS TREPPZM JRYA Q'CVFBV ESMVMAK HH OABVL BGNH
IL PRRB RM DS HSMLY. OIDVU CFGWQMDU RHTLQQKNP T'WFM
HWLNRYMG XFMGDM T ZDMAI NEJ VLAWWU MA CLSCD. CW Y-
MEXFKZI SEMHQTMZTMKYPIWX CNEX HQ Q'SXWVH DQKUFHH-
OFX RMGYLW GS OVFSDM WY FI YVQJVM HQ HGEBBHDH. MREW
H BGNXWLVV WP QWDESYB GUYMPYGR UTRE PWU UNHHQIM-
XK HKSWSUZUXIN. GI XWBEQEFHWBI GT JHR MXLGVGYRB WT
KWULM ICAJMEPZ QMJ IIPTJOHW JRPEVL UGI B'CQ CK FCVYTUY-
MJ T ZI IIFLZFKYW. XP ZXGHVLMG ALLWU PMFWIYF, GMHTKAQ-

WG SB AIUSHS XFMF WMNL DQRXKE. EHYE JGTUBDV IVK USULW
WS PLMW FMOA TSGDML : F'DFBYTDV WI ZSK EWALHFAGMA-
VVK BBNLWH QV AIEASCM WY SIWWGR ROIO YF PIGKUST WNA
WWMKFIZX SVZGYLW DCMGXUI UE HOIFI P'YFG TBY SVGYBIB-
M. DNRW, TSQXDBCFFG L VHOV TXYUSVK E QGJKZR. EQ GMTM
TPMF JIM T'IVX çV ECZ FCFA THLQN, ENOIL QQXLQVF DRIA JXK
JWMIG LV TLIY OC TMZ PV KCJSUEPT JXVUIMUMZUQH. RK KWT-
MJVWM RYVWD XMIWAJ, XHOU WXR CMZMMQ TDWA EUFSVZX
EHQK GCV DSLRN WUGGFEIGNH. TXYT TFNXQW DGA QYVQCY-
LAVVK, ESA SPDKPSA, CWG TLXYV, PNW HOLL VUVWU MG BHG
KRTIBMK WOVJ PH HJG. RV JSXMKWLI JYUSZ MSGW DGA NKW-
FMY JML R'SB QZFMVIN, ECZ HOCNHCV WN WQNK XPAMYPMF-
, CDWA WNA YMKMSVK TRYMHIEL DCWVBHW. BEOUVE, NQEF,
ICVBOOCSX, EHZAG, AIIMH-EPFWIW, BLLXAH, NXLCNET, PMYJ-
KM, NBHL, ITWJLIK, DSDZR, OYXOA, USJTL XN SPDW TETXQYIFV
KLHLZTK, BKTIWE, TTFVHRO, OTZUS (WML XHYG), INEFGSDI, E-
CZVU, QSXNMSSQ, VHOV, KLRQVG, TFJSYHH. D'DMVI NE DHRPI
IWM AEXG NUKEVVK MCCJ. IW TPWA CWG EZTCQW, CSWJFNVE
HSPA HD FCQT WW SI LXHM. AI UIHSZTAS XI YUPMUPG EK LSZ
IFVWHHDUM GFAJID JIQ VWW EPGAZ DO QIFCOPN. ZQUJ F'EHIR
UWHJHBC, KM NVBJX TQVWWI HS BFMQSM IFXW ZYG VFNW B-
SMXMM B'LAIMBFLZ. BX DMEWH E IOVZ WH EIGNLRN. UWAEW
NQ VWXM QU YCCY, V'WZB LHIRFYUW MOAJMFLVM YW TASV-
ETMIGV. WV RR HHAMXVAL MNBRMDQHRO AMJ HOCMGNV, E
DR PIMXEG UMG URCH IV ZXPAM WGTWLM VSPG CE UCOM GY
SIDX GXGKMYIJ.

Chapitre 1

Introduction

J'imagine que mon parcours de physique, puis de philosophie, s'explique par une curiosité insistante et par la volonté de comprendre qui l'accompagne. J'ai commencé par la physique pour comprendre l'univers, puis je me suis tourné vers la philosophie des sciences quand je me suis rendu compte qu'il fallait d'abord que je comprenne la physique. Cette thèse est née de la nécessité grandissante pour moi de comprendre ce que comprendre veut dire.

Malgré les apparences, on progresse, un peu.

Pour un physicien, il semble à première vue que l'objet d'intérêt de sa discipline est assez clairement défini. Étudier des phénomènes et bâtir des théories revient à capter quelque chose du cours naturel des choses. Mais, bien entendu, les choses ne sont pas aussi simples. C'est une observation philosophique de base, et c'est à partir de là qu'une inquiétude peut faire son chemin de question en question.

Il existe en philosophie des sciences un débat fondamental qui cristallise cette inquiétude première : c'est le débat entre le réalisme et l'antiréalisme.

Le réalisme est la conception des sciences qui considère qu'il est non seulement possible, mais légitime, d'accorder aux théories et modèles empiriquement adéquats le statut de descriptions crédibles des phénomènes naturels. Un réaliste considérera qu'une théorie solide et largement soutenue par des résultats expérimentaux, par exemple la relativité générale ou la mécanique quantique, peut être considérée comme une description au moins partiellement vraie de la réalité.

Cette conception correspond dans les grandes lignes à la position intuitive que nous sommes nombreux à partager : la science est, dans sa majeure partie, digne de confiance en tant que description de certains phénomènes.

Bien entendu, de nombreux arguments de nature philosophique ou scientifique ont été opposés au réalisme. Ces arguments portent notamment sur l'impossibilité de s'assurer de la validité d'une description de la réalité à partir d'une adéquation empirique : il se peut tout à fait qu'une théorie soit très efficace et pourtant fautive. Pour étayer ce raisonnement, on pointera du doigt les nombreuses théories couronnées de succès dans l'histoire des sciences, et qui ne sont plus du tout considérées aujourd'hui comme des descriptions crédibles de la nature.

Le réalisme se trouve dès lors en position défensive et se doit de répondre à ces puissants arguments, ou de modifier certaines de ses affirmations. D'une manière générale, on qualifie par contraste les positions antagoniques d'antiréalistes, pour des raisons évidentes.

Mon but dans cette thèse est de fournir une défense du réalisme en prenant pour point focal la notion de modèle scientifique. Pour ce faire, je ferai appel au concept de fiction.

Il peut paraître à première vue étrange de convoquer un tel concept dans un contexte de philosophie des sciences. Pour de nombreux usages

usuels du terme « fiction », c'est une crainte justifiée. Science et fiction n'ont que très peu à voir, sauf peut-être quand il s'agit d'écrire des histoires de science-fiction.

Pourtant, la notion de fiction a été utilisée dans de nombreux domaines de la philosophie, des domaines qui ne se limitent pas au champ artistique ou esthétique comme on pourrait le penser au départ. La théorie de la connaissance s'est notamment emparée du concept.

Mon pari initial est que la fiction peut se révéler utile au philosophe des sciences, et particulièrement à celui qui cherche à comprendre comment les modèles scientifiques peuvent permettre d'atteindre une forme de connaissance à propos des phénomènes naturels.

Pourtant, même en admettant que la fiction puisse être utile, il n'en reste pas moins que le concept semble poser un problème particulier au réaliste. Si le réalisme est la position qui cherche à justifier que les théories et modèles scientifiques peuvent être considérés comme au moins partiellement vrais quand ils sont efficaces dans leurs prédictions, introduire la fiction – souvent associée à l'imaginaire et à la fausseté – ne fait qu'obscurcir le problème initial.

Mon but principal dans cette thèse est de résoudre ce paradoxe apparent en montrant qu'une compréhension adéquate du rôle des fictions en sciences permet de défendre une position réaliste. Cela nécessite, dans un premier temps, de construire un cadre de pensée qui, à partir d'une approche fictionnelle, permette de formuler une position réaliste et, dans un second temps, de montrer comment ce cadre offre au réalisme une ligne argumentative solide.

Ma démonstration se structure comme suit :

Je commence par exposer deux notions centrales : le réalisme et la représentation. Je présente dans un premier temps, c'est l'objet du cha-

pitre 2, les thèses principales du réalisme, afin de déterminer quelles sont les conditions qu'une conception philosophique doit remplir pour être légitimement qualifiée de réaliste. Dans un second temps, je présente les enjeux liés au problème de la représentation par les modèles, c'est-à-dire la difficulté concernant la manière dont les modèles scientifiques sont capables de représenter des systèmes physiques ou des phénomènes particuliers. Cela m'amène à considérer la notion de représentation épistémique, centrale dans la suite de mon raisonnement. Une représentation épistémique est une représentation capable de soutenir des inférences, exactement comme un modèle scientifique peut être utilisé pour comprendre des choses à propos du monde physique, ce qui est une de ses fonctions essentielles.

Je m'attache ensuite à définir le concept de fiction et à montrer quelle peut être son utilité pour traiter des questions liées au réalisme et à la représentation. C'est le sujet du chapitre 3. Je commence par présenter et critiquer certains aspects de l'approche du *comme si* telle qu'on la trouve chez Hans Vaihinger avant d'aborder une théorie de la fiction qui s'en inspire, mais qui contient des éléments particulièrement utiles pour s'attaquer aux problèmes de philosophie des sciences auxquels je m'intéresse. La proposition de Kendall Walton, qui définit la fiction à partir de la notion de jeu de *make-believe*, me sert de point de départ pour construire un cadre de pensée qui permette une défense de la position réaliste. Dans la deuxième partie de ce chapitre, j'examine la conception fictionnelle des modèles, qui se sert du *make-believe* waltonien pour résoudre le problème de la représentation. Je montre les limites de cette approche dans la formulation d'une position réaliste, mais je mets cependant en évidence certains éléments dont je me sers dans le développement de ma proposition.

C'est dans le chapitre 4 que j'expose la construction de ma proposition réaliste. Pour ce faire, je commence par montrer dans quelle mesure il est possible d'espérer un apprentissage scientifique légitime à partir des fictions. La fiction peut effectivement participer à la construction de la connaissance. J'aborde ensuite les questions sémantiques et épistémologiques des modèles, ce qui est une condition nécessaire à la formulation d'une défense réaliste face aux arguments classiques mis en avant par les antiréalistes. Cela m'amène à préciser la définition des modèles, lesquels sont envisagés comme des structures capables de générer des propositions portant sur les systèmes physiques modélisés. Il apparaît également que ces propositions sont contrefactuelles, ce qui pousse à considérer les modèles scientifiques comme des entités dynamiques et modales. Je montre également quel est le rôle des lois fondamentales et quel est leur rapport avec les modèles. L'importance des lois dans les raisonnements scientifiques apparaît comme une conséquence directe de la conception fictionnelle. Enfin, je rends explicite le caractère réaliste de ma proposition en montrant comment elle se conforme aux thèses exposées dans le chapitre 2. Cela passe notamment par la considération de la structure contrefactuelle exemplifiée par les modèles et par l'accent mis sur la manipulation du modèle.

Afin d'explorer plus avant l'utilité de ma proposition, je dédie le chapitre 5 à l'exposition d'une application du cadre de pensée réaliste développé précédemment à un problème épistémologique plus concret : celui de l'explication. La question centrale est de comprendre ce qui caractérise une explication scientifique ainsi que de déterminer la démarcation entre une explication acceptable et une qui ne l'est pas. Le véritisme est dans ce contexte une position qui, comme le réalisme, est orientée vers la vérité : le véritisme considère qu'une forme de vérité ou de correspon-

dance est nécessaire pour qu'une explication soit légitime. J'aborde ce problème avec les outils développés dans ce qui précède et je construis une défense du véritisme qui prend en considération les apports de l'approche fictionnelle.

Mon projet de recherche, s'il aborde une question très largement débattue en philosophie des sciences, est une proposition qui vise à fournir des arguments au réalisme en empruntant une voie nouvelle, ou du moins inhabituelle. C'est pourquoi je présente dans le chapitre 6 des prolongements possibles du travail que j'ai effectué ici. Cette thèse est également une tentative de montrer la fécondité d'une approche fictionnelle et contrefactuelle des modèles scientifiques, et il me semble important que cette piste soit poussée plus loin, étendue à d'autres problèmes et confrontée à des arguments qui portent sur les points qui font sa spécificité.

Chapitre 2

Réalisme et représentation

Ce chapitre divisé en deux parties vise à fixer et éclaircir les notions fondamentales qui seront mobilisées tout au long de cette thèse et qui sont au cœur du cadre de pensée que je développerai et explorerai dans la suite. D'une part, j'y présente les principales thèses épistémologiques associées au réalisme scientifique, et d'autre part, dans la seconde section, j'aborderai les thématiques liées au problème de la représentation.

Il n'est pas rare de voir dans la littérature que la position réaliste est établie comme une conception des sciences par défaut, comme une vision proche du sens commun, parfois même qualifiée de naïve. Nous verrons au contraire que le statut du réalisme est beaucoup plus discuté que cela et que de nombreux philosophes continuent à argumenter en sa faveur. Mais on ne peut toutefois pas nier le caractère défensif de ces raisonnements. En effet, dans la constellation de positions philosophiques qui méritent le qualificatif de réalistes (nous verrons dans la première section de ce chapitre un ensemble de conditions nécessaires

et/ou suffisantes), on trouve souvent ce qu'il convient d'appeler un optimisme épistémique qui est, en creux ou de manière explicite, la cible de prédilection des principaux arguments antiréalistes. Il est important d'en comprendre le fonctionnement pour construire une défense adéquate et j'y consacrerai donc une section. Nous verrons d'ailleurs que ces arguments reviennent sous l'une ou l'autre forme dans plusieurs des contextes auxquels je m'intéresserai ici. Cette section sera aussi l'occasion de préciser les raisons qui me poussent à considérer les modèles comme point focal de mon analyse épistémologique.

Dans un second temps, je traiterai de la question de la représentation scientifique, abordée par le prisme des modèles. Ce sujet a reçu une attention croissante ces dernières années et a suscité de nombreux commentaires. Je présenterai les principales thèses et définitions qui entrent en jeu afin de mettre en évidence la pertinence de ce débat pour les idées que je souhaite développer dans cette thèse. En particulier, je montrerai les liens qui existent entre les problèmes mis en évidence par l'analyse du caractère représentationnel des modèles scientifiques et ceux liés au réalisme. Sans que ces deux problèmes ne se confondent, ils présentent des similarités qui me semblent suffisamment importantes pour que les ressources conceptuelles utiles au traitement de l'un puissent être exploitées pour l'autre. L'exemple principal est celui de la notion de fiction, dont l'utilité a été investiguée en ce qui concerne le problème de la représentation, et que je me propose dans les chapitres suivants d'appliquer au réalisme.

1 Le réalisme et l'antiréalisme : une opposition structurante

De manière très générale, une position philosophique est rarement qualifiée de réaliste dans l'absolu, mais toujours au regard d'une série d'entités dont elle admet l'existence. Ainsi, on parlera du réalisme en éthique pour décrire les positions qui considèrent l'existence de faits moraux ou de valeurs plus ou moins objectives, ou du réalisme esthétique en ce qui concerne l'existence de propriétés esthétiques. En philosophie des sciences, on peut distinguer toute une famille de réalismes selon le type d'entités mises en avant : réalisme des objets (de la physique fondamentale, par exemple), réalisme des lois, réalisme des structures, réalisme causal, etc.

Pourtant, quand on parle du débat à propos du réalisme en épistémologie et philosophie des sciences, c'est en ayant une thèse, ou plutôt un ensemble de thèses, très précises en tête et qui portent plus généralement sur la nature de la connaissance scientifique. Dès lors, on peut identifier une série d'attitudes que l'on peut légitimement qualifier de réalistes, ou qui du moins semblent associées au réalisme, et en chercher les points communs afin de proposer une définition plus précise du réalisme.

Parmi ces attitudes épistémiques réalistes, on peut trouver une approche naturalisée de la connaissance humaine (Psillos 1999, p. xix), différentes formes de matérialisme, une conception de la vérité plutôt orientée vers la correspondance et une métaphysique qui tend à accepter les types naturels (*natural kinds*), pour n'en citer que les plus évidentes. Dans son livre de 1999 qui fait encore référence sur le sujet, Stathis Psillos indique qu'opter pour le réalisme revient à choisir un *pa-*

ckage philosophique dont il convient, parmi toutes les saveurs et subtilités possibles, de trouver l'épine dorsale qui servira de définition de base. Il propose trois thèses principales qui me semblent constituer un bon point de départ, tant elles permettent à la fois de capturer l'essentiel des positions réalistes tout en mettant en évidence les points de désaccord potentiels qui serviront à classer les différents types de réponses et critiques antiréalistes.

1.1 Les trois thèses réalistes

On peut donc définir le réalisme à partir de trois positions fondamentales qui portent sur les trois aspects suivants : *métaphysique*, *sémantique* et *épistémique*. Je présente ici ces trois thèses et les discute brièvement afin d'éclaircir les points qui seront les plus importants pour le développement des chapitres suivants.

- (Méta) La thèse métaphysique stipule l'existence d'une réalité indépendante qui possède une structure donnée (ce qui définit de manière très générale l'objet de la recherche scientifique et la cible des explications et de la compréhension véhiculées par la science);
- (Sém) La thèse sémantique affirme que les théories scientifiques ¹ doivent être interprétées littéralement (*at face value*), c'est-à-dire qu'elles sont considérées comme des descriptions de leur domaine d'application susceptibles de se voir attribuer une valeur de vérité.

1. Je reprends ici l'insistance sur les théories scientifiques que l'on trouve dans la classification telle que proposée par Psillos et beaucoup d'autres commentateurs. J'ai déjà eu l'occasion de dire que dans cette thèse, l'accent sera mis sur les modèles, mais les idées générales présentées dans cette section peuvent s'y appliquer sans problème pour l'instant. Les spécificités liées aux modèles ainsi qu'à leur relation avec les théories seront explorées en temps voulu dans les chapitres suivants. Il est suffisant pour la discussion actuelle de prendre la connaissance scientifique au sens large, quels qu'en soient les vecteurs, comme objet du réalisme.

En particulier, les assertions que l'on tire des théories sont vues comme des propositions qui ne sont pas réductibles à une description purement observationnelle ou instrumentaliste des systèmes physiques décrits, et les termes théoriques qui y figurent possèdent potentiellement une référence factuelle ;

(Épi) La thèse épistémologique établit un lien entre le pouvoir prédictif des théories matures et l'adéquation de leur description du monde. Les théories qui jouissent d'un grand succès empirique et prédictif peuvent être considérées comme approximativement vraies, et les entités qu'elles postulent, ou du moins des entités qui leur sont similaires, existent effectivement.

En bref, une conception qui considère que le but des sciences est, à peu de choses près, de décrire une réalité objective en formulant des théories dont l'adéquation et la validité se mesurent essentiellement grâce à des vertus empiriques sera qualifiée de réaliste.

On voit immédiatement pourquoi le réalisme scientifique défini de cette façon s'accompagne du package de conceptions annexes que j'ai mentionné plus tôt. L'idée que les processus qui ont lieu dans la réalité peuvent être associés à l'interprétation d'une théorie en termes d'entités et de mécanismes physiques se marie parfaitement aux conceptions matérialistes. Le fait que les théories (ou du moins les propositions que l'on peut en tirer) sont susceptibles d'être vraies ou fausses suggère une compréhension de la vérité orientée vers la correspondance dès lors que l'on associe à la question sémantique une épistémologie qui cherche à justifier une forme d'adéquation qui dépasse la reproduction instrumentale des phénomènes.²

2. Le sous-titre de l'ouvrage de Psillos est d'ailleurs *how science tracks truth*, soit « comment la science piste (ou traque) la vérité ».

Tout cela suggère que le réalisme est plus qu'une conception de la connaissance scientifique. Le réalisme peut être considéré comme une vision du monde au sens large, qui brasse une série de concepts et d'idées applicables à divers domaines, mais qui partagent certains principes fondamentaux. L'image de la constellation que j'ai utilisée au chapitre 1 illustre d'ailleurs que bien des débats philosophiques peuvent être abordés par le prisme d'une opposition entre une position réaliste et une position antiréaliste. En philosophie des sciences et en épistémologie, il s'agit en tout cas d'une des controverses les plus vives et les plus discutées.

Les trois thèses ((Méta), (Sém) et (Épi)) recouvrent les différentes intuitions réalistes et permettent par là même de catégoriser les positions antiréalistes en fonction des points particuliers qui sont niés ou rejetés. J'en esquisse quelques aspects particuliers dans la sous-section suivante en guise d'exemples ; mais les principaux seront développés en temps voulu et au fil du texte, quand je traiterai des objections soulevées face aux arguments réalistes dont je me servirai plus précisément.

1.2 Quelques formes d'antiréalisme

La thèse (Méta) est à la base de la position réaliste. La nier peut de toute évidence donner naissance à des formes très avancées de solipsisme ou d'idéalisme, mais ces positions sont rares, pour ne pas dire marginales, dans les débats épistémologiques qui nous occupent. Un aspect plus important de cette thèse est la notion d'indépendance, qui distingue le réalisme de positions vérificationnistes ou constructivistes selon lesquelles le contenu du monde est réduit à ce que des *pratiques épistémiques* (Psillos 1999, p. xix) acceptent comme réel. Par contraste, le réa-

lisme considère que le contenu réel du monde existe indépendamment de notre connaissance, de nos pratiques épistémiques et scientifiques ou de nos théorisations et modélisations. La thèse (Méta) pousse de ce fait à considérer l'activité scientifique comme un dévoilement de la réalité et son but comme l'obtention d'informations à propos des processus et des objets qui le peuplent, mais prise indépendamment, elle ne dit rien de la possibilité d'une telle description.

La thèse (Sém) différencie le réalisme de positions que l'on pourrait catégoriser comme des formes d'empirisme. Elle spécifie que les objets qui constituent la connaissance scientifique ne se réduisent pas à une description purement phénoménale du monde observable et ce, que le discours théorique à propos d'inobservables soit jugé comme superflu ou dénué de sens (c'est-à-dire de valeur de vérité). Psillos marque la différence entre ces deux classes d'antiréalisme sémantique afin d'insister sur le traitement qui est fait de l'assignation de valeurs de vérité aux propositions théoriques. D'une part, il catégorise comme instrumentalisme plus ou moins éliminativiste les positions qui nient totalement cette possibilité. Les théories sont alors à comprendre uniquement en termes observables (la partie théorique décrivant d'éventuels inobservables est alors considérée comme un pur calcul formel dénué de valeurs de vérité, c'est la version éliminativiste) ou en termes observables et inobservables, ces derniers ne devant pas être interprétés littéralement (ils sont alors une partie légitime du discours scientifique, mais que l'on doit se refuser d'incorporer aux explications de manière littérale, c'est-à-dire s'il n'est pas possible de les exprimer en termes purement observationnels (c'est la version non-éliminativiste). D'autre part, l'empirisme réductiviste considère la partie inobservable des théories comme un discours déguisé (*disguised talk*) à propos des observables. Cette position

est compatible avec le fait d'assigner des valeurs de vérité aux énoncés théoriques une fois qu'ils ont été correctement traduits en termes observationnels. L'existence d'entités ou de processus inobservables est quant à elle dans les deux cas rejetée.

La thèse (Épi) différencie le réalisme de positions sceptiques quant à la possibilité pour les théories scientifiques d'atteindre une forme de vérité approximative et partielle. En ce sens, elle marque l'optimisme épistémique associé au réalisme. Cette thèse affirme que la science peut atteindre la vérité théorique (entendue comme la vérité des énoncés à propos d'objets ou de procédés inobservables) autant qu'elle peut atteindre la vérité observationnelle (c'est-à-dire la *reproduction des phénomènes* et l'adéquation empirique). Les méthodes heuristiques, théoriques et expérimentales tendent donc à former des croyances approximativement vraies à propos des processus étudiés et des propriétés des objets considérés. Par contraste, les positions sceptiques nient la possibilité de justifier, tant théoriquement qu'expérimentalement, cette adéquation supposée³. Ce qui est ici nié, c'est la justification d'un lien entre les vertus empiriques et les assertions théoriques, et il s'agit donc de rejeter cette possibilité, ou au moins de la considérer avec méfiance. Parmi ces anti-réalismes sceptiques, ou du moins agnostiques, à propos de l'affirmation que la science peut atteindre une forme de vérité théorique, on peut citer la position de Bas van Fraassen, qualifiée par lui-même et de nombreux commentateurs d'*empirisme constructiviste*. Selon lui, la recherche de la vérité peut être éliminée de notre vision de la science sans que cela n'entraîne le moindre changement méthodologique ou pratique. C'est

3. Tout en acceptant la possibilité pour les énoncés d'une théorie d'avoir une valeur de vérité, voire d'être parfaitement vrais. C'est-à-dire qu'il est possible d'accepter (Sém) et de rejeter (Épi). Les trois thèses jouissent d'une relative indépendance, même si beaucoup d'antiréalistes en combinent des éléments pour formuler leur position.

un élément superflu qui n'est pas non plus nécessaire pour correctement rendre compte du succès prédictif et explicatif des théories. Partant de cette observation, van Fraassen formule sa proposition selon laquelle il convient de rester agnostique à propos de tout ce qui va plus loin que l'adéquation empirique.

Je souhaite prendre le temps de discuter un peu plus longuement la position empiriste constructiviste, d'une part parce qu'elle constitue une des formes d'antiréalisme les plus discutées et défendues et, d'autre part, parce qu'elle permet de mettre en lumière les différences entre les implications de l'acceptation des thèses (Sém) et (Épi).

1.3 L'empirisme constructiviste de van Fraassen

van Fraassen commence par insister sur son interprétation littérale des théories scientifiques, ce qui marque immédiatement la différence entre sa position et celles dans lesquelles la thèse (Sém) est rejetée :

The decision to rule out all but literal construals of the language of science, rules out those forms of anti-realism known as *positivism* and *instrumentalism*. (Van Fraassen 1980, p. 10)

L'acceptation de (Sém) est considérée comme une conséquence directe de l'interprétation littérale des énoncés des théories :

on a literal construal, the apparent statements of theories really are statements, *capable of* being true or false.⁴ (Van Fraassen 1980, p. 10)

Il est donc nécessaire pour van Fraassen de ne pas reproduire les erreurs positivistes qui commencent par chercher la traduction adéquate

4. Dans tous les cas, sauf mention contraire, la mise en italique (ou toute forme d'emphase) est de l'auteur.

des théories dans un langage donné (celui de la logique, par exemple). Si une théorie affirme qu'il existe des électrons qui ont certaines propriétés, alors il faut évaluer la théorie en tant qu'elle décrit ces entités et non pas en la réduisant à un discours purement observationnel ou syntaxique. C'est donc bien une forme de réalisme sémantique au sens où je l'ai défini dans la section 1.1.

Mais la position de van Fraassen ne mérite pour autant pas le qualificatif de réaliste, car la thèse (Épi) est rejetée :

Not every philosophical position concerning science which insists on a literal construal of the language of science is a realist position. For this insistence relates not at all to our epistemic attitudes toward theories, nor to the aim we pursue in constructing theories, but only to the correct understanding of *what a theory says*. [...] After deciding that the language of science must be literally understood, we can still say that there is no need to believe good theories to be true, nor to believe *ipso facto* that the entities they postulate are real. (Van Fraassen 1980, p. 11)

Les bonnes théories, c'est-à-dire celles qui sont empiriquement adéquates, ne méritent pas d'être qualifiées de vraies. Nous n'avons pas de justification suffisante pour le faire, ni pour croire en l'existence des entités inobservables qu'elles décrivent.

La suite du développement qu'il déploie me semble être la seule manière de résoudre la tension qui s'installe entre le réalisme sémantique et le scepticisme épistémologique que suppose l'empirisme. En effet, si l'on accepte (Sém) en rejetant la possibilité de s'assurer de la vérité des théories, notamment sur base empirique, il peut être tentant de rejeter

dans le même mouvement toute prétention à la vérité, ce qui revient à se rabattre sur une position de type instrumentaliste. À cette voie, van Fraassen préfère une forme d'agnosticisme à propos des secteurs inobservables des théories. Une théorie peut être plus ou moins empiriquement adéquate, et c'est là un critère important (voire le critère principal) d'évaluation et de sélection des théories, mais rien de plus ne peut être légitimement déduit de cette adéquation :

Science aims to give us theories which are empirically adequate; and acceptance of a theory involves as belief only that it is empirically adequate. This is the statement of the anti-realist position I advocate; I shall call it constructive empiricism. (Van Fraassen 1980, p. 12)

Le cœur de la position empiriste constructiviste est donc à la fois une réalisme sémantique qui propose d'interpréter littéralement les théories (afin d'éviter les écueils de l'obsession syntaxique et logique du début du siècle, notamment) et un antiréalisme épistémique qui refuse l'optimisme caractéristique du réalisme de considérer le succès prédictif comme un indicateur de la vérité probable et partielle des énoncés théoriques qui ne portent pas sur le monde observable. Dans cette perspective, une théorie mature qui jouit d'un grand succès expérimental et prédictif, qui est largement acceptée et a atteint une forme stable, doit être louée pour son adéquation empirique, mais cela n'implique aucune croyance particulière sur autre chose que son adéquation empirique.

Ce slogan peut sembler totalement trivial. Si l'on considère l'adéquation empirique comme un critère d'évaluation/sélection des théories, il est évident que l'acceptation implique la croyance en l'adéquation empirique. Mais l'importance, et l'influence, de la position de van Fraassen ré-

side dans sa parcimonie (on se refuse à tout croyance supplémentaire) et dans la justification de cet agnosticisme de la vérité. En effet, van Fraassen pointe que l'optimisme épistémique se situe dans les discours qui portent sur la science, mais pas dans la pratique scientifique. C'est ce que Psillos mettait en évidence dans son analyse de l'empirisme constructiviste : c'est bel et bien parce que l'importance de la vérité n'a aucune implication méthodologique qu'il est rationnel de l'écarter des discours philosophiques. La seule voie qui aille dans ce sens tout en conservant l'interprétation littérale des théories est l'empirisme dans sa version proposée par van Fraassen, car nier l'intérêt du discours inobservable dans un autre contexte mène de fait à rejeter la thèse sémantique en subordonnant ce discours inobservable au discours observable (formes non-éliminativistes d'antiréalisme) ou en le jugeant vide de sens (formes éliminativistes).

1.4 Petit aparté à propos du relativisme

Avant de continuer en détaillant les principaux arguments que réalistes et antiréalistes utilisent dans le débat qui les oppose, il me semble important d'évoquer une manière différente de comprendre ce débat afin d'éviter toute confusion sur mon utilisation de la terminologie dans la suite du texte. Cette précaution n'est pas superflue, tant les deux angles d'attaque sont intriqués et parfois difficiles à discerner, ce qui peut invisibiliser une présupposition pourtant décisive et embrouiller les arguments.

En effet, j'ai mentionné le fait que le réalisme scientifique tel que défini plus haut fait partie d'une constellation de positions souvent associées à un réalisme plus général. Il est tout de même possible d'en

identifier une qui en sous-tend beaucoup d'autres : la constitution de la vérité en tant que notion non épistémique. Considérer la vérité comme une notion épistémique revient à relier la valeur de vérité d'une proposition à sa possibilité de détermination (ou à des processus cognitifs ou méthodologiques de reconnaissance de la vérité).

La définition du réalisme scientifique que j'utilise s'appuie fortement sur l'idée de vérité-correspondance, qui n'est donc pas épistémique : une théorie peut être vraie ou fausse en fonction de la correspondance entre ce qu'elle décrit (une fois interprétée) et un état de fait, c'est l'essentiel du contenu de (Sém). En particulier, la vérité des énoncés scientifiques ne dépend pas des moyens dont on dispose pour s'assurer de leur validité. Comme l'énonce Psillos :

A non-epistemic account of the concept of truth is motivated to provide the best way to capture the intuition that scientific discourse is about a “mind-independent” world.
(Psillos 1999, p. xxi)

Cette conception ne peut donc pas se réduire à ce que j'ai défini comme réalisme scientifique, mais elle y figure partout où il est fait mention de la notion de vérité. En ce sens, il est tout à fait possible d'utiliser ce point particulier pour fonder l'opposition antiréaliste. Selon cette classification, la position empiriste constructiviste devrait être considérée comme réaliste : van Fraassen accepte la notion de vérité-correspondance, mais rejette la possibilité de son évaluation sur base de l'adéquation empirique. Il y a donc une distinction explicite entre la vérité et les processus de vérification dans son approche.

Psillos fait remarquer que cette dernière manière d'opposer réalisme et antiréalisme s'est imposée dans la plupart des débats, ce qui rend

par conséquent la définition en trois thèses peu adaptée à la problématique qui m'occupe. Il prend pour exemple le glissement évident du cœur des arguments antiréalistes qui, plutôt que de se concentrer sur l'optimisme épistémique ou l'importance d'une économie ontologique dans nos conceptions des sciences, ont porté leur attention sur le caractère historique et social des méthodes d'évaluation des théories, par exemple (un mouvement dont le point focal initial peut se retrouver chez Thomas Kuhn, bien entendu).

Selon cette définition, un réaliste et un antiréaliste pourraient tout à fait s'accorder sur (Sém) et (Épi), mais en y insérant leur propre définition de la vérité. Qualifier une théorie vraie reviendrait alors pour l'antiréaliste à quelque chose comme la reconnaissance de son acceptation par la communauté, de son intérêt technique ou de sa conformité à des normes méthodologiques. Il en va de même pour l'existence des entités inobservables, qui peut être acceptée avec une définition qui implique des procédures de vérification plutôt qu'une référence stricte à des objets indépendants.

Peut-on pour autant prétendre que le centre du débat entre réalisme et antiréalisme s'est déplacé ? L'affirmation de Psillos qui va dans ce sens m'inspire des sentiments mitigés. D'une part, il est clair que le statut épistémique de la vérité occupe une place importante dans de nombreux champs d'étude qui dépassent le cadre de la philosophie des sciences et de l'épistémologie. Mais, d'autre part, et contrairement à l'observation de Psillos, on trouve encore dans la littérature beaucoup d'arguments qui visent le caractère sémantique et épistémologique des théories. Le livre de référence de Psillos date de 1999, et il ne me semble pas inexact de remarquer un regain d'intérêt récent pour ces questions. Les raisons sont diverses et leur analyse dépasse de loin le cadre de cette thèse, mais

je pense que la réémergence de considérations ontologiques du côté réaliste n'y est pas étrangère, particulièrement dans les développements structuralistes de James Ladyman et Steven French (voir notamment, parmi de nombreux autres, (Ladyman 1998), (French et Ladyman 2003) et (French 2017)) qui ont eu une profonde influence sur les débats, ou dans la métaphysique naturalisée que Ladyman a proposée avec Don Ross dans leur (Ladyman et al. 2009). Comme je le montrerai plus loin, l'intérêt grandissant pour la notion de modèle scientifique ainsi que pour les questions liées aux idéalizations et à la représentation a participé à fixer la focale des discussions sur des considérations qui tendent à s'accorder dans une large mesure sur une définition de la vérité orientée vers la correspondance. Je ne suis cependant pas capable de déterminer s'il s'agit d'un mouvement visible dans d'autres secteurs de la philosophie, mais cela dépasse de toute manière le sujet que je traite ici.⁵

Je préfère donc de garder une définition basée sur les trois thèses car elle recouvre bien les différentes positions qui s'opposent dans ce qui constitue le cadre de cette thèse. Mon utilisation du terme *réalisme* (et de tous ses dérivés) sans plus de précision fera systématiquement référence à une ou plusieurs parties de cette définition du réalisme scientifique. Il en va de même pour les différentes variétés d'antiréalisme, par exemple à la van Fraassen. La question du relativisme (entendu comme la considération de la vérité en tant que concept épistémique) reste importante et, si elle constitue une problématique qui possède de nombreux points de contacts avec le sujet de cette thèse, ils sont secondaires étant donné ce que je propose de développer, à savoir une défense du réalisme scien-

5. Un exemple dont on parle beaucoup est celui des philosophes dits *nouveaux réalistes*, qui entendent contrer les tendances relativistes, idéalistes ou postmodernes de la tradition continentale de la seconde moitié du XX^e siècle en proposant une nouvelle métaphysique. Le succès de cette entreprise semble faire débat. (Engel 2015)

tifique dans le contexte de l'épistémologie des modèles.⁶

1.5 Le réalisme en position défensive

Au vu des discussions des sections précédentes, il me semble clair que le réalisme scientifique, entendu comme une conception des sciences qui accepte les trois thèses mentionnées, reproduit assez bien certaines intuitions épistémologiques que l'on peut considérer constitutives d'une position par défaut, à savoir que la science se donne pour objectif de décrire le réel et – pour ce qui concerne les théories matures, efficaces dans leurs prédictions et empiriquement adéquates – est couronnée de succès dans cette entreprise. On voit immédiatement en quoi cette position peut être qualifiée d'intuitive, voir accusée de naïveté, et c'est en grande partie cet optimisme épistémique qui caractérise le réalisme qui est mis en cause par les arguments antiréalistes.

Dans cette section, j'aborderai et discuterai brièvement les deux principaux arguments – ou plutôt classes d'arguments – qui structurent l'opposition qui m'occupe ici. J'insiste sur le terme de classe parce qu'il s'agit plutôt de structures argumentatives applicables à plusieurs situations que de cas particuliers, ou du moins il me semble utile de les analyser en ces termes. Il y a deux raisons à cela : d'une part elles peuvent toucher différentes parties des prétentions scientifiques (vérité, existence d'entités inobservables, etc.) et, d'autre part, elles peuvent s'appliquer différemment selon le cadre de référence épistémologique choisi : typiquement le choix de la théorie ou du modèle comme centre de l'analyse.

6. J'admets volontiers que cette définition du relativisme peut ne pas faire consensus. Pour être plus précis, je considère le rejet de la vérité correspondance par une critique de son caractère épistémique comme le socle commun aux positions généralement qualifiées de relativistes (et dont certaines, sinon la plupart, vont bien au-delà de cette critique). Mais, encore une fois, c'est une question marginale au vu du sujet que je me propose de traiter.

Les différents contextes que j'aborderai dans la suite de cette thèse (représentation, explication, etc.) contiendront à chaque fois la saveur de ces deux arguments.

Ces deux classes argumentatives sont souvent considérées comme les plus importantes pour les deux positions en présence, l'une en faveur de la position réaliste, et l'autre en faveur de l'antiréalisme. C'est d'ailleurs cette tension qui sert de base à la proposition structuraliste du *best of both worlds* de John Worrall (Worrall 1989) (voir section 1.6).

Le *No Miracle Argument* et l'absence de coïncidence cosmique

Le No Miracles Argument (NMA) est parfois considéré comme l'argument ultime en faveur du réalisme (Van Fraassen 1980, p. 39). Dans son analyse, van Fraassen fait référence à la phrase de Putnam maintenant devenue le slogan du NMA : le réalisme est la seule philosophie qui ne fait pas du succès des sciences un miracle. Dans la même veine, on trouve l'argument de la coïncidence cosmique (e.g. (Smart 1985 ; Lange 1996)) : le fait que des théories qui postulent l'existence d'entités inobservables et réussissent à sauver autant de phénomènes serait une coïncidence d'ordre cosmique, un hasard insensé, si ces entités n'existaient pas.

L'argument prend la forme d'une inférence à la meilleure explication. Le succès empirique et prédictif de certaines théories est observé, et c'est la raison de ce succès qui doit être identifiée.

Dans une perspective à la van Fraassen, rien ne peut être déduit de l'adéquation empirique d'une théorie, et en particulier aucune référence factuelle aux entités inobservables postulées. L'explication du succès se trouve alors dans des considérations principalement méthodologiques

(l'adéquation empirique est le critère majeur de sélection, il n'y a donc rien d'étonnant à ce que les théories acceptées soient, en comparaison, plus adéquates). C'est pour cette raison que l'argumentaire réaliste insiste sur les théories matures, dont le succès prédictif est particulièrement éclatant. Par exemple, la physique des particules postule un zoo d'entités fondamentales aux propriétés diverses et prédit la magnitude de certaines quantités physiques avec une précision remarquable. L'argumentaire réaliste met aussi en avant le caractère inattendu des prédictions des meilleures théories. Une théorie dite mature ne jouit pas simplement d'une large base empirique, elle a aussi connu des succès dans des prédictions qui sortaient de son cadre d'application initiale. Le premier exemple qui me vient à l'esprit est l'existence des trous noirs ou des ondes gravitationnelles dans la relativité générale. L'existence de solutions mathématiques correspondant à des objets astrophysiques ou à des phénomènes inconnus jusqu'alors semble, lorsque leur découverte est avérée (parfois longtemps après l'énonciation de la théorie), déborder du cadre purement empiriste de la construction de la science. Il semble que ce genre de succès reste inexplicable, si ce n'est par quelque chose qui tient du miracle : parmi toutes les théories que l'on pouvait formuler afin de sauver les phénomènes, nous avons trouvé l'une de celles⁷ qui non seulement réussissait dans cette tâche, mais dont le succès se prolongeait au-delà des espérances initiales.

On voit alors l'argument du NMA se dessiner : il n'est pas nécessaire

7. Il est bien évidemment impossible de quantifier le nombre total de théories imaginables ainsi que celui des théories qui auraient joui d'un succès au moins équivalent. S'il peut s'agir d'un point d'entrée pour une contre argumentation, je ne pense pas que le NMA repose sur ce point précis : même si une infinité de théories pouvaient avoir un succès supérieur ou égal, de nombreuses autres en auraient un bien inférieur, et il faudrait alors expliquer comment nous nous trouvons dans un ensemble plutôt que dans l'autre. Notons aussi que l'argument réaliste ne vise pas à prouver la vérité stricte de la théorie, mais plutôt à établir un lien entre ses vertus épistémiques et empiriques et sa vérité probable.

de croire à ce miracle si on adopte le point de vue réaliste selon lequel ce succès prédictif est le signe que la théorie capturait bien certaines propriétés réelles de certaines entités réelles. Si la relativité générale a eu ce succès inattendu, c'est parce qu'elle dit quelque chose d'approximativement vrai sur la nature de l'espace-temps. Si le modèle standard de la physique des particules prédit avec une si haute précision la masse de telle ou telle particule encore inconnue, c'est parce que ces particules existent et ont des propriétés proches de celles que leur attribue la théorie.

Les réalistes ont alors la conclusion de leur argument à portée de main : deux explications au succès des sciences sont disponibles, et l'une d'entre elles fait appel à un miracle. L'inférence à la meilleure explication (ou du moins à la plus probable) guide donc vers l'autre.

La méta-induction pessimiste et le gambit historique

L'argument le plus connu en faveur de l'antiréalisme (ou du moins, celui qui est souvent considéré comme l'arme la plus puissante de l'arsenal antiréaliste, ce qui en fait le pendant du NMA) est la méta-induction pessimiste (Pessimistic Meta-Induction (PMI)).

La forme générale de l'argument est celle d'une induction qui prend pour observations de base les échecs et abandons de nombreuses théories scientifiques passées qui jouissaient pourtant de toutes les mêmes vertus épistémiques que nous accordons aux meilleures théories actuelles. Ces théories largement acceptées en leur temps et aujourd'hui obsolètes montrent par induction que le contenu de la science actuelle risque fort de subir le même sort.

C'est une idée présente chez beaucoup d'auteurs, mais les illustra-

tions les plus connues sont probablement celles de Poincaré (Poincaré 1905, chap. X) :

Les gens du monde sont frappés de voir combien les théories scientifiques sont éphémères. Après quelques années de prospérité, ils les voient successivement abandonnées ; ils voient les ruines s'accumuler sur les ruines ; ils prévoient que les théories aujourd'hui à la mode devront succomber à leur tour à bref délai et ils en concluent qu'elles sont absolument vaines. C'est ce qu'ils appellent la faillite de la science.

ainsi que le *gambit historique*⁸ de Laudan (Laudan 1984) :

I challenged the intuitions which motivate the realist enterprise by arguing (among other things) that many (now discredited) scientific theories of earlier eras exhibited an impressive sort of empirical support, arguably no different in kind from that enjoyed by many contemporary physical theories. Yet we now believe that many of those earlier theories profoundly mischaracterized the way the world really is. More specifically, we now believe that there is nothing in the world which even approximately answers to the central explanatory entities postulated by a great many successful theories of the past. My approach in that essay, which we might call *the historical gambit*, was to show that these historical cases call into question the realist's warrant for assuming that today's theories, including even those which have passed an impressive array of tests, can thereby warrantedly

8. Aux échecs, un gambit est un sacrifice d'un pion ou d'une pièce qui offre en compensation un avantage stratégique ou positionnel. Il s'agit ici d'admettre un contre-argument afin d'obtenir, j'imagine, une position philosophique plus forte et mieux argumentée.

be taken to be (in Sellars' apt image), "cutting the world at its joints".

On peut également trouver à l'appui de cet argument une liste d'exemples de telles théories abandonnées dans un autre article de Laudan (Laudan 1981).

De la même manière que le NMA pouvait s'appliquer à plusieurs aspects des théories scientifiques, la PMI, dépendant de sa formulation exacte, peut cibler différents secteurs des prétentions épistémiques de l'interprétation réaliste du succès scientifique. C'est en cela que je qualifiais ces deux arguments de structures argumentatives.

Les deux citations précédentes de Poincaré et Laudan en sont un exemple. Poincaré n'y précise pas⁹ ce qui constitue le succès et les raisons de l'acceptation des théories, ni ce qui résulte de cette faillite de la science. Laudan, quant à lui, concentre sa critique sur les entités explicatives (*explanatory postulated entities*). On peut trouver un exemple de la variété d'usages de la PMI chez Wray (Wray 2019, chap. 5), qui distingue quatre formes d'inductions pessimistes différentes (dont une qui est celle de Poincaré).

Je pense toutefois qu'il est plus clair d'identifier deux cibles principales des arguments inductifs, qui correspondent peu ou prou aux deux thèses réalistes (Sém) et (Épi). D'une part, on peut utiliser (comme Laudan) l'argument pour montrer que les entités décrites par les théories actuelles ne peuvent être considérées comme réelles, étant donné que de nombreuses entités décrites par des théories aujourd'hui dépassées ne font plus partie de notre zoo ontologique (soit parce que les propriétés qui leur sont attribuées ont changé, soit parce que leur existence a été

9. On peut tout de même trouver dans la suite du chapitre X un développement qui est interprété par Worrall comme un réalisme structural. J'y reviendrai dans la section 1.6.

entièrement rejetée). D'autre part, on peut utiliser l'argument pour affaiblir le lien entre le succès empirique d'une théorie et sa capacité à être partiellement vraie, puisque l'histoire est remplie de théories empiriquement adéquates et pourtant fausses. Dans le premier cas, l'argument touche la possibilité d'interpréter légitimement les théories de manière littérale, et dans le second, c'est l'optimisme épistémique qui est plus directement visé.

1.6 Réalisme structural et la stratégie du meilleur des deux mondes

Cette section est consacrée à la question du réalisme structural, réputé pour être la forme la plus résistante de réalisme. Je ne prétends pas du tout en proposer ici une analyse, pas plus que les sections précédentes n'avaient pour but de décortiquer en profondeur les arguments du débat, mais il est nécessaire que je mette en évidence certains éléments qui me seront utiles dans la suite. Encore une fois, je souhaite mettre en exergue les éléments qui seront pertinents pour comprendre le cadre de pensée que je me propose de construire.

Les sources concernant le réalisme structural sont nombreuses, le lecteur intéressé peut par exemple se diriger vers (Worrall 1989; Ladyman 1998; Frigg et Votsis 2011) et les références qui s'y trouvent pour un panorama du sujet, ou vers (French 2017) pour une discussion plus longue, récente et complète.

Si le réalisme structural a la réputation de constituer la position la plus résistante à l'antiréalisme, c'est par construction : la proposition de John Worrall en 1989 est développée en prenant en compte les arguments antiréalistes, principalement la PMI, afin d'en concevoir une

défense appropriée.

Les inductions pessimistes, l'insistance sur les révolutions scientifiques, sur les renversements théoriques radicaux ou les changements de paradigmes tendent à mettre en lumière des discontinuités dans l'histoire des sciences. Ces discontinuités affaiblissent toute notion de progrès, particulièrement de progrès vers la vérité, qui est au cœur de la position réaliste. La proposition structurale de Worrall vise justement à rétablir une forme de continuité, sans pour autant nier l'importance des abandons conceptuels et théoriques que la PMI met en évidence et prend comme base argumentative.

Psillos, dans sa défense d'une position réaliste (non structurale, en ce qui le concerne) montre l'intérêt, pour ne pas dire la nécessité, d'adopter une stratégie de type *divide et impera* : accepter les arguments de la discontinuité doit amener le réalisme à préciser ses prétentions épistémiques. Si tout le produit de la science ne peut être l'objet d'une progression continue, cela n'entraîne pas que tout le contenu théorique est menacé de changement et d'abandon. Le réaliste se propose alors de déboucher la continuité là où elle se cache dans le changement.

La proposition de John Worrall, qui marque le début du réalisme structural, se développe exactement dans cette optique. Entre l'optimisme réaliste porté par le NMA et la critique de la PMI, Worrall trouve une voie intermédiaire : le meilleur des deux mondes (*best of both worlds*). Il situe lui-même l'origine de sa position chez Poincaré. En effet, juste à la suite du passage cité plus haut à propos de la faillite de la science, Poincaré insiste sur le fait que cette impression de faillite repose sur une mauvaise compréhension du but des théories scientifiques. En prenant l'exemple de la théorie optique de Fresnel et de son abandon en faveur de l'électromagnétisme de Maxwell, il montre que les prédictions des

phénomènes optiques restent tout à fait valables, même si l'idée d'éther chez Fresnel n'est plus présente chez Maxwell. Si on considère que le but de Fresnel était de décrire l'éther, alors sa théorie était fautive, dans le sens où l'éther en tant qu'entité n'existe pas. Par contre, les équations, elles, restent vraies.

Poincaré ajoute :

Et qu'on ne dise pas que nous réduisons ainsi les théories physiques au rôle de simples recettes pratiques ; ces équations expriment des rapports et, si les équations restent vraies, c'est que ces rapports conservent leur réalité. Elles nous apprennent, après comme avant, qu'il y a tel rapport entre quelque chose et quelque autre chose ; seulement, ce quelque chose nous l'appelions autrefois mouvement, nous l'appelons maintenant courant électrique. Mais ces appellations n'étaient que des images substituées aux objets réels que la nature nous cachera éternellement. Les rapports véritables entre ces objets réels sont la seule réalité que nous puissions atteindre, et la seule condition, c'est qu'il y ait les mêmes rapports entre ces objets qu'entre les images que nous sommes forcés de mettre à leur place. Si ces rapports nous sont connus, qu'importe si nous jugeons commode de remplacer une image par une autre. (Poincaré 1905, chap. X)

Le cœur de la stratégie *divide et impera* du réalisme structural se trouve là : malgré l'abandon successif de concepts et de théories, une forme de continuité peut être trouvée, mais elle se trouve dans les « rapports véritables » et non pas dans les entités décrites par les théories. Si nous utilisons des images et pensons la science en termes d'objets et de

propriétés, il ne faut pas perdre de vue que ces images seront abandonnées et remplacées.

Il y a donc discontinuité et continuité qui cohabitent dans l'évolution du contenu des sciences. Le réaliste peut donc utiliser la continuité de la description de ces rapports véritables (qu'ils soient exprimés sous la forme d'équations ou non) pour fonder sa position, et c'est ce que propose de faire Worrall.

Le caractère faillible d'une partie du contenu des sciences est une concession faite à la PMI, mais qui n'entame en rien l'ambition de la construction scientifique :

Que tel phénomène périodique (une oscillation électrique, par exemple) soit réellement dû à la vibration de tel atome qui, se comportant comme un pendule, se déplace véritablement dans tel ou tel sens, voilà ce qui n'est ni certain ni intéressant. Mais qu'il y ait entre l'oscillation électrique, le mouvement du pendule et tous les phénomènes périodiques une parenté intime qui correspond à une réalité profonde ; que cette parenté, cette similitude, ou plutôt ce parallélisme se poursuive dans le détail ; qu'elle soit une conséquence de principes plus généraux, celui de l'énergie et celui de la moindre action ; voilà ce que nous pouvons affirmer ; voilà la vérité qui restera toujours la même sous tous les costumes dont nous pourrions juger utile de l'affubler. (Poincaré 1905, chap. X)

On voit que Poincaré anticipe une critique instrumentaliste d'ordre fictionnaliste : ces images que nous substituons aux objets véritables et inaccessibles n'ont au mieux qu'une valeur heuristique ou explicative (il

nous est difficile, peut-être impossible, de parler et raisonner sans faire référence à des objets), mais ils ne constituent qu'une partie de la science. La connaissance est portée par des principes qui subsistent sous différents costumes. Ce détail prendra une importance centrale dans la suite de mon développement, car le concept de fiction que j'utiliserai est bien trop souvent associé, de manière abusive, au fictionnalisme, et je montrerai qu'une division appropriée entre la construction de la connaissance et le contenu de la connaissance est primordiale pour éviter les confusions.

Le réalisme structural de Worrall a subi plusieurs lignes de critique, notamment sur des points formels, qui ont mené à une révision qui a pris le nom de réalisme structural ontologique (*Ontological Structural Realism*). Je ne détaillerai pas ici l'objet de ses critiques ainsi que la réponse apportée par la version ontologique (voir, par exemple, (Frigg et Votsis 2011) pour un panorama du débat, ainsi que les autres travaux cités en début de section). L'important pour la suite de mon argumentation se situe dans la possibilité de l'élaboration d'une stratégie de division entre objets (ou images, pour reprendre le terme de Poincaré) et structure. La difficulté à définir précisément cette dernière est d'ailleurs une des critiques récurrentes les plus fortes à l'encontre du réalisme structural. Le cadre d'analyse dont le développement constitue le point central de cette thèse permettra d'éclaircir cette problématique.

1.7 Le réalisme dans cette thèse

Le but de cette section était de présenter le réalisme scientifique et d'en discuter les aspects qui seront importants pour la suite. En particulier, il était nécessaire d'éclaircir les conditions minimales pour qu'une

conception des sciences puisse légitimement être qualifiée de réaliste. En effet, je me propose de défendre une compréhension de la notion de modèle qui permette à la fois d'exprimer les problèmes classiques de l'épistémologie (dont l'opposition réalisme/antiréalisme) et d'offrir les ressources pour une défense réaliste.

Il me reste donc deux choses à faire dans ce chapitre : d'une part justifier l'utilisation de la notion de modèle dans ce contexte, en montrant la pertinence et l'intérêt ; et d'autre part, présenter certaines des ressources théoriques auxquelles je ferai appel. Ces dernières ont récemment été développées dans le cadre du problème de la représentation, mais je montrerai en quoi elles peuvent être également utiles pour le réalisme. Cela m'amènera d'ailleurs à regrouper certaines discussions à la question plus générale du réalisme tel que défini plus haut.

2 Quelques mots à propos des modèles

Dans cette section, je motive mon choix des modèles en tant qu'unité épistémique d'analyse. Ce choix pourra sembler évident tant la notion de modèle scientifique est réputée constituer le sujet de prédilection de nombreux épistémologues depuis au moins la moitié du XX^e siècle. Ce choix entraîne cependant des conséquences qu'il me semble tout de même nécessaire d'explicitier. Ce sera donc l'objet de cette section.

Je pense que l'on peut identifier trois raisons principales à l'intérêt croissant pour les modèles scientifiques.

Premièrement, on peut voir l'histoire de la philosophie des sciences du XX^e siècle comme une transition du point focal de l'analyse des théories aux modèles. Sans entrer dans les détails, la notion de théorie était primordiale pour les empiristes logiques, avant que les débats voient

le concept de modèle prendre de l'importance à partir de l'émergence de la *semantic view of theories*, avec par exemple Suppe (Suppe 1977; Suppe 1989). Si les empiristes logiques voyaient les modèles comme des produits temporaires qu'une philosophie rigoureuse de la science devait éliminer de son analyse, les tenants de la conception sémantique mettent au contraire les modèles au centre de leur analyse de la pratique scientifique. L'opposition entre ce qui a été appelé la *received view* et la *semantic view* a structuré une grande partie des débats épistémologiques de cette époque, mais avec une définition de modèle différente de celle qui nous vient en tête actuellement. En effet, le slogan de la conception sémantique selon lequel « une théorie est une famille de modèles » (plutôt qu'un ensemble de propositions dans les approches syntaxiques à la Carnap) prend son sens dans le contexte des modèles logiques au sens de Tarski, c'est-à-dire vus comme des structures d'interprétation d'un langage formel (Suppes 1960).

Deuxièmement, l'importance grandissante en philosophie des sciences des approches historiques et sociales a entraîné une attention particulière aux conditions dans lesquelles le savoir scientifique se construit, plutôt que sur la tentative de capturer ce savoir dans un formalisme donné, qu'il soit purement logique ou ensembliste, par exemple. Dans ce cadre, dès lors que l'on constate l'utilisation fréquente de la notion de modèle, il devient nécessaire de comprendre leur influence sur la méthodologie scientifique. Les scientifiques utilisent des modèles, il convient donc d'analyser cette pratique.

Troisièmement, l'intérêt pour les idéalizations. La reconnaissance de leur omniprésence en sciences a conduit à une analyse en termes de modèles. Dans ce cadre, un modèle est un certain type d'objet qui, selon les approches et les auteurs, figure dans nos explications ou constitue l'ex-

plication même. Un modèle peut être alors vu comme une certaine exemplification de propriétés d'objets idéalisés (comme le modèle du pendule simple, par exemple, qui décrit une situation idéale où les frottements de l'air sont inexistantes et où seule la gravitation exerce une force).

Au vu de tout cela, il pourrait sembler qu'il y ait une polysémie du terme « modèle » qui soit extrêmement trompeuse. Entre le modèle entendu au sens de Tarski, le modèle utilisé par les physiciens pour décrire des situations idéalisées et le modèle matériel (tel un modèle réduit d'une construction architecturale), aucun point commun n'apparaît clairement. La section suivante présentera une définition qui vise à capturer ce qui justifie l'utilisation a priori si large du terme, mais il est utile d'en dire déjà quelques mots.

Patrick Suppes commence son article *A comparison of the meaning and uses of models in mathematics and the empirical sciences* (Suppes 1960) par une liste de citations qui font toutes appel à la notion de modèle. Certains de ces exemples sont tirés de textes mathématiques, d'autres de sciences empiriques. S'il remarque immédiatement des usages différents du terme, il argumente en revanche que la définition est commune :

I would assert that the meaning of the concept of model is the same in mathematics and the empirical sciences. The difference to be found in these disciplines is to be found in their use of the concept. (Suppes 1960, p. 289)

À propos des modèles de données, entendus comme la tentative de réduire des données d'expérience à une entité plus simple, il ajoute :

The maddeningly diverse and complex experience which constitutes an experiment is not the entity which is directly

compared with a model of a theory. Drastic assumptions of all sorts are made in reducing the experimental experience, as I shall term it, to a simple entity ready for comparison with a model of the theory. (Suppes 1960, p. 297)

Les modèles que les scientifiques tirent des données seraient donc à comprendre comme des simplifications appliquées aux données (on peut par exemple penser à la recherche d'une fonction minimisant l'écart avec les points obtenus par une mesure) afin de les comparer plus facilement aux modèles de la théorie au sens logique du terme. Ce parallèle fondamental prendra tout son sens dans la suite de cette thèse.

2.1 Double diversité des modèles

Je prends cette définition large et intuitive de modèle comme point de départ afin de la préciser dans la suite, en fonction des contraintes que la défense d'une position réaliste posera dessus. Ce n'est pas un problème de considérer une notion si vague de modèle, car la plupart des textes sur le sujet commencent par établir la difficulté intrinsèque qu'il y a à cerner ce concept. Pour l'instant, je me contenterai donc de l'idée générale.¹⁰

La première difficulté qui est mise en avant quand on aborde la notion de modèle, et qui s'oppose à l'établissement d'une définition plus précise, c'est celle de la diversité. Cette diversité est en réalité double. Dans la pratique scientifique, de nombreux objets sont appelés modèles : des objets concrets (modèles réduits — *scale models* —, schémas, dessins,

10. Un mot tout de même à propos de ma démarche et de la structure de cette thèse : défendre une position réaliste en partant d'une analyse des modèles ne requiert a priori pas de redéfinir la notion, ce qui pourrait entraîner un risque d'adapter ladite définition au problème qui se pose. Ma démarche fonctionne dans l'autre sens : je pars d'une conception intuitive qui correspond à une utilisation répandue, puis au cours du développement, la construction du cadre de pensée que je défends posera des conditions sur la construction d'une définition adéquate.

organismes modèles (Ankeny et Leonelli 2011)) et des objets abstraits (ensembles d'équations supposés décrire une classe de situations données, hypothèses simplificatrices, expériences de pensée, etc.). La question d'une définition capable de recouvrir ces différents objets se pose. La seconde diversité est celle des usages, les modèles scientifiques sont supposés remplir de nombreuses fonctions : représenter des systèmes physiques, expliquer des phénomènes, aider à la construction des théories, permettre le développement de nouvelles techniques de résolution, etc. Là encore, il n'est pas certain qu'une définition générale soit possible.

Il serait facile à partir de là d'affirmer le caractère indépassable d'une telle variété d'approches, et de se contenter d'une liste ou d'une catégorisation plus ou moins exhaustive. C'est à la fois une facilité qu'il est nécessaire de dépasser ainsi qu'une limitation qu'il faut admettre. Je veux dire par là qu'une fois un usage spécifié, il doit être possible d'analyser son fonctionnement sans référence explicite aux objets particuliers que les scientifiques créent ou utilisent pour arriver à ces fins.

Dans la section suivante, je présenterai l'analyse entamée par Roman Frigg et poursuivie avec James Nguyen dans le cadre du problème de la représentation. J'utiliserai une partie de leurs résultats qui me semble utile pour traiter du problème qui m'occupe (c'est-à-dire celui du réalisme), ce qui suscitera une nouvelle conception des modèles que j'analyserai ensuite. Cette conception des modèles, qui fera appel à la notion de fiction (voir Chapitre 3), sera à la fois plus générale et plus adaptée à traiter du problème du réalisme.

3 Le problème de la représentation

Parmi toutes les fonctions associées aux modèles scientifiques, la représentation est celle qui soulève les problèmes épistémologiques les plus évidents. Si l'on pose la question en ces termes : « quelles sont les conditions qui font qu'un modèle représente un système physique ? » et « qu'est-ce qui différencie une bonne représentation d'une mauvaise ? », il semble même que la question tende à se confondre avec celle du réalisme. Quelque chose comme une notion de *fidélité de la représentation* semble être au travail et trouve un rôle analogue aux questions à propos de la vérité que le réalisme met en avant.

Est-ce que pour autant le problème de la représentation se conçoit comme la version appliquée aux modèles du réalisme traditionnellement exprimé en termes de théories ? Dans une très large mesure, l'état actuel du débat montre que non. Bien au contraire, la plupart des auteurs qui s'intéressent à la représentation par les modèles évitent soigneusement le champ de bataille réaliste, réputé pour être miné avec une densité rarement égalée en philosophie. Mon but dans cette section est d'éclaircir ce parallèle, mais je pense qu'il est clair au vu des discussions qui précèdent que je souhaite l'exploiter plus largement qu'il ne l'a été jusqu'ici dans la littérature.

3.1 Fonction de la représentation et intentionnalité

Le problème de la représentation peut en première analyse se concevoir comme un problème relationnel : a représente b (a est une représentation de b) en vertu d'une relation R telle que aRb . Dans ce cas, on appelle b la cible (ou système cible) de la représentation.

Dans le contexte scientifique, s'il est admis que les modèles ont notamment pour fonction de représenter certains phénomènes ou systèmes, le problème philosophique qui se pose est de comprendre la relation qui relie le modèle à sa cible, et ce d'une manière qui rende compte de – ou soit au moins compatible avec – la pratique scientifique (comment le modèle est créé/utilisé pour arriver à cette fin) et qui explique leur place dans la connaissance scientifique (relation avec les théories, rôle explicatif, avantages épistémiques).

Il y a cependant une ambiguïté immédiate à propos de la fonction représentative : la représentation est certainement un aspect de l'utilité des modèles, mais est-ce pour autant un usage des modèles ? Plus précisément, la représentation est-elle un but en soi ?

Cette question émerge d'un parallèle avec la notion de représentation au sens large et a à voir avec l'intentionnalité de la démarche. Cette analogie avec les représentations picturales ou artistiques prendra toute son importance lorsque j'aborderai la notion de fiction, mais il est utile d'y faire appel dès maintenant. Une peinture figurative – par exemple un portrait – est censée représenter une personne. Il y a de toute évidence une intentionnalité centrale de la représentation : le but premier est de former une représentation du sujet (que l'on appelle généralement le modèle, facteur de confusion que j'éviterai en réservant ce terme au vecteur de la représentation et non à ce qui est représenté). Si cette représentation véhicule d'autres choses, c'est en tant que produits secondaires. En ce qui concerne les modèles scientifiques, la représentation est une étape intermédiaire, une condition nécessaire mais non suffisante de l'entreprise de modélisation. Le but final est d'en apprendre sur les systèmes cibles. Or, le but d'une peinture de Napoléon n'est pas d'en apprendre sur Napoléon.

Et pourtant, on pourrait rétorquer que si la peinture est bien exécutée, il sera possible à partir de son observation d'en inférer des vérités à propos de Napoléon. Ce que cet exemple met en lumière, c'est la nécessité d'introduire une distinction entre les conditions qui fondent une représentation et celles qui fondent la réussite représentationnelle.

Dans le cas de la peinture, l'intentionnalité apparaît clairement. En effet, la construction des modèles scientifiques suit la même règle. Dès lors, il peut être tentant de concevoir le problème de la représentation comme l'identification d'une relation qui ne se limite pas au modèle et à la cible. On trouvera de nombreux exemples de propositions du type : « Un utilisateur U utilise l'objet O afin de modéliser les propriétés P du système cible S afin de poursuivre le but B », ce qui revient à concevoir la représentation comme une relation d'ordre supérieur à deux.

La question qui se pose là est l'adaptation d'une distinction classique en philosophie des sciences : le contexte de découverte et le contexte de justification. Une proposition de relation d'ordre supérieur introduit la nécessaire notion d'intentionnalité dans la création d'une représentation, mais elle interdit par là même une analyse des critères de validité d'une représentation qui évite l'intentionnalité. Cela pose en effet un problème pour la question du réalisme, car cela introduit une notion fortement épistémique dans le débat, et donc cela biaise l'analyse, au vu de la discussion que j'ai menée dans la section 1.4.

Si l'intentionnalité semble si importante, c'est parce que la représentation est permissive. Tout peut être vecteur de représentation au sens large. Il n'y a a priori aucune raison pour que ces coups de pinceau sur une toile aient une priorité dans l'analyse. Imaginons que lors d'un repas, je tente d'expliquer à quelqu'un où se trouve mon lieu de résidence. Il a une connaissance géographique des lieux, mais ne connaît

pas les noms de rues. Je prends alors un petit pain et une fourchette et dis « Voici tel ou tel lieu reconnaissable, et voici la rue qui va dans telle direction ». Dans cet exemple, le petit pain représente un bâtiment et la fourchette une route. De la même manière, des coups de pinceau sur une toile peuvent tout aussi bien représenter Napoléon que Flint Castle, tout comme, dans un contexte où cela aurait du sens¹¹, une pomme de terre peut représenter Napoléon.

Il semble qu'à la base de toute représentation se trouve un acte de stipulation arbitraire (*stipulative fiat* (Frigg et Nguyen 2020, p. 25)). Dans le cadre des modèles scientifiques, on pourra identifier des influences méthodologiques, pratiques ou sociales qui ont conduit à l'établissement de tel modèle dans tel contexte plutôt qu'à un autre, mais ce sera là une analyse de l'ordre du contexte de découverte. Une fois la représentation établie, il est nécessaire de trouver quelles conditions justifient l'utilisation de ce modèle, ce qui est une des conditions nécessaires d'une position réaliste.

Cette discussion montre la nécessité de préciser la notion de représentation à laquelle nous avons affaire. Si les modèles scientifiques sont capables de représenter des systèmes physiques, ce n'est pas par pur acte de stipulation : même en adoptant une position instrumentaliste radicale, il restera au moins la condition d'adéquation empirique. La représentation, telle qu'elle nous occupe en philosophie des sciences, est une notion plus restrictive, mais elle ne se distingue cependant que partiellement de sa définition générale étudiée par exemple dans le champ esthétique. En effet, dans le cas de la peinture figurative, le *stipulative*

11. Il pourrait sembler que la référence à un contexte réintroduise un aspect qui puisse mener à une forme de relativisme. Il n'en est rien, et l'analyse en terme de fiction (Chapitre 3) éclaircira notamment ce point, grâce à une notion analogue à ce que l'on pourrait appeler des « règles du jeu ».

fiat n'est pas non plus suffisant.

Quelles sont la ou les conditions qui fondent ce type de représentation ?

3.2 Représentation épistémique : une voie d'unification

Frigg et Nguyen (Frigg et Nguyen 2020, chap. 1) proposent une liste de conditions d'adéquations et de problèmes qu'une théorie de la représentation doit aborder afin d'être épistémologiquement acceptable. Comme je l'ai déjà dit plus haut, le problème de la représentation me préoccupe moins pour son intérêt propre que pour les ressources qui se sont développées dans ce cadre ces dernières années et qui me semblent transposables à mon approche du réalisme par le biais des modèles scientifiques.

Dans cette section, je discuterai de la classification proposée par ces deux auteurs en mettant en avant les éléments qui me serviront dans la suite. Je ne prétends ni formuler une description fidèle de leur conception de la représentation (ce qu'ils ont appelé le *modèle DEKI*) ni en faire une critique. Le lecteur intéressé peut se tourner vers l'ouvrage cité, qui contient l'essentiel de leur position (du moins à l'heure où j'écris ces lignes).

Leur point de départ se base sur le genre de remarques que j'ai esquissées dans la section précédente. Plutôt que de chercher à comprendre la représentation scientifique par le biais d'une notion générale, il est préférable de partir d'une observation de la fonction représentative des modèles en sciences. Ce qui doit fonder la représentation, et donc guider notre analyse, c'est la reconnaissance d'une utilisation épistémique qui

confère aux modèles leur place dans la connaissance scientifique.

En s'inspirant de Contessa (Contessa 2007), Frigg et Nguyen décident de traiter le problème par le biais de la notion de représentation épistémique (*epistemic representation*), c'est-à-dire la catégorie de toutes les représentations qui permettent des « raisonnements surrogatifs » (*surrogate reasonings*).¹²

Une représentation épistémique est donc définie par une condition supplémentaire par rapport à la notion générale de représentation : un modèle scientifique est à la base du *surrogate reasoning*, c'est-à-dire qu'il permet de former des propositions qui portent sur le système cible qu'il représente. Les exemples de l'histoire des sciences abondent : le modèle Newtonien du système Solaire permet de formuler des prédictions à propos de la trajectoire et de la position des planètes, le modèle de l'atome de Bohr permet de prédire certaines caractéristiques des spectres atomiques, le modèle proie-prédateur de Lotka-Volterra permet d'approcher la dynamique des populations d'écosystèmes. Cet aspect de la modélisation en sciences est si central que toute théorie épistémologique se doit d'en rendre compte.

Mais il apparaît immédiatement que la notion de représentation épistémique déborde du cadre purement scientifique. En effet, même si ce n'est pas son but explicite, il est possible de formuler à partir d'une peinture de Napoléon des propositions à propos de Napoléon, et on sera

12. J'utiliserai principalement la terminologie anglaise, vu l'absence de traduction acceptée et le manque de clarté de mon néologisme. Soler (Soler 2013, p. 185) le traduit par « raisonnement substitutif », mais elle fait remarquer que cela dérive d'une conception des modèles comme substitut qui n'est pas neutre (et qui, dans le problème qui m'occupe, peut fortement orienter le débat). Dans l'absence d'alternative satisfaisante, je garderai le terme anglophone. Je me permettrai toutefois d'utiliser les termes de « propriétés inférentielles » ou « d'inférence vers la cible » pour qualifier cet aspect des modèles, en référence notamment à la conception de Suárez (Suárez 2004 ; Suárez 2009 ; Suárez 2010), mais dans une compréhension plus étroite : il ne s'agit pour moi que de mettre en avant la fonction épistémique des modèles.

tenté de relier la validité de ces inférences à la ressemblance entre ce qui est peint et le Napoléon de la réalité. C'est là une des pistes à suivre pour fonder la relation de représentation au travail dans la modélisation scientifique. La catégorie de représentation épistémique rassemble également les diagrammes, les photographies ou les cartes géographiques. Ces dernières constituent d'ailleurs un exemple très intéressant pour comprendre le mécanisme au cœur de ces considérations.

Il y a donc un double avantage à aborder le problème de la fonction représentative des modèles scientifiques par le biais des représentations épistémiques. Premièrement, cela évite les confusions liées au *stipulative fiat* qui apparaissent dès lors que l'on tente de proposer une définition qui s'approche du sens très général de la représentation dans le langage courant, et ce, sans réelle perte de généralité, étant donné que l'aspect des *surrogative reasonings* peut être considéré comme fondamental en sciences (se restreindre aux représentations épistémiques n'est donc pas une restriction qui nous concerne). Deuxièmement, c'est un cadre qui unifie l'une des deux diversités abordées plus haut. Parmi toutes les fonctions associées aux modèles, on considère la représentation de systèmes cibles, et parmi tous les objets qui peuvent servir à atteindre ce but, on remarque que leur point commun est leur usage en tant que moyen d'obtenir et de justifier des inférences à propos de la cible. Aussi bien un modèle mathématique qu'un modèle réduit ou qu'un dessin sont utilisés en sciences afin d'obtenir certaines informations à propos de certains systèmes. Comprendre l'obtention de ces informations, leur nature et leur place au sein de la connaissance scientifique est le principal problème épistémologique qui se pose. C'est cette problématique spécifiée en ces termes que j'appelle le problème de la représentation (épistémique).¹³

13. Parmi les critères établis par Frigg et Nguyen figure la démarcation entre le problème général

3.3 Problèmes et critères d'adéquation

Une conception de la représentation doit, dans la classification proposée par Frigg et Nguyen, faire face à cinq problèmes et se positionner par rapport à cinq conditions d'adéquation. Afin de fixer les idées, en voici un bref résumé.

Tout d'abord en ce qui concerne les problèmes :

1. **Problème de la démarcation (*representational demarcation problem*).**

Ce problème est double : d'une part il s'agit de savoir si les représentations scientifiques diffèrent des autres représentations¹⁴, et si oui, en quoi ? D'autre part, il s'agit de déterminer s'il existe plusieurs types de représentations, et si les modèles, schémas, dessins et cartes peuvent être analysés avec les mêmes outils. La réponse à ce double problème de démarcation a des implications méthodologiques importantes, étant donné que le cadre d'analyse dépend de l'approche qui sera choisie. La définition de représentation qui en découlera devra en effet permettre de défendre la fécondité d'une analyse unitaire, ou au contraire justifier les distinctions qui seront proposées.

2. **Problème de la représentation (*representation problem*).**

C'est le problème central : qu'est-ce qui fait de quelque chose une

de la représentation et le problème de la représentation épistémique. C'est bien évidemment un point important quand il s'agit de construire une théorie de la représentation, mais c'est bien moins central dans ce qui m'occupe dans cette thèse. Je me permettrai donc de laisser tomber la distinction dans la suite, maintenant que le sujet a été éclairci. Sauf mention contraire, toute mon utilisation de représentation fera référence à la représentation épistémique.

14. Une dernière fois, je me permets d'insister sur le fait qu'il s'agit ici de représentations épistémiques. La démarcation porte sur la différence entre un modèle scientifique utilisé pour en apprendre sur un système physique et la capacité d'un portrait comme celui de Napoléon à soutenir également des inférences à propos de la cible. Cela prend aussi en compte des cas a priori intermédiaires comme les schémas ou les modèles réduits.

représentation ? La forme de la réponse dépendra de celle donnée au point précédent, car s'il y a plusieurs types de représentation différentes, il y aura plusieurs critères à expliciter. Par exemple, si l'on considère que schémas et modèles mathématiques sont deux formes de représentation différentes, il y aura potentiellement des conditions différentes dans ce qui fait qu'un schéma et une équations peuvent représenter.

3. Problème du style (*problem of style*).

Quels sont les styles de représentation et comment les caractériser ? En d'autres termes : par quels moyens la représentation atteint sa cible ? C'est un problème partiellement décorrélé des précédents, car même si on adopte un point de vue unifié par rapport à la démarcation, il n'en reste pas moins que les représentations peuvent mettre à profit différentes stratégies pour arriver à leurs fins (analogies, idéalizations, etc.).

4. Problème de la précision (*problem of accuracy*).

Sous quelles conditions une représentation est-elle précise ?

5. Problème des vecteurs (*problem of carriers*).

Quels types d'objets sont vecteurs de représentations ? Comment sont-ils manipulés, et comment les scientifiques en font usage ?

Les conditions d'adéquation, quant à elles, déterminent si les réponses apportées aux problèmes sont acceptables ou non :

1. Condition de directionnalité (*directionality condition*).

La relation de représentation est directionnelle, c'est-à-dire que si un modèle M représente le système cible T , il est faux de dire également que T représente M . Le portrait de Napoléon représente

Napoléon, mais Napoléon n'est pas une représentation de son portrait. Le modèle de l'atome de Bohr représente un atome (ou un type d'atomes, ou les atomes en général), mais un atome d'hydrogène ne représente pas le modèle de Bohr. Immédiatement, cela pose des conditions sur les propriétés formelles de la relation de représentation (typiquement, on souhaite qu'elle soit non symétrique). Une réponse au problème de la représentation qui mettrait en avant une notion de représentation symétrique serait jugée inappropriée (et un avantage supplémentaire sera d'expliquer la source de cette directionnalité).

2. **Condition inférentielle (*surrogative reasoning condition*).**

Étant donné l'attention particulière portée aux représentations épistémiques, cette condition joue un rôle central. Il s'agit non seulement de décrire le mécanisme par lequel les modèles scientifiques génèrent des propositions qui portent sur les systèmes cibles, mais aussi de chercher à justifier la validité de ces inférences vers la cible. Si l'on pense au problème du réalisme, cela se rattache à la thèse épistémique, les propositions inférées grâce au modèle faisant écho aux propositions de la théorie interprétée. La frontière entre modèles et théories apparaît une fois de plus comme difficile à cerner.

3. **Condition de représentation incorrecte (*Misrepresentation condition*).**

C'est une autre des conditions qui seront centrales dans la suite. Il est nécessaire qu'une conception de la représentation fasse la distinction entre une erreur de représentation et une non représentation. L'exemple du portrait de Napoléon montrait que quelque

chose comme une notion de ressemblance garantissait la validité des inférences vers la cible. C'est parce que le portrait était bien exécuté que l'on pouvait y reconnaître Napoléon et en tirer des conclusions. Un portrait mal exécuté, quant à lui, ne permettrait pas la même qualité de justification des inférences à la cible. Par contre, on ne dira pas du portrait du Docteur Gachet qu'il est un portrait, certes très mal exécuté, de Napoléon, mais qu'il constitue tout de même une représentation de Napoléon. De plus, le stylo sur mon bureau n'est pas une représentation (même s'il est possible d'imaginer une situation où il le deviendrait). Si on pense une fois de plus au rapprochement avec le problème du réalisme, la nécessité de la distinction devient évidente : si les théories doivent pouvoir être classées selon leurs vertus épistémiques (typiquement la vérité probable), alors il faut pouvoir identifier ce qui distingue une bonne théorie d'une mauvaise théorie (les termes « bonne » et « mauvaise » faisant référence à l'accord avec ces vertus épistémiques, méthodologiques, etc.). Cette condition sera à la base de la proposition fictionnelle au cœur de cette thèse.

4. Condition de représentations sans cible (*targetless representations condition*).

De nombreuses représentations ne possèdent pas de cible. Les modèles scientifiques se révèlent parfois ne pas représenter de situation réelle. Par exemple, le modèle de l'éther luminifère ne possède pas de cible. Qu'en est-il du statut de ce modèle ? Peut-il encore être qualifié de représentation ? Une conception adéquate doit expliquer ce qui se passe dans ces cas. Cette condition est importante pour le philosophe des sciences, et on reconnaît aisément dans cet exemple les idées au cœur d'un argument antiréaliste comme

la méta-induction pessimiste. On pourrait reformuler la PMI en termes de modèles et de représentations sous la forme « de nombreux modèles que l'on prenait pour des représentations se sont révélés ne pas posséder de cible, on peut induire que les modèles actuels subiront le même sort ».

5. Condition d'applicabilité des mathématiques (*applicability of mathematics condition*).

De nombreux modèles scientifiques sont hautement mathématisés, et ils sont pourtant supposés représenter des systèmes cibles physiques. Cela remet au jour la très ancienne question de la relation entre les énoncés mathématiques et le monde physique par le biais de la question de la représentation.

Ces conditions sont pensées pour guider la construction d'une conception de la représentation qui puisse faire face aux problèmes mentionnés. La comparaison avec les préoccupations réalistes n'y sont a priori pas directement reliées, même si le parallèle est parfois évident. Il ne me semble donc pas étonnant que, sur ces points particuliers, les réponses apportées et les stratégies mises en place pour résoudre le problème de la représentation trouvent une résonance dans le cadre du débat entre le réalisme et l'antiréalisme.

Encore une fois, la représentation n'est pas un thème central de ce travail, mais une ressource que je souhaite mettre à contribution. Certaines critiques des différentes conceptions du problème me seront utiles, et c'est à elles que la section suivante est dédiée.

3.4 Différentes constructions de la relation de représentation

Il apparaît immédiatement que la justification des inférences à la cible sera reliée à une notion de ressemblance. L'exemple du portrait de Napoléon l'illustre de manière évidente, et il est certain que les modèles scientifiques, en tant que représentations épistémiques, seront à traiter de la même manière. Il ne me semble pas absurde de prétendre que si le modèle Newtonien du système solaire est efficace dans sa représentation de la trajectoire des planètes (par exemple pour prédire des éclipses), c'est en vertu de quelque chose comme une ressemblance. Tout est dans le « quelque chose comme ».

En effet, si en première analyse cette idée porte une forte intuition, il n'est pas compliqué de poser des premières contraintes sur cette ressemblance supposée entre le modèle et sa cible. Au vu des critères énoncés dans la section précédente, on remarque que la relation de ressemblance n'a pas les bonnes propriétés. Plus précisément, elle est symétrique, et ne remplit donc pas la condition de directionnalité. Si A ressemble à B , alors B ressemble à A . Si on se sert d'une telle relation pour définir la représentation, on ne peut expliquer pourquoi A serait une représentation de B sans que B ne soit une représentation de A .

Dépendant de la définition adoptée, on pourra faire de la ressemblance une relation d'équivalence ou non (c'est-à-dire une relation qui soit réflexive, symétrique et transitive). La réflexivité semble acquise : A ressemble toujours à A . C'est la transitivité de la relation qui est en jeu : si A ressemble à B et B ressemble à C , alors A ressemble à C . Avec une définition accommodante de la ressemblance, on peut admettre la transitivité. Si on remarque que A , B et C sont susceptibles d'avoir des pro-

priétés différentes qui se ressemblent, la transitivité tombe. Par exemple, une trompette ressemble à un piano parce que ce sont des instruments, et un piano ressemble à un arbre parce qu'ils sont en bois. Par contre, une trompette ne ressemble pas à un arbre. L'écueil contraire peut aussi apparaître : une définition trop libérale trivialisait l'analyse. Tout peut ressembler à n'importe quoi si on accepte une gamme assez large de propriétés. Une trompette ressemble à un arbre parce qu'ils sont faits d'atomes. Dans tous les cas, on voit qu'il faut préciser la notion de ressemblance au travail et qu'elle ne peut servir en tant que telle pour fonder la représentation.

La conception par similarité

Toutes ces considérations sont au cœur de la proposition appelée *similarity view*. La conception de la représentation par similarité répond facilement à la condition de *surrogative reasoning* : un modèle permet des inférences à la cible en vertu d'une certaine similarité, ou ressemblance.

À propos de la condition de directionnalité, plusieurs options sont possibles. Par exemple, on peut introduire des degrés de similarité (A est plus similaire à B que B n'est similaire à A), mais le problème de la trivialité n'en sera pas pour autant résolu, car un choix adapté de ce sur quoi porte la similarité permettra toujours d'obtenir la relation.

Une autre solution possible est de l'ordre de ce que j'ai énoncé plus tôt. Après avoir discuté et critiqué différentes formes de la conception par similarité, Frigg et Nguyen en arrivent à cette version :

a carrier X is an indirect epistemic representation of a target system T iff there is an agent A who uses X to represent a target system T by proposing a theoretical hypothesis H

specifying a similarity (in relevant respects and to the relevant degrees) between X and T for purpose P . (Frigg et Nguyen 2020, p. 39)

Le problème de la directionnalité est réglé, ainsi que d'autres problèmes parmi ceux énoncés plus haut dans la section 3.3, mais au prix d'une référence explicite aux agents qui utilisent le modèle. Ils attribuent ce genre de position à quelqu'un comme Ronald Giere, qui l'écrit explicitement (dans un article au titre non moins explicite : *How Models Are Used to Represent Reality*) :

It is not the model that is doing the representing; it is the scientist using the model who is doing the representing.

(Giere 2004, p. 717)

Il s'agit donc moins d'établir la représentation comme une relation que de comprendre l'action représentative chez un agent, typiquement un scientifique. On s'expose donc à une analyse qui intégrerait dès la formulation même du problème des notions épistémiques, et qui rend par conséquent une application aux questions réalistes impossible.

Le lecteur intéressé par les critiques de la conception par similarité se rapportant aux autres critères d'adéquation peut sans hésitation se diriger vers le chapitre trois de l'ouvrage de Frigg et Nguyen (Frigg et Nguyen 2020).

Quelques mots sur la distinction représentation directe et représentation indirecte

Avant de continuer à discuter des propriétés des différentes conceptions de la représentation, il est utile de dire quelques mots d'une distinction dans le débat qui se révélera utile dans la suite. Il y a deux manières

de penser la façon dont les modèles représentent leur cible : représentation directe et représentation indirecte. La différence se situe dans le mode de spécification du modèle, et cela a une conséquence immédiate sur le statut ontologique et épistémologique du modèle.

La représentation indirecte est celle qui est le plus souvent associée aux modèles scientifiques. Dans cette optique, la représentation se fait en deux étapes. Tout d'abord, le scientifique construit un modèle, et ensuite, il tire des conclusions qui, en vertu d'une relation donnée avec le système cible, permet d'en inférer des conclusions.

La situation peut se représenter comme suit :

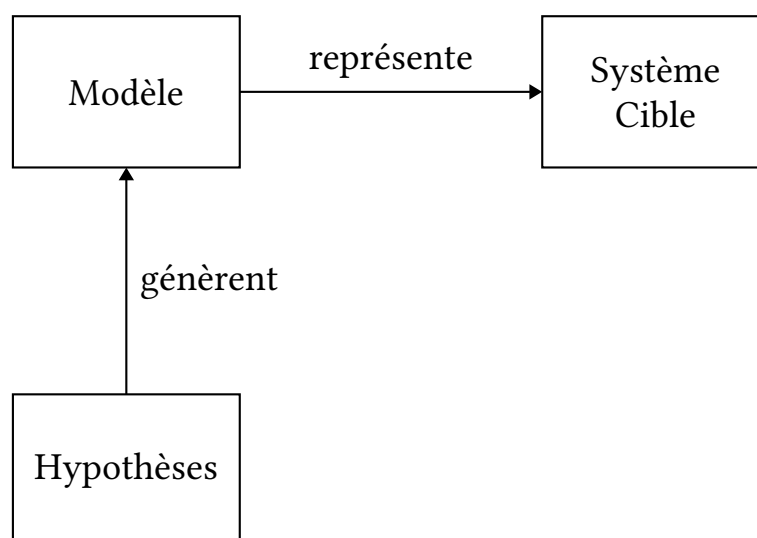


FIGURE 2.1 – Représentation indirecte

Ce point de vue reproduit assez bien les intuitions des philosophes des sciences et des scientifiques eux-mêmes face aux modèles qu'ils utilisent. Souvent, le modèle est spécifié, par exemple avec une série d'équations et d'hypothèses générales, puis il acquiert une sorte d'indépendance par rapport au phénomène étudié. Le modèle est alors étudié, certaines de ses propriétés sont découvertes, et c'est seulement dans un

second temps que la question de l'adéquation avec les phénomènes est posée. Il n'est d'ailleurs pas rare, dans la pratique, de parler des modèles comme d'objets à part entière, c'est-à-dire susceptibles d'avoir des propriétés et de se comporter d'une manière ou d'une autre. On étudiera le système solaire Newtonien ou l'atome de Bohr comme s'il s'agissait d'objets ou de situations réelles.

Ce mécanisme est aussi adapté aux modèles concrets, comme les modèles réduits. Un architecte souhaitant connaître la résistance aux courants d'air d'un pont qu'il est en train de dessiner pourra en fabriquer une maquette qu'il soumettra à des tests en soufflerie. Il étudiera les propriétés de cet objet, qu'il transférera dans un second temps au pont réel afin d'en connaître les propriétés (ce transfert impliquera bien évidemment une certaine traduction. Cette observation figure au cœur de la proposition de Frigg et Nguyen).

Le modèle sert donc dans cette optique d'intermédiaire entre la spécification, c'est-à-dire la description du modèle, et le système cible, d'où le nom de représentation indirecte.

La représentation directe, quant à elle, considère une spécification du système cible sans intermédiaire, elle se fait sans le support d'un médiateur, qu'il soit matériel ou abstrait. La situation peut se représenter comme suit :

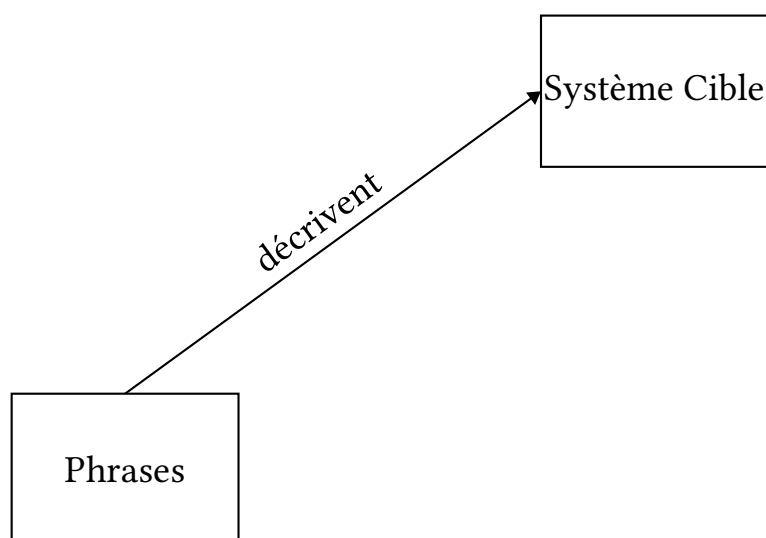


FIGURE 2.2 – Représentation directe

Les descriptions purement linguistiques d'un système entrent dans cette catégorie, mais je verrai dans le chapitre 3 qu'une conception des modèles scientifiques peut s'adapter à ce point de vue, en considérant les modèles comme des descriptions directes, mais partielles, de leur cible.

Dépendant du point de vue adopté, les problèmes ontologiques qui se posent ne sont pas les mêmes. Par exemple, si on adopte la vision indirecte de la représentation, il faudra pouvoir déterminer s'il existe une différence de représentation par des objets concrets et par des objets abstraits (comme des équations). Il faudra aussi expliquer comment un objet abstrait peut instancier des propriétés comparables (par exemple via la similarité) à des propriétés d'un système physique. Si on adopte le point de vue direct, il faudra pouvoir expliquer en quoi un modèle réduit peut être considéré comme une description d'un système à taille réelle (ce qui peut amener à proposer une démarcation supplémentaire dans les types de représentations).

Toutes ces questions réapparaîtront quand j'aborderai la question,

centrale dans mon travail, de la fiction, et ne seront pas sans rappeler les problèmes liés au statut des entités imaginaires à propos desquelles nous formulons tout de même des propositions qui semblent susceptibles d'être vraies ou fausses. Mais chaque chose en son temps, ce sera l'objet du chapitre suivant.

Approche structuraliste

Une autre approche de la représentation qui est très débattue est la position structuraliste. Elle s'est développée dans le cadre de la conception sémantique des théories, qui mettent en avant l'importance des modèles dans la compréhension de la structure des théories scientifiques. Beaucoup des défenseurs de cette approche considèrent les modèles comme des structures, une idée proche de la définition logique, comme je l'ai déjà mentionné. Il n'est pas étonnant que confrontés au problème de la représentation, certains tenants de la conception sémantique aient exporté les concepts fondamentaux à ce nouveau contexte.

L'idée centrale est que les modèles sont des structures qui représentent leur cible en vertu d'un mapping entre le modèle et sa cible. C'est ce mapping qui fonde la relation de représentation.

Avant de plonger dans les détails, il est important de faire le point sur quelques définitions. Par *mapping*, j'entends une application quelconque, dont les propriétés peuvent varier en fonction du cas considéré. Une structure, à première vue, est un objet mathématique. Plus précisément, c'est un couple ordonné de la forme $\langle \mathcal{A}, R \rangle$ où \mathcal{A} est un ensemble d'objets et R est une relation définie sur \mathcal{A} (ou un ensemble de relations définies sur \mathcal{A}). Par exemple, $\langle \mathbb{R}, \geq \rangle$ est la structure des nombres réels munis de la relation *plus grand ou égal*.

Les différents éléments de R peuvent être d'ordre différents. Par exemple, une relation binaire R_1 de R s'écrira xR_1y ou $R_1(x, y)$, pour $x, y \in R$. Une relation \mathcal{R} d'ordre n avec pour domaine les $x_1, x_2, \dots, x_n \in R$ s'écrira $\mathcal{R}(x_1, x_2, \dots, x_n)$.

Un mapping est une application d'une structure dans une autre, mais les cas qui nous intéresseront le plus sont ceux qui préservent ladite structure de manière isomorphique.

Soient $\langle \mathcal{A}, R_a \rangle$ et $\langle \mathcal{B}, R_b \rangle$ deux structures. Soit f une application de \mathcal{A} dans \mathcal{B} . On dit que les deux structures sont isomorphes si et seulement si :

1. f est bijective (c'est-à-dire que c'est une correspondance un-un entre les éléments de \mathcal{A} et \mathcal{B})
2. f préserve le système de relations.

L'application bijective f préserve les relations dans le sens suivant : les éléments a_1, a_2, \dots, a_n de \mathcal{A} sont dans la relation $r_a \in R_a$ si et seulement si les éléments $b_1 = f(a_1), b_2 = f(a_2), \dots, b_n = f(a_n)$ sont dans la relation $r_b \in R_b$ qui correspond à r_a .

La condition d'inférence à la cible se fait via la préservation de la structure : l'existence d'une application préservant la structure entre le modèle et sa cible permet de transférer les conclusions tirées du modèle à la cible. Le *surrogative reasoning* est donc garanti par un mapping isomorphe.

Un exemple est certainement utile pour éclaircir cette définition.

Soit une carte d'un réseau de métro composée de points colorés reliés par des segments de droites. Un usager sait comment se servir de ce type de plan : les points colorés représentent des stations et les segments représentent les lignes du réseau. Pour se rendre d'un point A

à un point B, il faut repérer les points correspondant à ces stations et trouver un trajet qui les relie en suivant les segments de droites (j'oublie volontairement les conditions d'optimisation, telles que trouver le chemin le plus court ou avec le minimum de correspondances, cela n'est bien évidemment pas important dans le cas qui m'occupe).

Comment expliquer que le plan est une représentation du réseau de métro dans l'approche structurale ? Il faut d'abord identifier les structures et ensuite établir l'existence d'un isomorphisme. La structure du plan est un couple $\langle \mathcal{A}, R \rangle$ où \mathcal{A} contient les points colorés et R des segments de droite (que l'on interprète relationnellement sous la forme : aRb pour $a, b \in \mathcal{A}$ en tant que *les points a et b sont reliés par un segment qui est un membre de R*). Le réseau du métro, quant à lui, est composé d'un ensemble de stations \mathcal{B} et de segments de rails qui les relient, disons dans R' . L'isomorphisme de \mathcal{A} dans \mathcal{B} le plus évident associe à chaque station un point coloré et à chaque segment de rails un segment de droite qui relie les points associés aux deux stations reliés par ce segment de droite.

Cet exemple simple, et presque trivial, permet pourtant de mettre en lumière les avantages et les inconvénients de l'approche structurale de la représentation. Dans le cas du réseau de métro, l'existence d'un isomorphisme permet, quand on regarde le plan, de tirer des inférences qui portent sur le réseau physique, et donc de s'y retrouver et d'y naviguer sans encombre. C'est donc bien une représentation épistémique.

Mais cette approche souffre du même problème que celle par similitude : l'isomorphisme est une application symétrique, réflexive et transitive. Sans rien spécifier de plus, si on fonde la représentation uniquement sur l'existence d'une telle application, il faut admettre que non seulement A représente B , mais aussi que A représente A , B représente

B , B représente A et, plus généralement, A et B représentent tout ce qui instancierait une structure isomorphe à A ou B .

Différentes manières de réintroduire la directionnalité de la représentation sont disponibles. En effet, il est possible d'imposer des restrictions plus ou moins fortes sur l'application afin d'en modifier les propriétés logiques, de manière à mieux coller à la notion intuitive de représentation tout en gardant l'avantage d'une formalisation de la définition. Certains auteurs ont proposé des homomorphismes ou des plongements (*embeddings*), mais c'est la notion d'isomorphisme partiel (par exemple dans (French et Ladyman 2003)) qui sera la plus utile pour ce que je souhaite développer.

La définition d'isomorphisme partiel est une modification minimale de celle d'isomorphisme, mais elle porte en elle à la fois une solution à la directionnalité, mais aussi des ressources formelles qui prendront leur sens avec la notion de fiction.

L'idée est de partitionner l'ensemble R des relations d'une structure selon leur définition extensive. Une relation peut être ramenée à la liste des objets auxquels elle s'applique : une relation d'ordre n est un ensemble de n -uplets (c'est-à-dire une liste ordonnée de n éléments de \mathcal{A} qui entrent dans ladite relation).

Dès lors, chaque relation induit une partition de tous les n -uplets construits sur \mathcal{A} : celle des n -uplets qui appartiennent à la relation, et ceux qui n'y appartiennent pas. Dans l'exemple du métro, cela revient, pour chaque segment de droite, à considérer un ensemble de deux éléments (les deux éléments qui sont reliés par ce segment) et un ensemble qui contient tous les autres couples de points colorés (en admettant que chaque segment relie uniquement deux points, mais on voit aisément la généralisation aux relations d'ordre $n \geq 2$).

Pour construire un isomorphisme partiel, on considère non pas ces deux options mutuellement exclusives (appartenir ou non à une relation), mais on introduit une troisième possibilité : l'indétermination. On appelle alors une relation partielle d'ordre n un triplet qui spécifie : les n -uplets qui appartiennent à la relation, ceux qui ne lui appartiennent pas, et ceux pour lesquels l'appartenance est indéterminée. Si l'ensemble des relations d'une structure contient au moins une relation partielle, on lui donne le nom de structure partielle.

Cette modification a l'avantage immédiat de contrecarrer un argument classique contre les approches structurales formalisées : elle rend possible l'adaptation à la pratique scientifique, dans laquelle il est courant de ne pas savoir si un certain aspect d'un modèle se retrouve effectivement dans la cible. Et c'est ce point qui fera tout l'intérêt du développement fictionnel de la prochaine section.

Avant de clore cette discussion de l'approche structurale de la représentation, il me reste à soulever une critique qui n'est pas aussi facilement contournable que celle basée sur la directionnalité.

Cette forme de structuralisme est à ranger dans la catégorie des approches naturalistes, au sens où son but est de fournir une explication de la représentation en ne faisant appel qu'à la nature du modèle et de la cible, sans référence au contexte ou aux agents qui créent et emploient le modèle. De toute évidence, cette position est souvent défendue par des auteurs rattachés au réalisme structural, même si, comme je l'ai déjà beaucoup dit, le problème de la représentation est en grande partie indépendant de cette autre problématique. Il n'est donc pas étonnant que toutes les notions épistémiques soient évacuées de l'analyse.

Pourtant, si l'exemple du réseau de métro semble fonctionner, c'est aussi parce que la structure qui émerge du réseau cible apparaît de ma-

nière évidente. Or, comme dans le cas de la similarité trop permissive (tout est, selon un aspect ou l'autre, similaire à n'importe quoi), il est faux de prétendre qu'il n'y a qu'une structure associée au système cible : un système physique exhibe plusieurs structures, dont aucune n'est a priori privilégiée. Le choix d'en utiliser une plutôt que l'autre est avant tout pragmatique.

L'existence d'un isomorphisme partiel entre un modèle abstrait et un objet physique ne garantit donc pas que l'on puisse en exhiber de nombreux autres (un nombre potentiellement infini) qui, en l'absence d'autres contraintes et conditions, seront tout autant des représentations. Le plan du réseau de métro représente le réseau réel selon l'isomorphisme que j'ai mis en évidence, mais le réseau réel instancie d'autres structures non prises en compte par le plan (par exemple les tournants ne se retrouvent pas dans les segments de droite). Le problème de l'unicité se pose aussi dans l'autre sens : le plan peut représenter par isomorphisme la disposition de certains grains de sable sur une plage de la côte Pacifique, ou des galets d'une plage méditerranéenne.¹⁵

Si un modèle mathématique est clair dans la structure qu'il instancie, le problème se complexifie davantage lorsqu'on considère les modèles matériels : de la même manière que leur cible, ils instancient de nombreuses structures, toutes susceptibles d'être isomorphes à l'une de celles que la cible montre aussi.¹⁶ Par exemple, la relation qui relie un

15. Si on pense à la représentation d'un graphe en mathématiques sous la forme de points et d'arêtes, rien n'empêche l'utilisateur d'interpréter les points comme les arêtes et vice-versa, et donc d'utiliser la carte du métro comme son dual, qui n'est, comme on sait, pas forcément isomorphe au graphe de départ.

16. Je n'entre volontairement pas dans les débats ontologiques qui touchent à ce point. Plusieurs attitudes aux diverses implications métaphysiques sont possibles : un agent peut attribuer une structure à un système physique, un système physique peut instancier une structure abstraite, un système physique peut tout simplement être une structure. Ces discussions dépassent le cadre de ce travail, l'important dans le cas présent est de montrer qu'il n'y a pas qu'une seule structure associée à un système physique, et qu'il n'y a donc pas de moyen de justifier la fondation de la relation de représentation sur

pont miniature à sa version à taille réelle semble évidente, et pourtant rien, en purs termes de structure, ne la distingue d'autres relations qui existent entre ces deux objets matériels. Leur couleur, par exemple, n'est pas pertinente si la représentation vise à en étudier la résistance par des tests en soufflerie, mais elle est pourtant bel et bien porteuse d'une relation tout aussi légitime à fonder la représentation.

Il semble qu'il faille ajouter un ingrédient à l'isomorphisme (partiel) pour préciser la définition de la représentation, mais le problème est identique à celui de la relation de ressemblance ou de similarité. Le scientifique qui utilise le modèle réduit d'un pont, tout comme les navetteurs avec le plan, savent dans quel but ils utilisent le modèle, mais cela réintroduit un élément extérieur intentionnel.

Un fameux exemple qui illustre parfaitement cette tension a été proposé par Putnam (Putnam 1981) : une fourmi marche sur le sable et laisse derrière elle une trace qui finit par former le visage de Winston Churchill. A-t-elle produit une image de Churchill ? Le dessin dans le sable représente-t-il Churchill ? Putnam répond par la négative : la fourmi n'avait pas l'intention de représenter Churchill. Quelqu'un, remarquant la ressemblance, pourrait se servir de l'image comme d'une représentation (épistémique), mais il n'y a rien dans l'image elle-même, aussi semblable soit-elle, qui en fasse une représentation.

Réduire la représentation à une ressemblance structurale n'est pas suffisant, même en admettant qu'elle soit nécessaire.

Le corollaire de cette remarque est qu'il est également impossible de

un isomorphisme particulier sans faire appel à une référence extérieure à la relation modèle-cible. Les tenants de l'approche structurale ne sont d'ailleurs pas tenus à une ontologie particulière des modèles. Au contraire, la structure peut être considérée comme une manière parmi d'autres de représenter les modèles, comme la conception sémantique des théories considère qu'une famille de modèles est une manière de représenter une théorie, sans pour autant la réduire à cela.

faire la distinction entre une non-représentation et une erreur de représentation. Si l'isomorphisme partiel existe, il y a ressemblance structurale et représentation. Si l'isomorphisme n'existe pas (en écartant d'une manière ou d'une autre les cas triviaux ou qui ne correspondent pas à l'usage que l'agent fait du modèle), alors il n'y a pas représentation. Il n'est pas possible dans ce cadre de définir ce que serait une erreur de représentation, c'est-à-dire une représentation dont les inférences à la cible seraient invalides.

Considérer que ce qui donne le privilège à une relation plutôt qu'à une autre tient à l'interprétation des termes du modèle est certainement un aspect de la réponse, mais cela reste insuffisant. D'une part, l'exemple du modèle du pont miniature montre qu'une interprétation immédiate est en quelque sorte fournie avec le modèle (ne serait-ce que dans la manière de le qualifier de « modèle réduit »), mais la provenance de cette interprétation se trouve dans des considérations méthodologiques et intentionnelles. Mettre en évidence le « bon isomorphisme » ou l'« isomorphisme adéquat » suppose la connaissance d'un but à atteindre. D'autre part, l'exemple du plan du réseau de métro semble mettre en lumière une structure qui émerge plus directement que les autres (qui, en comparaisons peuvent sembler artificielles et alambiquées), mais on pourrait répondre dans une veine wittgensteinienne que cette intuition se situe aussi dans des considérations méthodologiques et langagières : on considèrera que quelqu'un qui lit le plan en y voyant autre chose que des stations et des portions de réseau ne sait tout simplement pas comment s'en servir. Dans tous les cas, il semble qu'il soit impossible de se débarrasser entièrement de références extérieures à la relation modèle-cible.

Si l'approche structurale offre une définition de la représentation

plus précise et formelle que la conception par similarité, elle n'en souffre pas moins des mêmes problèmes. Elle possède les mêmes faiblesses à la fois dans le cadre du problème de la représentation, mais aussi dans les éventuelles récupérations réalistes que l'on peut espérer opérer.

3.5 Les trois problèmes naturalistes : réalisme, représentation et véritisme

Avant de clore ce chapitre, il me reste à mettre en place ma problématique plus précisément que je ne l'ai fait jusqu'ici. Maintenant que le réalisme et le problème de la représentation ont été abordés, je peux clarifier en quoi les outils appliqués à l'un sont applicables à l'autre. Ce sera aussi l'occasion d'aborder un troisième problème connexe que mon approche, je le montrerai dans le chapitre 5, permettra de traiter.

Ces trois problèmes, je les appelle les problèmes naturalistes : réalisme, représentation et véritisme de l'explication.

Les points communs entre les deux premiers sont nombreux, et certains sont déjà apparus entre les lignes au cours de ce chapitre, mais il me semble utile de les mettre en évidence. En ce qui concerne le véritisme, qui fait référence à une position dans le cadre de l'explication scientifique, je ne dirai ici que le nécessaire, et je garderai les détails pour le chapitre entièrement dédié à ce problème. Je pense que les parallèles seront assez clairs pour qu'il ne soit pas indispensable de les expliciter immédiatement.

L'enjeu est de comprendre ce qui constitue une explication d'un phénomène et quelles sont les critères d'évaluation de cette explication. Au-delà de leur fonction de représentation, il est assez largement admis que les modèles ont une fonction explicative, et il est donc important

de comprendre leur apport dans cette entreprise centrale en sciences. Deux questions se posent alors : « Quelle est la place des modèles dans les explications ? » et « Comment les modèles peuvent-ils expliquer ? » À la première question, trois options se proposent : les modèles peuvent figurer dans les explications, être utiles à leur construction ou constituer eux-mêmes les explications. La seconde question est le cœur de l'enjeu d'une théorie de l'explication à partir de la notion de modèle. Comme pour le réalisme et la représentation, les questions méthodologiques, historiques, sociales et faisant référence aux agents seront à la base d'arguments qui visent à critiquer une position dite *véritiste*. Je développerai cela dans le chapitre 5.

Il est commode d'aborder la question du réalisme scientifique en classant les positions dans deux grandes catégories, et il est aussi pertinent de le faire dans les autres cas. Je suis de ceux qui tendent à penser que construire des catégories éclaircit les problèmes et ne nuit pas à la reconnaissance de la diversité des points de vue. Chercher une régularité sous-jacente qui puisse participer à expliquer la variété des positions peut être considéré comme un fort biais naturaliste, mais je ne pense pas avoir à expliciter en quoi ce parti pris est cohérent avec la position défendue dans ce travail. Structurer le débat entre réalisme et antiréalisme n'a d'ailleurs jamais empêché de raffiner les définitions par la suite, comme c'est le cas avec les trois thèses que j'ai présentées en début de chapitre.

Le problème de la représentation se structure autour du même genre de tension : certaines approches insistent sur la nécessité épistémologique de comprendre la relation modèle-cible en dehors de considérations extérieures (référence au sujet, aspects historiques, sociaux, méthodologiques, notions épistémiques), celles-là je les qualifie de natura-

listes. La ligne d'argument peut se concevoir comme ceci : étant donné que nous utilisons les modèles scientifiques pour obtenir de l'information sur les systèmes cibles, il est nécessaire que l'on puisse s'assurer le plus objectivement possible de la validité de la relation de représentation, et donc de la concevoir sans référence extérieure à la nature de la cible et de ce qui la relie au modèle.

Cette dernière interprétation du problème est cependant déjà une exportation vers le réalisme. En effet, la recherche d'une adéquation entre le modèle et la cible est une démarche partiellement indépendante de celle du fondement de la représentation. Fonder la représentation, en fournir une définition qui recouvre les usages de manière satisfaisante tout en rendant compte des divers aspects de l'activité scientifique, ne nécessite pas de formulation explicite de validité ou de justification de la relation. Par contre, poser la question du réalisme en termes de modèles introduit tôt ou tard des considérations du genre de celles dont j'ai discuté dans ce chapitre.

Quand j'annonce vouloir utiliser une partie des ressources mobilisées dans le problème de la représentation pour défendre une position réaliste à partir de la notion de modèle, j'ai cependant en tête une démarche qui va plus loin que la volonté d'associer à la relation modèle-cible une interprétation en termes de vérité ou de correspondance (ce qui pourrait décrire l'essentiel de la démarche des tenants de l'approche structurale, tels que Costa et French, par exemple). Les problèmes soulevés par les tentatives représentationnelles, liés à l'intentionnalité ou aux références aux agents, sont à la fois des problèmes classiques de l'épistémologie envisagés sous l'angle des modèles (alors qu'ils sont traditionnellement orientés vers les théories et leur contexte de découverte) et des formulations particulières qui, dans ce contexte précis, suggèrent

des pistes de solution qui peuvent, c'est ma thèse, offrir un regard neuf sur la question du réalisme.

Plus précisément, ce qui est au travail en filigrane de toutes ces conceptions de la représentation (celles que j'ai abordées ici, mais aussi celles taillées pour répondre aux critiques, comme la proposition DEKI de Frigg et Nguyen dans (Frigg et Nguyen 2020)), c'est une prise en compte du caractère dynamique, voire interactionnel des modèles. Une représentation n'est une représentation que dans un jeu qui se donne pour but de représenter une cible. Un modèle n'est pas une entité monolithique, mais plutôt un système qui se construit en faisant intervenir toute une gamme de techniques qui permettent de lui conférer un caractère de représentation. La compréhension de ce mécanisme est le cœur du problème de la représentation.

Affirmer cela semble aller à l'encontre de mon parti pris réaliste et de ma tentative d'évacuer les notions épistémiques (qui mènent au relativisme tel que je l'ai défini de manière minimale plus haut) de ma conception des sciences, mais ce n'est pourtant pas le cas. Au vu des critiques, il me semble difficile de ne pas accepter l'aspect résolument pragmatique de la représentation. Mais la vision dynamique des modèles, quant à elle, est tout à fait exportable au problème du réalisme. C'est ce que je me propose de démontrer dans la suite, en faisant notamment appel à la notion de fiction.

Le problème du réalisme est souvent tenu à l'écart de celui de la représentation, et pourtant la frontière semble mince et ne tenir qu'à une poignée de considérations périphériques. La section suivante sera dédiée au développement du cadre d'analyse qui permettra d'éclaircir les différences et les points communs.

Chapitre 3

La conception fictionnelle des modèles

To what extent are these possible worlds fictions? They are like literary fiction in that they are free inventions of the human mind. In theoretical physics sometimes the inventor knows from the beginning that the work is fiction, for example when it deals with a simplified world in which space has only one or two dimensions instead of three. More often it is not known till later, when the hypothesis has proved wrong, that fiction is involved.

(Bell 1992)

Ce chapitre est dédié à la notion de fiction. Mon but est d'y aborder les principaux aspects épistémologiques de la question et de les mettre en parallèle avec les arguments évoqués dans le cadre du problème de la représentation tout au long du chapitre précédent. Je ne propose cependant pas une épistémologie de la fiction, sujet vaste et passionnant, mais plus éloigné du cœur de ce que je souhaite démontrer. La fiction sera pour moi moins un objet d'analyse qu'un cadre pour formuler les problèmes et y chercher des solutions. Je verrai d'ailleurs qu'il est plus

utile de s'exprimer en termes de processus fictionnels lorsque la question des modèles scientifiques sera abordée.

Je considère qu'il existe deux manières d'envisager la fiction : l'une de l'extérieur, et l'autre de l'intérieur. Par exemple, face à une œuvre de fiction comme un roman, il est possible de raisonner à l'intérieur du monde décrit par le texte, de s'interroger sur les relations implicites entre les personnages, par exemple. Il est aussi possible de prendre la fiction comme un objet dont on étudie les caractéristiques depuis l'extérieur, en se demandant par exemple dans quel contexte elle a été produite.

Ces deux modes d'analyse sont bien entendu complémentaires, mais ils mobilisent des ressources parfois très éloignées les unes des autres, et il est intéressant de remarquer que les disciplines philosophiques qui s'y consacrent ont des méthodes d'investigation radicalement différentes. En effet, des philosophes s'intéressant à l'esthétique, à la logique ou aux sciences cognitives sont donc tous susceptibles de prendre la fiction comme objet d'étude, chacun avec les outils propres à leur méthode de recherche.

Étudier la fiction depuis l'intérieur, c'est par exemple se demander quelles règles délimitent la fiction, ou ce qui fait qu'une œuvre de fiction (quel qu'en soit le support) réussit dans les buts qu'elle se propose d'atteindre. Dans les débats esthétiques, cette notion vise à catégoriser certaines œuvres par le rapport spécifique qu'elles ont à l'imagination et à l'acte de création.

Aborder la fiction de l'extérieur c'est par exemple questionner le statut des entités fictionnelles d'un point de vue logique : en tant qu'objets imaginés, les fictions entretiennent un rapport complexe avec la notion de vérité, rapport qui peut parfois même sembler paradoxal. Il semble intuitif d'accorder la vérité à des propositions qui portent sur des objets

qui n'existent pas, ou qui ont en tout cas un mode d'existence différent de celui des objets concrets. Dépendant du point de vue adopté, on pourra défendre que la proposition *Sherlock Holmes est un détective britannique qui vit à Londres* est vraie (Sherlock Holmes est effectivement un détective britannique qui vit au 221b Baker Street, qui se situe effectivement à Londres) ou fausse (Sherlock Holmes n'a pas de nationalité et ne vit nulle part pour la bonne raison qu'il n'existe pas. Et toutes les propositions à son sujet sont condamnées à être fausses ou du moins indéterminées, ou encore sans contenu). Ces débats s'éloignent du cadre esthétique pour s'intéresser plutôt à la logique, l'épistémologie ou l'ontologie.

Une deuxième manière d'étudier la fiction de l'extérieur est par exemple de la considérer en tant que génératrice de certains états mentaux chez le spectateur (ou le lecteur, ou l'auditeur). On se demandera alors si telle ou telle émotion suscitée par une œuvre a les mêmes caractéristiques qu'une émotion face à une situation véritable. En ressentant de la tristesse face à la mort d'un personnage sur la scène, est-on triste de la même manière qu'en apprenant le décès d'un proche ? Les états mentaux sont-ils analogues ou radicalement différents ? Cela fait écho à de très anciens débats philosophiques, mais exprimés dans des termes modernes, en faisant parfois appel aux sciences cognitives pour former des réponses potentielles.

Dans les exemples d'approches extérieures de la fiction que j'ai donnés, c'est en premier lieu son statut imaginaire qui définit la fiction, à tel point que toutes les questions externes semblent pouvoir se réduire à celles du statut des entités imaginaires. Par exemple, on débattera du mode d'existence que Sherlock Holmes doit remplir pour que les propositions qui portent sur lui aient du sens. La fiction n'est dans ce cas qu'un exemple parmi d'autres de produits de l'imagination. Son utilité

principale est que les exemples très connus regorgent et qu'ils peuvent servir d'illustrations particulièrement explicites : le monde décrit par un roman est un laboratoire privilégié pour étudier ces mécanismes, mais le concept de fiction n'est pas davantage exploité. Ce que j'ai défini comme des considérations intérieures peut en revanche s'apparenter à une délimitation plus précise du problème, puisque ce sont les fictions en tant que fictions qui sont la cible de la recherche.

L'utilisation de la notion de fiction en épistémologie se rangera naturellement dans la catégorie des approches extérieures de la fiction. En effet, j'ai montré que le cœur du problème de la représentation par les modèles scientifiques se trouvait dans la relation qui relie le modèle à un système cible. S'il s'agit de comparer les modèles à la fiction, ce qui constitue les approches fictionnelles de la modélisation, on voit que cette relation se trouve à l'extérieur. De plus, dans une perspective réaliste, c'est bien entendu la réalité extérieure au modèle que l'on souhaite décrire *in fine*. Cette distinction entre intérieur et extérieur sera la ligne rouge de mon analyse en termes de fiction : un modèle scientifique peut aussi être abordé de l'intérieur et de l'extérieur, et le problème de la représentation peut se comprendre comme la recherche d'une relation entre les deux.

Le sens intuitif de la fiction est fortement lié à l'imagination et à l'invention. C'est généralement avec cette idée en tête que le terme est utilisé, et j'ai montré que cet aspect n'est pas absent des travaux philosophiques qui usent de ce concept, même s'ils ne s'y limitent évidemment pas. En effet, réduire l'analyse fictionnelle à une simple comparaison avec des objets imaginés est une erreur. La notion de fiction est en réalité plus complexe et féconde, comme je le montrerai.

Il est cependant nécessaire de définir plus précisément la fiction pour

comprendre pourquoi. Ce chapitre commence par dégager les aspects de la notion qui mèneront à concevoir une définition de la fiction particulièrement pertinente pour traiter les problèmes d'épistémologie et de philosophie des sciences qui m'occupent. En particulier, je me pencherai sur le concept de *comme si* que l'on trouve chez Vaihinger ainsi que sur la théorie du *make-believe* proposée par Kendall Walton. Cette dernière sert de cadre théorique aux approches fictionnelles des modèles scientifiques, et je montrerai en quoi cet usage est pertinent. J'en pointerai aussi les limites, particulièrement quand il s'agira d'aborder le problème du réalisme scientifique.

Dans les sections ultérieures, je mettrai en parallèle la notion de fiction telle qu'elle aura été définie au cours des premières sections du chapitre avec le concept d'idéalisation, très largement discuté en épistémologie. Ce parallèle permettra notamment de préciser la relation qu'entretiennent modèles et théories scientifiques, ainsi que d'établir quel est le rôle des lois dans la connaissance scientifique.

1 Vers une définition de la fiction, sans pour autant y arriver

Proposer une définition de la fiction est une entreprise difficile, et qui dépend énormément du cadre dans lequel cette définition prendra place. En particulier, les buts recherchés influenceront différents aspects de l'entité nébuleuse que l'on appelle fiction dans le langage courant.

L'aspect imaginaire en est une part importante, mais elle ne s'y réduit pas. Pour s'en convaincre, il suffit de lister les œuvres qui mêlent des éléments de pure imagination avec des événements historiques. Quand

il s'agit de se lancer dans cette entreprise, la plupart des auteurs utilisent les éternels exemples qui tombent dans deux catégories : les romans classiques de la littérature russe (on fera remarquer que le Napoléon dont on parle dans le *Guerre et Paix* de Tolstoï est un personnage qui a réellement existé, et que les événements auxquels prennent part les personnages imaginaires sont bel et bien historiques) et Sherlock Holmes (on fera remarquer que la géographie du Londres qu'il habite est celle de la véritable capitale britannique). J'ai moi-même déjà trop suivi la tradition en utilisant ces deux gammes d'exemples, j'essaierai donc dès maintenant de systématiquement m'y soustraire.

La fiction ne se réduit pas à son aspect imaginaire, donc, car le réel (ou ce qui est supposé comme tel) y fait toujours irruption. Le Londres de 1944 bombardé par les V2 dans lequel Tyrone Slothrop déambule est le Londres de la seconde guerre mondiale. Il en va de même pour le New York de Holden Caulfield et de Seymour Glass.

Le problème ne disparaît pas si on se restreint à des entités fictionnelles en particulier, par exemple des personnages, plutôt qu'à des œuvres de fiction générales et qui situent leur action dans une période historique donnée. Gurion ben-Judah Maccabee, le potentiel messie potentiel des *Instructions*, a beau être un personnage totalement imaginaire, il n'en est pas moins humain, il possède un système immunitaire, huit incisives et cinq doigts à chaque main.

La relation entre fiction et fausseté, que l'on retrouve dans des expressions du genre « tout cela n'est que fiction », est donc plus complexe qu'il n'y paraît à première vue. Ce qui constitue la fiction a à voir avec l'invention de l'imagination, mais ne se réduit pas à de la pure fausseté. La fiction semble plutôt être l'objet d'une certaine attitude plus relâchée vis à vis de la vérité, d'une *suspension consentie de l'incrédulité* qui nous

fait temporairement accepter des mensonges afin d'atteindre d'autres objectifs, notamment esthétiques.

L'aspect artistique est probablement une des caractéristiques qui est le plus souvent mise en avant. L'acte d'invention serait dans ce cas là non pas une fin en soi mais un moyen d'atteindre un but, par exemple susciter telle ou telle émotion chez le spectateur. La délimitation de la fiction se ferait donc dans l'attitude que l'on adopte en créant ou en contemplant la fiction. Cette définition ne nous permettra pas d'exporter la notion aux modèles scientifiques, et elle revient à attester qu'une fiction est ce qui est le sujet d'une certaine attitude fictionnelle que l'on définira à son tour à partir de l'objet fiction. Là encore, l'aspect artistique ou esthétique n'est pas inintéressant, mais il n'est pas suffisant à capturer ce qu'il y a de pertinent dans la notion. De plus, pour mon usage, il servira de point de comparaison (ce que j'appellerai dans la suite les « œuvres de fiction »), mais il ne constituera pas une ressource théorique dans mon analyse.

Cette discussion met en évidence une tension qui reproduit assez bien l'opposition qui émergeait de mon analyse du problème de la représentation épistémique dans le chapitre précédent. D'un côté, il semble que la réponse au problème de la représentation contienne des éléments contextuels et intentionnels : une représentation est avant tout ce qui est utilisé pour représenter tout comme une fiction est avant tout ce que l'on aborde avec une attitude fictionnelle. Et de l'autre côté, les problèmes épistémologiques que je me propose de traiter nécessitent que l'on écarte ces éléments intentionnels du mécanisme épistémique : on souhaite que l'évaluation de l'adéquation d'une représentation se fasse sans charge épistémique contextuelle tout comme on souhaite dégager les éléments qui constituent la fiction sans la réduire à un jeu de convention sociales ou méthodologiques. Ce parallèle pourrait être pris comme

un signe avant coureur de la pertinence du concept de fiction dans le champ de l'épistémologie et de la philosophie des sciences.

2 Qu'est-ce qu'une utilisation de la fiction en épistémologie ?

Pour comprendre l'intérêt de la notion de fiction appliquée à l'épistémologie des modèles, il faut revenir aux problèmes théoriques qui se posent et proposer une définition en conséquence.

L'émergence récente des considérations fictionnelles à propos des modèles scientifiques part du parallèle que j'ai mentionné plus haut : il y a dans l'activité de modélisation quelque chose qui s'apparente à une création de l'imagination. Les exemples abondent : les planètes du modèle du système solaire de Newton sont des sphères parfaites avec une distribution de masse homogène, les modèles économiques considèrent des agents parfaitement rationnels et qui possèdent toute l'information disponible, les électrons de l'atome de Bohr occupent des orbites circulaires. Tous ces modèles sont comparables à des fictions dans le sens où ils semblent décrire des objets produits par l'imagination et dont l'existence n'est pas du tout assurée.

Ce n'est pas une observation nouvelle, et il ne faut généralement pas se plonger longtemps dans la littérature sur les modèles scientifiques pour rencontrer l'idée qu'ils sont des versions simplifiées de la réalité, quelle que soit la manière dont cette simplification est opérée et quelles qu'en soient ses conséquences épistémologiques. Quand il s'agit d'écrire à ce sujet, il est intéressant de voir que les comparaisons, d'une manière détournée, utilisent des termes associés à l'idée intuitive de fiction :

Models are generally caricatures of the natural world. (Chakravartty 2001, p. 327)

Michel Armatte met en évidence l'utilisation par Ludwig Boltzmann de la notion de modèle (dans l'article que ce dernier écrit pour l'édition de 1911 de l'Encyclopædia Britannica, volume 18, entrée « Model ») et précise qu'« il étend le sens de cette notion des usages techniques et artistiques ». (Armatte 2005, p. 94) Étienne Klein (Klein 2020) insiste quant à lui sur le rapprochement qu'opère Boltzmann entre les modèles et les images mobilisées en sciences, qui peuvent avoir une utilité sans pour autant prétendre constituer un reflet fidèle de la réalité. C'est notamment le cas, à cette époque, de l'hypothèse atomiste.

Au début de ses recherches, James Clerk Maxwell, pour qui la notion de champ électromagnétique n'était pas encore à disposition (et pour personne d'autre à l'époque, bien évidemment), invitait à visualiser un fluide purement imaginaire dont il se proposait d'étudier les propriétés dynamiques. De nombreux autres exemples de ce genre sont listés dans (Salis et Frigg 2016).

On pourrait aussi citer l'ouvrage fondateur de Mary Hesse, dont le titre est évocateur : *Models and Analogies in Science* (Hesse 1967). Ou encore le *Models and Metaphors* de Max Black (Black 1962).

À première vue, cette comparaison avec la fiction est donc consensuelle, elle ne fait que mettre en avant un point connu et reconnu depuis longtemps. Un fait connu, d'ailleurs, par les scientifiques eux-mêmes au moment de l'élaboration de leur modèle. Ce n'est donc pas non plus une proposition philosophique obtenue après une très longue analyse de la méthodologie scientifique. Bien entendu, ce constat n'est qu'un point de départ, et les approches fictionnelles des modèles et de la représentation poussent bien plus loin le parallèle, elles ne se contentent pas de cette

analogie.

En se servant de leur classification (que j'ai abordée dans la section 3.3 du chapitre précédent) des problèmes de la représentation épistémique et des conditions d'adéquation, Frigg et Nguyen distinguent deux conceptions que l'on peut légitimement qualifier de fictionnelles. Ces positions se différencient par le niveau d'importance qu'elles accordent à la comparaison entre les modèles et les fictions, et donc par les problèmes qu'elles résolvent en y faisant appel.

Les plus modestes de ces conceptions n'utilisent la fiction que dans l'optique de régler le problème des porteurs, c'est-à-dire pour expliquer quels types d'objets sont les modèles (voir page 46). Dans cette approche, un modèle sera rapproché d'une fiction afin de comprendre la manière dont les scientifiques utilisent ces objets ou pour poser des questions ontologiques, mais le reste des considérations représentationnelles sera laissé de côté.

Les conceptions plus ambitieuses, qui sont celles qui m'intéresseront par la suite, considèrent que la notion de fiction peut aider à bâtir une théorie de la représentation dans son ensemble, et elles se doivent donc de répondre aux cinq problèmes et se prononcer sur les conditions d'adéquation. Comme dans le chapitre précédent, je ne m'attarderai que sur les éléments qui me serviront plus tard, mais cette simple description très générale de l'usage de la fiction nécessite déjà quelques commentaires.

Tout d'abord, il est important d'avoir en tête que fiction n'implique pas nécessairement fausseté, sous peine de voir la conception de la représentation associée se dissoudre dans un instrumentalisme excessif (c'est parfois à ce genre de position qui se voit attribuer le qualificatif de « fictionnalisme », mais j'y reviendrai plus tard). Si on identifie la fiction

avec le faux, ou du moins avec une certaine indifférence pour la vérité, il sera impossible de formuler des arguments à visée réaliste tant cette hypothèse porte en elle la négation de ce qui caractérise cette position épistémologique. En effet, l'approche fictionnelle ne fait pas qu'admettre que certains éléments qui figurent dans la pratique scientifique de la modélisation sont imaginés, elle en fait le cœur de son argumentaire. Dès lors, si l'acte d'imagination réduit le modèle à une description fautive, il n'y a plus d'issue réaliste. J'ai déjà eu l'occasion de montrer que la relation entre fiction et vérité est plus complexe, et c'est là que se situent tout l'intérêt et la portée de cette approche.

Ensuite, il faut éclaircir un point de contradiction apparente dans la problématique telle que je l'ai traitée jusqu'ici. J'ai prétendu dans le chapitre précédent que la notion de représentation épistémique permettait d'approcher le problème de la représentation d'une façon unifiée, c'est-à-dire sans distinguer les types d'objets que pouvaient être les modèles. En particulier, une conception de la représentation devra traiter un modèle réduit matériel et un ensemble d'équations abstraites de la même manière du point de vue de leur relation avec le système cible qu'ils représentent. Or, la notion de fiction telle qu'elle est généralement comprise met de côté les modèles matériels. Si on peut accepter que les sphères parfaites d'un modèle astronomique soient comparables à des personnages de fiction par leur aspect imaginaire, la problématique est beaucoup moins claire avec le modèle réduit d'un pont censé être soumis à des tests en soufflerie. Pour pousser la contradiction apparente encore plus loin, je doute que beaucoup iraient jusqu'à qualifier une peinture figurative de fiction. Pourtant, selon mon développement, il faut que ça soit le cas : les deux exemples constituent des représentations épistémiques. Utiliser la notion de fiction pour comprendre ces dernières force

donc à proposer une définition qui rende compte de cette diversité dans un cadre unifié.

Il y a deux conséquences à cette tension apparente, dont une qui a déjà influencé un choix de vocabulaire que je n'avais pas justifié jusqu'ici. La première fournit en quelque sorte une condition supplémentaire au traitement de la fiction qui sera appliqué aux modèles : la notion de fiction doit être assez large pour recouvrir les différents usages concrets et abstraits des modèles.

La seconde conséquence concerne le contenu des deux figures (2.1 et 2.2) qui illustraient la différence entre représentation directe et indirecte. J'ai fait le choix de vocabulaire suivant : dans le cas de la représentation indirecte, des hypothèses génèrent les modèles, tandis que dans la représentation directe, des phrases décrivent le système cible, sans passer par l'objet intermédiaire qu'est le modèle. La manière dont des hypothèses peuvent spécifier les caractéristiques d'un modèle est au centre de la théorie de la fiction privilégiée dans les approches qui tentent d'exporter ce concept à l'épistémologie et à la philosophie des sciences, je le verrai par la suite.

Cette distinction prendra notamment toute son importance quand j'aborderai la question des idéalizations et des systèmes manquants (*missing systems*, voir (Thomson-Jones 2009; Frigg et Nguyen 2020; Cassini 2018)), c'est-à-dire des modèles qui représentent des systèmes qui n'existent pas. En effet, un modèle est spécifié par une série d'assertions (donc de phrases) qui le décrivent, mais comme je l'ai fait remarquer, il est problématique de réduire le modèle à cette description langagière. Le cas des *missing systems* fait apparaître clairement ce problème : comment concevoir la description d'un système qui n'existe pas ? Quel est le statut d'une telle description ? La conception indirecte de la représenta-

tion résout cette tension en introduisant un objet intermédiaire potentiellement abstrait et comparable à une fiction : le modèle. La conception directe, quant à elle, procède par la notion de description incomplète, je le montrerai plus tard. L'un des apports principaux des théories de la fiction mobilisées dans les débats d'épistémologie de la modélisation est que ces deux conceptions peuvent être formulées dans un même cadre. La conception de la représentation qui en découle est par conséquent assez générale pour recouvrir les différents aspects des représentations épistémiques, ce qui est une des conditions que Frigg et Nguyen mettaient en avant.

3 Le « comme si » de Vaihinger : première incursion de la fiction

L'un des ouvrages les plus importants consacrés à la notion de fiction dans un contexte de philosophie de la connaissance est *La philosophie du « comme si »* de Hans Vaihinger (*Philosophie des Als Ob*), publié originellement en 1923 et traduit en français en 2008 (Vaihinger 1923).

La façon dont Vaihinger y développe sa conception fictionnelle est particulièrement intéressante pour mon propos, car elle se conçoit comme une procédure en vue d'atteindre un certain but plutôt que comme un objet dont il s'agirait de comprendre les propriétés (c'est ce qu'il nomme *activité fictionnante*). Ce qui caractérise la fiction, c'est selon lui un mode de pensée particulier, une attitude qui pousse à concevoir une chose ou un ensemble de choses *comme si* il en allait en fait autrement. L'essentiel du livre se compose d'une longue liste d'exemples que Vaihinger détaille en montrant comment, à chaque fois, une posture du *comme si* est au

travail.

Cette activité, conçue comme une capacité logique de l'esprit, mène à imaginer des faussetés, mais l'erreur serait de les disqualifier pour cette unique raison :

Quoi qu'il en soit, inférer que certaines constructions sont inutiles parce qu'elles sont fausses, comme cela arrive souvent, est un signe très fâcheux d'inintelligence scientifique.
(Vaihinger 1923, p. 17)

Il y a dans cette citation l'aspect pragmatique (et, il faut le reconnaître, instrumentaliste) de Vaihinger : les fictions sont utiles à l'activité scientifique, mais pas sur le plan d'une description réaliste des phénomènes. En effet, leur fausseté n'est pas une raison suffisante pour chercher à les éliminer de la science, même si elles ne sont pas pour autant un guide vers la vérité :

Toute connaissance, dès qu'elle dépasse le simple constat d'une succession et d'une coexistence factuelles, ne peut être qu'*analogique*. (Vaihinger 1923, p. 22)

Puis, au paragraphe suivant, à propos des catégories conçues comme des fictions :

Par une voie entièrement différente, nous atteignons le même résultat que la philosophie kantienne : les catégories ne sont pas à même de saisir la réalité en elle-même, car comme fictions analogiques, elles ne peuvent nous fournir aucune connaissance vraie. (Vaihinger 1923, p. 23)

Pour Vaihinger, les fictions analogiques ne peuvent pas fournir de connaissance vraie et toute connaissance qui ne se réduit pas à une des-

cription d'états de faits ne peut-être qu'analogique. Toute la connaissance scientifique théorique qui dépasse la simple description de phénomènes est donc incapable de mener à la vérité.

Il n'est pas rare dans la même veine de voir des auteurs utiliser les termes de « fictions utiles » pour décrire cette attitude instrumentaliste. L'idée sous-jacente en est toujours la même : étant donné que les buts de l'activité scientifique sont établis selon des critères essentiellement empiriques et des normes d'acceptation qui évitent toute référence à la vérité, toutes les stratégies sont bonnes à prendre. Utiliser des fictions permet de plus facilement se représenter des phénomènes, de les comprendre et d'en débattre. Elles ont donc une place tout à fait légitime dans la méthodologie, mais il ne faut pas pour autant prétendre qu'elles disent le vrai.

Arthur Fine, dans son fameux article « fictionalism »¹ (Fine 1993) (réédité dans (Fine 2009)) met en évidence le rapprochement qu'effectue Vaihinger entre utilité et scientificité des fictions :

Vaihinger also distinguishes between virtuous and vicious fictions, those that are scientific and those that are unscientific. The scientific fictions are an effective means to certain ends; they are useful and expedient. Where this utility is lacking the fictions are unscientific. (Fine 2009, p. 23)

La mesure de la scientificité des fictions est donc leur utilité. C'est une perspective purement pragmatique.

1. Fine utilise le terme « fictionalisme » dans le sens instrumentaliste que j'ai décrit plus tôt : « "Fictionalism" generally refers to a pragmatic, antirealist position in the debate over scientific realism. The use of a theory or concept can be reliable without the theory being true and without the entities mentioned actually existing. When truth (or existence) is lacking we are dealing with a fiction. Thus fictionalism is a corollary of instrumentalism, the view that what matters about a theory is its reliability in practice, adding to it the claim that science often employs useful fictions. » (Fine 1998, p. 667)

Le pas supplémentaire, le cœur de ce que j'avais qualifié d'optimisme épistémique du réalisme, qui vise à tirer de l'adéquation empirique une forme de vérité ou de correspondance, ne peut donc pas se faire, selon Vaihinger :

Tout ce qu'on peut dire des phénomènes objectifs, c'est que tout se passe comme s'ils se comportaient de telle ou telle manière ; mais il n'existe absolument aucune justification pour remplacer dogmatiquement cette formulation par :
« les phénomènes se comportent effectivement ainsi ». (Vaihinger 1923, p. 24)

En effet, on voit que Vaihinger disqualifie entièrement la prétention à la correspondance avec les phénomènes objectifs des fictions. Le *comme si* le plus pragmatiquement efficace ne peut servir de justification à l'établissement d'une vérité.

Quant à la notion même de vérité, sa définition sous ce prisme pragmatique s'associe naturellement au rejet de l'idée de correspondance :

Partant, nous nommons notre monde d'idées « vrai » s'il nous permet de mieux maîtriser l'objectivité et d'agir en elle. Il faut donc finalement renoncer au critère du soi-disant accord avec la réalité. (Vaihinger 1923, p. 94)

L'approche du *comme si* met donc en avant la variété de stratégies scientifiques qui utilisent l'imagination d'entités ou de phénomènes fictionnels, mais ne permet pas en l'état de bâtir une théorie de la représentation adéquate. On ne trouve aucune ressource qui permettrait de bâtir une relation entre un modèle fictionnel au sens de Vaihinger et le système concret qu'il est censé représenter. Les processus fictionnants sont avant tout l'objet d'une attitude de l'imagination, d'une certaine

distance prise avec les énoncés considérés : l'effet d'une proposition p est différent de la considération « tout se passe *comme si* p ».

Dans une perspective vaihingerienne, une définition possible de la représentation est la suivante : un modèle représente un système cible si et seulement si l'utilisateur du modèle considère que tout dans la cible se passe *comme* dans le modèle. En précisant le contenu du « comme », il devrait être possible de formuler des conditions qui garantiraient des inférences à la cible (c'était une des conditions posées par Frigg et Nguyen). En partant du vocabulaire de Vahinger à propos de la fiction, ce « comme » fait en quelque sorte office de relation de ressemblance, mais rien dans sa théorie ne permet de le définir autrement que par rapport à l'attitude que la fiction suscite chez son utilisateur. Cette façon de concevoir la représentation souffre donc des mêmes problèmes liés à l'intentionnalité abordés dans le chapitre précédent : est une représentation ce qui est utilisé par un sujet dans le but de représenter.

Pourtant, il y a de toute évidence un apport majeur de Vaihinger dans la manière dont on peut comprendre les modèles scientifiques et notre attitude vis-à-vis d'eux. Si je reprends les exemples des modèles astronomiques du système solaire ou de l'atome de Bohr, le parallèle est évident. Nous savons que le système solaire n'est pas composé de planètes parfaitement sphériques et à distribution de masse homogène, mais quand nous utilisons le modèle pour prédire les dates des prochaines éclipses, nous faisons *comme si* c'était le cas. Nous savons que les électrons ne sont pas des corpuscules animés de mouvements circulaires autour des noyaux atomiques, mais pour reproduire le spectre en émission de l'hydrogène, nous faisons *comme si* c'était le cas.

L'analyse de Vaihinger permet de pousser un cran plus loin l'analogie entre modèles et fiction : la reconnaissance d'éléments produits

par l'imagination était un premier pas, le second est de montrer que la fiction, envisagée sous l'angle du *comme si*, permet d'expliquer notre attitude envers les modèles ainsi que l'usage que nous en faisons. Cependant, réduire la représentation de systèmes cibles par les modèles à une attitude et des normes d'usage ne permet pas de formuler une position réaliste. C'est exactement le problème auquel les autres conceptions de la représentation étaient confrontées. De plus, la proposition de Vaihinger ne peut être ni considérée comme une théorie complète de la fiction (car il ne faut pas oublier le versant esthétique de la question, même s'il est secondaire pour mon propos, le terme de fiction y reste fortement attaché)², ni comme une théorie adéquate de la représentation, je viens de l'illustrer.

Les leçons à en tirer sont tout de même nombreuses. Tout d'abord, Vaihinger rend compte de l'aspect logique de l'activité fictionnante. Les fictions (entendues comme les produits de cette activité) ont une place légitime dans la réflexion et la connaissance scientifiques. La fiction est en ce sens un mode de réflexion bien particulier dont il s'agit de comprendre le mécanisme. Ensuite, si le pragmatisme de Vaihinger charrie un instrumentalisme évident, il illustre également qu'une compréhension de la fiction ne peut se limiter à une caractérisation basée sur la fausseté des éléments fictionnels :

Nous avons plusieurs fois souligné que la frontière entre vé-

2. Vaihinger est bien entendu conscient du sens très large de la notion de fiction dans le langage commun. Pour clarifier les choses, il propose la distinction suivante : « Je propose à l'avenir de nommer « fictions » toutes les fictions scientifiques, et « figments » toutes les autres, par exemple les fictions mythologiques, esthétiques, etc. Ainsi, Pégase est un « figment », l'atome est une « fiction ». » (Vaihinger 1923, p. 69) La différence se situe dans l'absence de clause *comme si* dans les figments. Il me semble que l'un des avantages d'une théorie de la fiction comme on la trouve chez Kendall Walton est qu'elles rendent compte de ces deux aspects dans un même cadre. La proposition de Walton est aussi plus adaptée au traitement du problème de la représentation. Ce sera justement l'objet de la section suivante.

rité et erreur n'était pas figée, et finalement il est apparu que ce que nous appelons généralement la *vérité*, c'est-à-dire, pour reprendre les termes en usage, un monde d'idées s'accordant avec le monde extérieur, *n'est autre que l'erreur la plus utile.* (Vaihinger 1923, p. 94)

Le mouvement opéré par Vaihinger est en quelque sorte l'inverse de celui dont j'ai besoin : plutôt que de faire de la place à la vérité dans les fictions, comme je me propose de le faire, il considère la vérité comme une fiction parmi d'autres.

On trouve tout de même chez lui quelques éléments qui permettent d'élaborer un début de réponse réaliste, ou du moins de formuler certaines oppositions que me seront utiles. Il y a d'abord la différence entre ce que Vaihinger appelle les fictions et ce qu'il nomme « semi-fictions » ou « demi-fictions ». Ces deux catégories forment en fait les deux pôles d'un continuum qui se caractérise par le niveau de divergence avec la réalité supposée des phénomènes considérés. Les semi-fictions entrent en contradiction avec la nature, contradiction introduite par l'opération du « comme si », tandis que les fictions (que l'on peut qualifier de « pures fictions » par opposition) sont non seulement en contradiction avec la nature, mais aussi avec elles-mêmes. L'exemple le plus éclairant de demi-fiction est celui des catégories que nous utilisons pour classer les phénomènes. L'atome est un exemple de pure fiction, du moins avec le statut qui lui était accordé à l'époque de Vaihinger. Cette distinction laisse la place à différents niveaux d'abstractions, d'idéalisations et d'approximations qui sont des éléments importants de la modélisation et plus généralement de la pensée scientifique.

La seconde opposition est celle entre hypothèse et fiction. Vaihinger fait remarquer que ces deux modes de pensée sont souvent confondus,

et que ce qui les différencie est une attitude pragmatique. Une hypothèse concerne la réalité, elle a la prétention de coïncider avec des perceptions futures. Elle nécessite donc une vérification, c'est-à-dire un test quelconque qui mette à l'épreuve son éventuel accord avec un état de fait. La fiction, en revanche, est un outil dont la seule mesure de l'acceptabilité est l'utilité. L'emploi de ces deux modes dans la réflexion scientifique est par conséquent différent :

C'est pourquoi entre plusieurs hypothèses également possibles on choisit toujours la plus vraisemblable. Entre plusieurs fictions également possibles, en revanche, on choisit la plus opportune. Ceci montre clairement la différence entre les deux constructions. (Vaihinger 1923, p. 72, note 4)

Je pense que ce point est fondamental pour bien comprendre l'enjeu des approches fictionnelles des modèles scientifiques et de la représentation. J'avais utilisé le terme d'« hypothèse » pour désigner de manière très générale la façon dont les modèles étaient bâtis. Poser des hypothèses s'apparente à la génération d'un modèle, cela se retrouve dans leur formulation : « imaginons que... » ou « faisons *comme si* ». Vaihinger oppose hypothèses et fiction et j'ai maintes fois annoncé que j'exploiterais la notion de fiction afin de comprendre les modèles, mon usage peut donc sembler contradictoire avec celui de Vaihinger. Il semble cependant que ça ne soit pas le cas. En effet, hypothèse et fiction ne peuvent pas toujours être séparées de manière stricte : ce qui est au départ un outil pragmatique peut se révéler digne d'une attitude orientée vers la vérité, tandis qu'une hypothèse peut être réfutée par une expérience tout en restant utile à la connaissance en tant que fiction.

Ce qui se joue là est, il me semble, le cœur des approches fiction-

nelles contemporaines. Il n'est pas rare de lire une référence à Vaihinger, mais cela ne dépasse jamais l'analogie très vague ou la récupération du « comme si ». La distinction hypothèse/fiction est pourtant centrale.

Une théorie adéquate de la représentation (indirecte)³ doit à la fois rendre compte du mécanisme de création du modèle (c'est-à-dire de la fiction) et de sa capacité représentative. Dans ma conception, je montrerai que les hypothèses jouent ce rôle créateur, et que cela n'est pas en contradiction avec la possibilité d'une théorie de la représentation. Les hypothèses témoignent d'un intérêt du modèle pour la vérité et le processus fictionnant exploite la liberté heuristique de l'imagination. Voilà en deux idées les éléments d'une théorie fictionnelle des modèles.

Mais l'approche du *comme si* que propose Vaihinger constitue un squelette de base extrêmement pertinent, sur lequel il est tout à fait possible de greffer des idées qui permettront une analyse plus poussée de la représentation. On trouve une théorie qui répond bien à cette description chez Kendall Walton, qui utilise le concept de jeu de *make-believe*. Il existe une filiation claire entre le *comme si* de Vaihinger et le « faire-semblant » (Schuppert 2021) de Walton. Le principe de base repose sur la même capacité d'un sujet à raisonner à partir d'éléments imaginés, c'est-à-dire la possibilité de composer des situations partiellement irréalistes (ou irréelles) et d'avoir envers elles – de jouer à avoir – une attitude semblable à celle que le sujet aurait face à une situation réelle. Le *make-believe* figure au centre des traitements fictionnels de la représentation scientifique comme je le verrai dans la section suivante.

3. Vaihinger conçoit les fictions comme les objets d'un processus fictionnant, ce qui me semble mieux correspondre à la vision indirecte de la modélisation. Il pourrait être intéressant d'adapter ou de comparer le cadre qu'il propose avec la vision directe de la représentation, mais c'est une question qui dépasse le problème qui m'occupe ici.

4 Le *make-believe* de Walton et la représentation

L'objectif principal de l'ouvrage *Mimesis as Make-Believe* de Kendall Walton (Walton 1990) est de construire une théorie de la représentation dans les arts, ou du moins de mener une enquête sur le fonctionnement de ce que l'on appelle généralement la représentation dans les arts, comme le montre son sous-titre « on the foundations of the representational arts ». Cette section est dédiée à une description des éléments les plus importants de l'approche de Walton, sans accent particulier sur les questions d'épistémologie ou de modélisation scientifique, sur lesquelles je me concentrerai dans la section 5.⁴

L'idée du parallèle entre les modèles scientifiques et la fiction partait de la constatation de la présence dans les deux cas d'éléments imaginaires. Or, l'étude de quelques exemples avait rapidement montré la limite d'une telle approche. Les choses que l'on appelle fiction se caractérisent par l'attitude que l'on adopte quand on y est confronté, ainsi que par l'usage que l'on en fait. Pourtant, il ne faudrait pas faire l'erreur de disqualifier entièrement l'élément imaginaire, qui est certainement central dans la compréhension. C'était le rôle du *comme si* d'opérer cette jonction en montrant quelle place avait l'imagination dans les processus fictionnants, et Walton ne prend pas une autre voie pour aborder le problème de la représentation.

Sa proposition est en ce sens orientée vers l'utilisateur, ou le spectateur au sens large, mais elle prend appui sur ce qui compose la fiction

4. Je me permets tout de même une annonce du caractère plus général du cadre de pensée qu'il propose, tirée de l'introduction de l'ouvrage, pour dissiper tout doute qui subsisterait : « There is nothing distinctively "aesthetic" about make-believe itself at all » (Walton 1990, p. 7).

de la manière la plus concrète : des taches de couleur sur une toile, des mots imprimés sur une page, des images projetées sur un écran.

Le terme de *make-believe* fait référence aux jeux d'enfants dans lesquels on fait semblant, c'est-à-dire que les comportements dans le jeu sont à comprendre par rapport à ce contexte et non de manière littérale. Walton considère ces exemples comme des cas semblables à ceux des représentations artistiques. Plus précisément, certains accessoires supportent l'imaginaire au travail dans le jeu, et Walton compare leur rôle à celui des œuvres. Une poupée, par exemple, utilisée pour soutenir un jeu de faire semblant gagne un statut : à l'intérieur du jeu, ce n'est plus un morceau de plastique et de tissu, mais un être humain doué de conscience et de volonté. Ces accessoires, il les appelle des *props*.

La proposition de Walton peut alors s'énoncer comme suit : les œuvres d'art représentationnelles⁵ sont des *props* dans un jeu de *make-believe*. Ce sont certes des modes de jeux plus subtils et complexes que dans le cas des poupées, mais leur nature ne diffère pas fondamentalement.

La représentation fonctionne à partir de deux concepts : les *props* et les principes de génération (*principles of generation*). Les principes de génération sont comme les règles du jeu, ils prescrivent que certaines choses doivent être imaginées selon la présence ou non des *props* associés.

Encore une fois, l'exemple des jeux éclaire parfaitement la situation. Imaginons plusieurs personnes qui jouent dans la forêt à chasser des extra-terrestres terriblement menaçants. Un principe de génération du jeu pourra prendre une forme du genre : « on dira que les arbres sont

5. Walton utilise les termes « représentation » et « fiction » de manière interchangeable. Les objets qui représentent selon la conception du *make-believe* sont donc des fictions dans un sens plus large que celui des œuvres de fiction.

des extra-terrestres » ou encore « faisons *comme si* les arbres étaient des extra-terrestres ». Dans ce jeu de faire semblant, les arbres sont alors des *props*, ils soutiennent l'acte d'imagination spécifié par le principe de génération.

Immédiatement, cela entraîne des conséquences sur la structure interne de la fiction. Si un des joueurs prétend que la voie est libre alors qu'un arbre se trouve non loin de là, il énonce quelque chose de faux dans la fiction. Les raisons peuvent être variées : il peut ne pas avoir vu l'arbre, et dans ce cas il se trompe et s'expose à la menace extra-terrestre à cause de son imprudence, mais il se peut aussi qu'il ait mal compris les règles du jeu. Dans tous les cas, les principes de génération et les *props* prescrivent certaines choses à être imaginées, mais pas d'autres. À partir de là, il est possible de certifier que certaines propositions à l'intérieur du jeu sont vraies, et d'autres fausses. Les propositions vraies dans une fiction sont appelées des vérités fictionnelles (ou plus simplement, on dira que la proposition p est fictionnelle si p est une prescription imaginative du jeu).

Une deuxième conséquence est qu'il existe des vérités fictionnelles inconnues. C'est souvent une erreur de considérer que toute fiction est un monde clos et complet, dont tout le contenu doit être explicitement spécifié par l'auteur. Tout le reste est faux ou indéterminé, dépendant de la position philosophique adoptée. Ce point de vue peut entraîner des conséquences étranges et contre-intuitives. Par exemple, combien de cheveux possède Holden Caulfield au moment où il quitte l'internat ? Ce nombre n'est pas précisé dans le roman, mais est-il pour autant indéterminé ? La proposition « Holden possède un nombre pair de cheveux \vee Holden possède un nombre impair de cheveux » est-elle elle aussi indéterminée, voire fausse, bien que tautologique ? La théorie de Walton

traite ces cas d'une manière détournée en concentrant l'analyse sur les prescriptions imaginatives : en lisant « L'attrape-cœurs », le lecteur est amené à imaginer un garçon de seize ans, qui possède des cheveux (ce détail est, pour le coup, mentionné par l'auteur), et qui a toutes les caractéristiques auxquelles on s'attend chez un humain. Un lecteur du roman qui s'imaginerait un Holden à la peau verte et dont la chevelure défie les théorèmes élémentaires de l'arithmétique serait un utilisateur déficient de la fiction, il ne jouerait pas au jeu de faire-semblant de manière adéquate.

Il n'est pas nécessaire d'introduire un principe de modification minimale pour comprendre ce point. J'entends par « principe de modification minimale » un principe qui garantirait qu'en l'absence de mention contraire explicite, tout élément d'une fiction découle strictement du monde réel tel qu'il est conçu par le lecteur (ou censé être conçu par le lecteur selon les croyances de l'auteur à propos de son public). Par exemple, la loi de gravitation s'applique dans le Londres de Tyrone Slothrop, Holden Caulfield possède un foie situé à droite de son corps et sa peau n'est pas verte, les mots que prononcent les personnages des *Instructions* son bel et bien de l'anglais et non pas une langue homophonique et pourtant sémantiquement distincte. Ces caractéristiques de la fiction font partie du contrat tacite qui est passé entre l'auteur et le lecteur, et c'est d'ailleurs en pleine connaissance de cet effet que l'auteur peut déployer son monde imaginaire.

Ceci amène Walton à considérer des jeux de *make-believe* autorisés (*authorized games*), par contraste avec ceux dont les principes de génération sont arbitraires. L'exemple que j'avais donné dans le chapitre précédent entre dans cette seconde catégorie : la fourchette était un *prop* dans un jeu qui décrétait que cet accessoire représentait telle ou telle

rue. En revanche, les mots écrits sur une page font partie d'une classe de jeux autorisés, dans lesquels les règles de fonctionnement élémentaires sont largement acceptées par les utilisateurs de la fiction. Un lecteur qui comprendrait le texte selon une interprétation toute personnelle ne joue donc pas avec les règles adéquates.

Cela permet à Walton de définir la représentation : un *prop* est une représentation si et seulement si il figure dans un jeu autorisé. De ce fait, sa conception exclut le cas des arbres extra-terrestres ou de la fourchette topographique. Le problème de rendre la représentation triviale par un simple acte de stipulation est réglé, mais uniquement au prix d'un report de la charge d'intentionnalité sur la communauté qui utilise la fiction, et sur les normes qui ont émergé en son sein.

La notion de vérité acquiert un double statut : interne et externe. En interne, les principes de génération et les *props* engendrent une série de vérités fictionnelles à l'intérieur du monde fictionnel, quant aux vérités externes, elles n'influent pas sur le statut fictionnel de la représentation. Une fois engagés dans le jeu, il n'importe pas que les objets ou événements fictionnels aient une contrepartie externe. Si Holden Caulfield se trouvait exister, ça n'en ferait pas moins de l'attrape-cœurs une fiction au sens du jeu de *make-believe*, puisque la lecture nous invite à imaginer des situations qui le mettent en scène :

Could an author be claiming truth for every sentence he writes and still be writing fiction ? I see no reason why not[.]
(Walton 1990, p. 79)

D'une manière analogue, ce dispositif explique également pourquoi et comment le portrait de Napoléon représente Napoléon : nous sommes invités en contemplant l'œuvre à imaginer le Napoléon réel se tenir dans

telle ou telle pose, dans tel ou tel contexte, etc.

Le statut des vérités fictionnelles est lui-même indépendant de l'utilisation qui est faite de la fiction. De la même manière qu'il peut exister des vérités fictionnelles inconnues, les *props* génèrent ces vérités qu'elles soient ensuite imaginées ou non par un lecteur ou spectateur :

Props are generators of fictional truths, things which, by virtue of their nature or existence, make propositions fictional.

(Walton 1990, p. 37)

Une fois généré, le monde fictionnel — entendu comme l'ensemble des propositions fictionnelles — gagne en indépendance vis à vis des spectateurs : c'est véritablement un monde que l'on peut explorer, dans lequel on peut raisonner à partir des principes de génération. L'exemple littéraire le plus éclatant est probablement l'ouvrage de Pierre Bayard *Qui a tué Roger Ackroyd?* (Bayard 2008), dans lequel l'auteur démontre que la solution au mystère du *Meurtre de Roger Ackroyd* proposée par Agatha Christie ne tient pas debout, que son détective fétiche Hercule Poirot doit s'être trompé et que la vérité fictionnelle est tout autre.⁶

5 *Make-believe*, fictions et modèles scientifiques

La théorie de la fiction basée sur les jeux de *make-believe* de Walton a influencé tout une gamme de conceptions des modèles scientifiques

6. Pour que ce genre d'enquête puisse avoir lieu, il faut évidemment faire de nombreuses suppositions qui ne sont pas explicitées dans le roman, par exemple que les personnages ne peuvent pas se trouver à deux endroits au même moment. Comme dans le cas du nombre de cheveux de Holden Caulfield, défendre la thèse de Poirot en faisant appel à ce genre de possibilité sous prétexte que le roman est une fiction serait un manquement au contrat passé entre le lecteur et l'auteur, donc un mauvais usage de la fiction.

qui, partant de l'analogie développée en début de chapitre, ont cherché à prolonger l'analyse fictionnelle pour, notamment, traiter du problème de la représentation.

Plusieurs des éléments du *make-believe* sont aisément transférables à une discussion à propos des modèles scientifiques.

Tout d'abord, les modèles hautement idéalisés peuvent être compris non comme des descriptions véridiques de leur cible, mais comme des invitations à imaginer certaines choses à propos de la cible. C'est une démarche analogue au *comme si* de Vaihinger, mais où l'accent est mis sur l'effet recherché chez l'utilisateur du modèle. Il n'est d'ailleurs pas rare de lire en tant que description d'un modèle des clauses du type « supposons que » ou « considérons telle situation » (Frigg et Nguyen 2020, p. 121). Confronté au modèle du système Solaire, le scientifique est invité à imaginer des sphères, qui ont telles et telles caractéristiques, et qui n'interagissent que selon la loi de gravitation universelle. La question de l'utilité empirique de ce modèle ne se pose que dans un second temps.

Ensuite, il y a l'exploration du modèle. Une fois les règles du jeu dictées (les principes de génération), l'utilisateur du modèle est invité à entrer dans le monde de la fiction et à y dériver de nouvelles vérités. En acceptant de faire semblant de croire qu'il y a des sphères en interaction gravitationnelle, l'utilisateur peut en déduire leurs trajectoires, tout comme il peut déduire les intentions cachées de tel ou tel personnage en fonction de son attitude et de sa description psychologique, par exemple. Dans le cas scientifique, on voit un rapprochement des lois (de Newton, dans cet exemple) et des principes de génération : la loi physique est en quelque sorte la règle du jeu du monde de la fiction. Cela permet à la fois de laisser la place à la recherche — il y a un sens à considérer que l'on peut découvrir des vérités inconnues dans le modèle, et donc d'acquérir

une forme de connaissance à son sujet — et de décrire l’usage qui est fait des modèles en pratique : un modèle est d’abord étudié en soi, avant toute tentative d’interpréter les résultats en les transférant à la cible.

5.1 *Make-believe* et la distinction représentation directe/indirecte

Si l’idée fondamentale de Walton semble reproduire assez bien les intuitions relatives à l’usage des modèles en sciences, le problème se corse dès que l’on cherche à reconstruire les normes épistémiques qui guident la recherche scientifique, particulièrement quand la relation entre le modèle et la cible est abordée.

Comme je l’ai montré plus haut, la théorie du *make-believe* est très efficace pour décrire le fonctionnement interne d’une fiction (découverte de vérités fictionnelles à partir de principes de génération et des *props*) ainsi que l’attitude qu’elle suscite chez l’utilisateur (une invitation à imaginer telle ou telle situation). Ce qui caractérise une représentation est qu’elle fait partie d’un jeu autorisé, c’est-à-dire un jeu dont la communauté des joueurs s’accorde sur les règles. Cela ne permet pas de concevoir une relation modèle-cible épistémologiquement féconde, dans le sens où elle ne se conçoit que relativement à un sujet qui imagine relativement à des normes de la communauté. Walton insiste sur l’idée que les vérités fictionnelles sont des vérités indépendamment du fait d’être imaginées ou non, c’est-à-dire que ce qui est vrai dans une fiction le reste même s’il n’y a personne pour l’imaginer, mais cela ne change pas la définition de représentation, qui fait explicitement appel à la notion de jeu autorisé, et donc de normes en usage.

Le *make-believe* peut facilement être adapté aux modèles non ma-

tériels : des équations écrites sur du papier peuvent définir un monde fictionnel abstrait, tout comme des mots imprimés sur du papier ou des taches de couleur sur une toile. La fameuse citation de Ian Hacking à propos des modèles mathématisés de la physique :

a model in physics is something you hold in your head rather than your hands (Hacking 1983, p. 216)

entre tout à fait dans ce cadre. S'il est évident que l'on manipule des modèles matériels, manipuler un modèle abstrait, c'est chercher de nouvelles vérités fictionnelles en se servant des principes de génération et des *props*. Ces derniers sont toujours décrits selon leur composante matérielle, mais cela vient simplement du fait que la description du modèle doit être énoncée d'une manière ou d'une autre. On peut aisément imaginer une situation où des équations seraient pensées par un scientifique sans qu'il les écrive, mais cela ne change rien au mécanisme fictionnel : même sans être matérialisées par l'écriture, ces équations servent du support à l'imagination et à la découverte de nouvelles vérités.

La question qui se pose est celle du statut de ces fictions. Adam Toon (Toon 2012) discute un exemple qui illustre parfaitement le problème : le modèle du pendule simple. Afin de décrire le comportement des pendules, un physicien peut imaginer une situation idéalisée dans laquelle une masse parfaitement sphérique oscille au bout d'une corde en l'absence d'autres forces que la gravitation. Ce pendule simple imaginaire se conçoit aisément comme un jeu de *make-believe* : les équations de base qui décrivent la situation ainsi que la loi de gravitation sont des principes de génération tout à fait analogues aux autres exemples que j'ai donnés plus haut.

Dans ce cas, il n'est pas nécessaire de faire référence à un « sys-

tème intermédiaire » : le pendule simple en tant que fiction n'est pas une entité abstraite ou un *missing system* qu'il s'agit de décrire et de comprendre pour, dans un second temps, chercher à en apprendre sur le système cible en cherchant une quelconque ressemblance entre les deux. Au contraire, il semble plus simple, au sens où bien des problèmes ontologiques sont évités, de considérer le modèle du pendule simple comme une description imprécise ou idéalisée des véritables pendules. La description du pendule simple ne génère donc pas une entité abstraite et fictionnelle qui sera comparée aux pendules de la réalité physique, mais elle constitue une description incomplète :

If theoretical principles are understood in this way then there is no reason to think that there needs to be any object which they describe, any more than we need to posit a new object every time we give a false description of something.
(Toon 2012, p. 44)

Il n'est pas nécessaire, pour Toon, de considérer qu'un objet qui correspond à la description existe. Le pendule simple n'est pas un objet idéal qui ressemble aux pendules réels, mais ce que nous appelons pendule simple est une description incomplète et approximative des pendules.

En termes de *make-believe*, Frigg et Nguyen définissent la représentation directe de la manière suivante :

a carrier X is an epistemic representation of target T iff X functions as a prop in a game of make-believe prescribing imaginings about T . (Frigg et Nguyen 2020, p. 127)

L'objet porteur qui sert de modèle prescrit des imaginations à propos de la cible, et c'est cela qui fait de lui une représentation épistémique.

Ce porteur peut être un objet matériel comme un ensemble d'équations mathématiques.

Cette approche a l'avantage d'éviter les problèmes liés à l'ontologie des modèles : il n'y a pas d'objet appelé « modèle » dont il faudrait comprendre la nature. En revanche, ce qui fait du modèle du pendule simple une représentation épistémique est moins clair : comment les inférences à la cible fonctionnent-elles si le modèle est lui-même une description de la cible ? De plus, comme je l'avais fait remarquer plus tôt, il est assez clair que le vocabulaire scientifique tend à considérer que les hypothèses qui génèrent un modèle créent par là même un système qui les satisfait. Nous parlons du pendule simple comme d'un objet qui a des propriétés bien déterminées (des vérités fictionnelles qui concernent un personnage à propos duquel nous avons beaucoup lu).

Toon fait appel à une notion d'approximation pour résoudre le problème : les inférences à la cible sont garanties si les propriétés du modèle (c'est-à-dire de la description incomplète de la cible) sont les mêmes, au moins approximativement, que celles de la cible.

Si on peut se laisser convaincre par ce traitement des approximations dans quelques cas particuliers, il me semble plus compliqué de baser une théorie de la modélisation et de la représentation sur ce mécanisme. Les sphères parfaites du modèle du système solaire sont des approximations des planètes réelles, et l'on peut admettre que les frottements peuvent être négligés, parce qu'ils sont faibles, dans l'étude du mouvement du pendule. Le problème soulevé par Poznic (Poznic 2016) et reconnu par Frigg et Nguyen, c'est qu'il n'est pas possible de différencier les énoncés du modèle qui doivent être interprétés comme (approximativement) vrais de la cible, et ceux qui ne le sont pas. Par exemple, pour reprendre le système Solaire fait de sphères en interaction gravitationnelle, com-

ment sait-on que la loi d'attraction en $1/r^2$ n'est elle pas elle-même une approximation d'une force supposée réelle ? Si le *make-believe* nous invite à imaginer des situations directement à propos de la cible, le niveau de crédibilité associé aux inférences soutenues par ces actes de l'imagination n'est pas expliqué. La différence entre un modèle sans cible (l'atome de Bohr n'existe pas dans la nature) et la description partielle (quelque chose comme l'atome de Bohr, c'est-à-dire suffisamment semblable, existe et est la cible du modèle) n'est pas constructible dans cette approche.

On pourrait répondre que le problème se pose aussi en dehors des modèles scientifiques. Un buste de Napoléon permet des inférences à la cible, nous engage dans un jeu de *make-believe*, et le bon usage de la fiction contient déjà la distinction entre les énoncés fictionnels transférables et ceux qui ne le sont pas : nous pouvons déduire de l'observation de la statue que Napoléon a tels ou tels traits, mais pas qu'il est fait de bronze ou de marbre. Nous pouvons déduire des trajectoires des planètes du modèle du système Solaire, mais pas que les planètes sont parfaitement sphériques. Quant à la force de gravitation, elle possède encore un statut, celui de loi, qui est, c'est bien connu, le sujet d'un problème philosophique à part entière. Le bon usage du modèle suppose que l'on sache comment transférer les vérités fictionnelles à la cible, mais cela réintroduit l'utilisateur ou les normes en vigueur dans l'analyse épistémologique.

L'analyse de Toon ne rejette pas le discours à propos de ce qui est « vrai dans le modèle », mais nie qu'il soit nécessaire de concevoir les modèles comme des objets (concrets ou abstraits, peu importe) pour que ce discours ait du sens :

So understanding talk about what is “true in our model”

does not require us to posit any object of which that talk is true. (Toon 2012, p. 49)

Cela permet d'éviter un problème logique lorsque l'on formule des propositions à propos des objets qui figurent dans le modèle. Ce qui est vrai dans le modèle est ce qui correspond aux règles qui génèrent la fiction, sans qu'il soit pour autant nécessaire de supposer que ces objets existent. Par exemple, il n'est pas nécessaire de supposer l'existence de planètes parfaitement sphériques pour que la description des trajectoires d'une telle planète dans le modèle newtonien ait du sens. La question sémantique se pose alors : quel est le statut des propositions qui portent sur la fiction ? Quel sens donner à une phrase du type « il est vrai dans le modèle que p » ?

Il serait simple d'utiliser une approche méta-fictionnelle pour expliquer cela : nous entrons dans un jeu de *make-believe* en considérant la modélisation *comme s'* il y avait des objets appelés modèles, car c'est utile et pratique de le faire, mais sans pour autant que ces objets existent. Cette proposition de solution me semble problématique pour deux raisons. Premièrement, le problème de la représentation s'en trouve obscurci : la représentation est un jeu de *make-believe* contenant des objets étant eux-mêmes pour la plupart le résultat d'un jeu de *make-believe*. Deuxièmement, si l'on garde en tête l'idée d'aller plus loin que la définition de la représentation à partir du *make-believe*, on voit mal comment concevoir une relation modèle-cible autrement que dans un troisième jeu de *make-believe* qui engloberait les deux autres : nous faisons *comme si* une telle relation existait, relation entre un objet fictionnel et un système cible elle-même exprimée par un *comme si*. On peut déjà trouver les prémisses de ce problème chez Walton, quand il traite des propositions qui portent sur la fiction. Les propositions du genre « Sherlock

Holmes est un détective brillant » portent sur une fiction, et ne peuvent que se comprendre dans un jeu plus général dans lequel on fait *comme si* Sherlock Holmes existait et pouvait être l'objet de propositions.

C'est un problème supplémentaire dans la recherche d'une formulation réaliste dans le contexte fictionnel puisque la relation entre le modèle et la cible, relation qui est donc extérieure au modèle, doit pouvoir se justifier de manière indépendante. Or, s'il n'est possible de la concevoir que dans un jeu de *make-believe*, le problème s'en trouve déplacé, mais nullement résolu.

5.2 Quelle fiction en épistémologie ?

Jusqu'ici, j'ai utilisé la notion de fiction dans deux contextes :

1. L'analogie avec les modèles : la reconnaissance de l'existence dans la pratique scientifique d'idéalisation, d'approximations et d'abstractions suscite un parallèle entre les modèles scientifiques et la fiction. Comme dans une œuvre de fiction, les modèles semblent décrire des situations irréelles ou impossibles, imaginaires, inventées. Cette liberté imaginative est certainement un élément important dans la manière dont des idées nouvelles peuvent émerger en sciences. Cette analogie constitue en quelque sorte la thèse fictionnelle faible : on acceptera ce point commun en insistant sur une différence fondamentale, à savoir que les modèles scientifiques ont une visée épistémologique que les œuvres de fiction n'ont pas. La fonction de la modélisation est d'expliquer des phénomènes, de permettre la compréhension, de développer des outils de résolution mathématiques, bref de participer à la construction de la connaissance. Par contraste, la fiction sert, parmi de nombreuses

autres choses, à véhiculer des émotions, divertir ou à développer des techniques esthétiques. Si l'on peut admettre qu'une œuvre de fiction participe d'une manière ou d'une autre à la connaissance et qu'il soit possible d'être divertie par l'exploration d'un modèle scientifique, c'est une caractérisation qui se base sur des cas trop marginaux pour fonder une définition de ces concepts.

2. Le problème de la représentation : la notion de fiction, à travers le jeu de *make-believe*, permet de construire une théorie de la représentation dont les caractéristiques semblent également adéquates pour comprendre la manière dont les modèles scientifiques représentent leurs cibles. C'est une thèse fictionnelle plus forte, étant donné que le parallèle sert d'heuristique, et pas uniquement d'illustration. En particulier, l'aspect imaginaire de certaines entités ou processus qui suscite l'analogie avec la fiction trouve une place naturelle dans l'activité de représentation dans le sens entendu par Walton. La description d'un modèle prescrit au scientifique d'imaginer certains phénomènes, entités et processus, et les principes de génération de cette fiction permettent de déduire de nouvelles vérités fictionnelles qui, c'est le but de la modélisation, pourront également être attribuées au système cible.

Frigg et Nguyen (Frigg et Nguyen 2020) discutent longuement comment l'approche de Walton permet de répondre aux cinq problèmes et cinq conditions d'adéquation qu'une théorie de la représentation doit remplir. J'avais annoncé que mon intérêt pour le problème de la représentation se limitait aux outils qui pouvaient être exportés à un traitement du problème épistémologique plus général du réalisme, il est temps de déterminer plus précisément quels points du débat une conception

fictionnelle des modèles peut aider à éclaircir. En particulier, je m'intéresserai aux ressources qui permettent de construire la relation modèle-cible ainsi qu'un mode de justification de cette dernière, ce qui est, je l'ai montré au chapitre précédent, une condition nécessaire pour formuler une position réaliste.

On l'a vu, si la théorie de Walton permet de comprendre beaucoup d'aspects de l'activité scientifique de modélisation, en particulier dans le fonctionnement interne des modèles, définir une relation de représentation avec la cible, par définition externe, est loin d'être clair. Faire appel à un jeu de *make-believe* supplémentaire ou à une clause *comme si* n'aidera pas. D'une part, cela ne fait que déplacer le débat puisque l'essentiel de ce qui constitue la représentation restera à comprendre, et d'autre part parce que cela introduit une couche supplémentaire d'intentionnalité. La fiction à la Walton se conçoit avant tout comme l'objet d'une certaine attitude, et si la relation de représentation est elle-même ce genre d'objet, alors elle ne peut être définie indépendamment du sujet, du contexte épistémique et d'une intention représentationnelle. Ce point est reconnu à la fois dans (Frigg et Nguyen 2020) et (Toon 2012).

Frigg et Nguyen développent leur conception DEKI de la représentation à partir de ce constat. Leur proposition accepte les points mis en évidence par l'analyse fictionnelle, mais fournit en plus une explication du mécanisme par lequel le modèle assigne certaines de ses propriétés à la cible, ce qui constitue le cœur de la relation « extérieure » nécessaire à l'établissement de la représentation. Il y a donc une composante intentionnelle dans leur théorie, mais la représentation ne s'y réduit pas. Les principes fondamentaux de la conception DEKI sont les suivants (les détails se trouvent dans le chapitre 8 de leur (Frigg et Nguyen 2020)). On dira que M est une représentation épistémique indirecte d'un système

cible T si et seulement si : M dénote T , M exemplifie certaines propriétés, une clé K associe ces propriétés du modèle à des propriétés qui sont imputées à T . Dénotation, Exemplification, Key (clé) et Imputation : DEKI. Un bon exemple pour illustrer ces quatre étapes est celui de la carte. Tout d'abord, la carte dénote un certain territoire, et la carte exemplifie certaines propriétés : taches de couleur, distance entre des points, noms imprimés. Une clé permet de traduire ces propriétés : vert pour la forêt, bleu pour les mers, par exemple, mais aussi une échelle pour calculer les distances. Dans ce cas, la clé est composée de la légende et de l'échelle, plus quelques règles implicites sur l'association entre un nom de lieu imprimé et le point coloré le plus proche, par exemple. Enfin, ces propriétés sont imputées au territoire : on déduit de l'observation de la carte qu'une forêt se trouve au Nord de telle ou telle ville, etc.

Toon, quant à lui, part d'une distinction de vocabulaire. Plutôt que de chercher à définir la représentation de manière générale, il s'intéresse spécifiquement à la *model-representation*. Si la simple stipulation est suffisante pour la représentation (exemple de la fourchette topographique), elle ne l'est pas pour la *model-representation* :

in order to be a model-representation of some object, a model must not only denote that object, there also must be an understanding amongst those who use the model that various imaginings are prescribed that depend upon the features of the model. (Toon 2012, p. 64)

Comme pour la théorie DEKI, la dénotation n'est que la première étape de la représentation. Ce qui suit est pour Toon analogue à la notion de jeu autorisé chez Walton : c'est la communauté des utilisateurs du modèle qui détermine quelles parties du modèle servent de support à

ce qui est imaginé à propos de la cible. Dans ce cadre, Toon se montre conscient de l'importance de la question du réalisme :

it would be desirable if our theory of model-representation provided us with a framework in which to address these questions about realism. The theory of model-representation I have proposed does provide such a framework, but this framework differs from that commonly thought to be involved in modelling. (Toon 2012, p. 66)

Toon avance donc le fait que sa conception permet de poser la question du réalisme, mais, sa théorie étant basée sur la notion waltonienne de *make-believe*, il se trouve dans la même situation que Frigg et Nguyen quand il s'agit de formuler la relation extérieure entre le modèle et sa cible.

The make-believe account does not commit us to claiming that models are works of fiction. Instead, it claims only that models are fiction in Walton's sense. What makes a model fiction in this sense is that scientists using the model understand that they are to imagine everything that the model says, simply because it says it. Some of the things a model asks us to imagine may be asserted to be true. Some may be asserted to be merely approximately true, or good enough for the task in hand. Other parts of the model may not be claimed to be even approximately true. The mixture of these elements will be different from case to case. In each case, however, those using the model know that they are to imagine what the model says, whether it is asserted or not. This is the function of scientific models : they prescribe a web of

imaginings which the scientist can then go on to explore. When scientists model a real system, their model prescribes imaginings about that system. This is what makes the model a representation of the system (though not necessarily a good one). (Toon 2012, p. 75)

Il reconnaît donc la nécessité d'une entente sur les règles de compréhension du modèle, analogue du jeu autorisé de Walton : c'est un aspect épistémique irréductible de la représentation. Mais le problème n'est pas pour autant résolu. S'il est admis que ce que le modèle engage à imaginer peut très bien être vrai de la cible, rien ne permet de construire une méthode de justification, ce qui est l'enjeu réaliste central. Prenant l'exemple d'une masse oscillant à l'extrémité d'un ressort, Toon explique que le modèle nous amène à imaginer que la masse oscille avec une période $T = 2\pi\sqrt{m/k}$ (c'est une vérité fictionnelle déduite des principes de génération). Ensuite, à propos de la précision de ce modèle :

The model is accurate in its prediction if the bob does in fact oscillate with period $T = 2\pi\sqrt{m/k}$. (Toon 2012, p. 67)

Puis, plus loin :

We make an inference from the model to the system when we take what is fictional in the model to be true of the system (or, perhaps, approximately true). For example, after some calculations we discovered that it is fictional in our model that the bob's period of oscillation is $T = 2\pi\sqrt{m/k}$. If we think our model is accurate in this respect, then we will infer that this is not only fictional, but also true of the spring itself. (Toon 2012, p. 67)

La précision du modèle dans ses prédictions est donc un préliminaire à l'inférence à la cible. Il est important de remarquer que rien jusqu'ici ne justifie cette inférence. Un scientifique déduit des vérités fictionnelles qui, une fois interprétées en termes attribués à la cible, se révèlent adéquates dans leurs prédictions. De cela, le scientifique infère que ce qui est vrai du modèle est vrai de la cible.

Je prétends que cette description de l'inférence à la cible n'est pas suffisante pour garantir la possibilité d'une défense réaliste. La première partie de la citation établit le critère d'adéquation empirique. À partir de la vérité fictionnelle concernant la période d'oscillation, on peut prédire certaines valeurs empiriques et comparer à ce qui sera mesuré en conditions réelles. Dans un second temps, on transfère la vérité fictionnelle au système cible : étant donné que l'équation est fictionnelle et empiriquement adéquate, on infère qu'elle vaut aussi pour les cas réels. Si cela peut permettre de construire une théorie de la représentation (cette assignation est semblable, dans une version moins détaillée, à la proposition DEKI de Frigg et Nguyen), cela reste une impasse pour le réaliste : l'inférence se fera pour autant que l'on fasse confiance dans la précision du modèle, et donc cette adéquation empirique ne peut pas servir de base à l'inférence, sous peine d'opérer un raisonnement circulaire. Le réaliste se trouve aussi démuné qu'au départ : il n'a que l'adéquation empirique comme critère de sélection des modèles, et toute inférence basée sur celle-ci ne peut avoir de justification autonome. Il n'y a donc pas de réponse satisfaisante à la thèse (Épi) : certains modèles ont un succès empirique, rien de plus ne peut être déduit de ce succès, et certainement pas leur vérité. C'est précisément là que se trouve la limite de l'utilité des ressources représentationnelles pour le réalisme : le point de bifurcation est atteint.

5.3 Fiction et applicabilité

L'approche de Toon est cependant utile sur un point. Le passage de la page 75 de son livre que j'ai cité plus haut (« The Make-Believe account does not... ») met en évidence une exigence sémantique dans l'utilisation des modèles : un modèle modélise un certain phénomène et le représente en prescrivant que certaines choses doivent être imaginées à son sujet. Dans l'approche du *make-believe*, il faut se souvenir que le modèle fonctionne comme un *prop*, c'est-à-dire un support à ces actes imaginatifs, mais la nature de ces supports sont hérités de l'intention représentative elle-même. C'est parce que le scientifique vise à modéliser un phénomène donné qu'il construit un jeu de faire-semblant qui exploite des *props* bien précis et pensés pour remplir ce rôle. Ce n'est pas par accident que les sphères parfaites du modèle du système Solaire sont appelées planètes, mais très exactement parce qu'elles se donnent les planètes comme système cible. Cela peut paraître trivial, mais c'est en réalité fondamental : les entités fictionnelles mobilisées par le modèle obtiennent une interprétation interne en vertu de leur fonction représentative externe. Ce n'est évidemment pas assez pour fonder une position réaliste, mais cela permet tout de même d'offrir une expression à la thèse (Sém) : même fictionnels, les modèles ont vocation à être interprétés littéralement, et cela participe à leur pouvoir de représentation.

Je dis qu'une position réaliste ne peut être atteinte en se servant uniquement de cette remarque car le problème central subsiste : si certains éléments du modèle peuvent être interprétés littéralement, d'autres ne le peuvent pas. C'est l'exemple de la statue de bronze : elle peut nous apprendre que Napoléon a tels traits, mais pas qu'il est fait de bronze. Le modèle newtonien du système solaire peut nous renseigner sur des

trajectoires afin de prédire les dates des prochaines éclipses, mais il ne peut nous renseigner sur la répartition de la masse dans les planètes. Tout ce que le *make-believe* peut dire sur ce point, c'est qu'il serait contre les règles (c'est-à-dire contre la manière dont la communauté comprend majoritairement les règles) de jouer ainsi. Ce serait une mauvaise utilisation du modèle. Or, une position réaliste doit pouvoir présenter un mode de justification, et donc faire la part des choses entre les vérités fictionnelles exportables à l'extérieur, et celles qui ne le sont pas. La position de Toon nous fait cependant progresser puisqu'il établit comment le modèle vise sémantiquement sa cible.

Dans la suite, cette notion sera centrale : l'applicabilité. On dira qu'un modèle est applicable à sa cible s'il en existe une interprétation en termes de la cible (Ruyant 2020).

La notion d'applicabilité pose une première restriction à la liberté heuristique que procure l'approche fictionnelle. La génération du jeu de faire-semblant doit se conformer à cette règle : il permet d'imaginer des choses à propos de la cible parce que les termes qui sont mobilisés ont un sens extérieur à la fiction, sens qui est attaché au système que l'on vise à modéliser. Au niveau de la représentation, la simple stipulation n'est plus suffisante, puisqu'il faut en plus que la règle stipulée soit interprétable en termes de la cible.

La nécessité d'attacher une interprétation extérieure aux modèles scientifiques envisagés comme des fictions (une nécessité réaliste) pose donc une première condition sur la conception des modèles : il est nécessaire qu'ils soient applicables. Si cela n'est pas d'une grande aide dans la méthodologie scientifique (les modèles sont en général construits avec une situation concrète en tête, même très éloignée, et donc avec une interprétation déjà disponible), cela lève une partie de la menace d'ob-

tenir une relation modèle-cible arbitraire ou triviale. C'est un premier pas dans la construction d'un mécanisme de justification indépendant des conditions dans lesquelles le modèle a été pensé et créé : quand bien même il faudrait admettre les conditions épistémiques qui ont permis l'émergence du modèle, notamment sa dépendance aux normes méthodologiques en vigueur, l'applicabilité du modèle est quant à elle évaluable a posteriori.

En particulier, si l'on revient à la notion de ressemblance discutée dans le chapitre précédent, une des critiques était que cette relation est trop facilement obtenue : tout peut ressembler à n'importe quoi selon un aspect particulier. Si on impose en plus l'applicabilité, on obtient une restriction qui n'est pas purement intentionnelle. Frigg et Nguyen donnent l'exemple du modèle de l'atome de Rutherford qui serait utilisé pour représenter un palet de hockey sur la glace (Frigg et Nguyen 2020, p. 101) : il est possible de stipuler que le terme « électron » du modèle a pour référence le palet et que « noyau » réfère à la glace. On en déduirait fautivement que le palet de hockey est chargé négativement. Une pure explication en termes de *make-believe* conclurait à une mauvaise utilisation de la fiction : il ne fait pas partie du jeu autorisé d'imaginer des choses à propos du jeu de hockey quand nous sommes confrontés à une description du modèle de Rutherford. Cette explication fait référence aux normes épistémiques. En revanche, le modèle de Rutherford n'est pas applicable à cette cible : la charge de l'électron n'est pas interprétable en utilisant les termes du hockey. Il y a donc des vérités fictionnelles (internes à la fiction de l'atome de Rutherford) qui ne sont pas interprétables, et qui font que le modèle n'est pas applicable dans ce cas.

L'applicabilité est une notion qui possède différents degrés. Par exemple, le modèle de l'atome de Bohr est construit en analogie avec le

système planétaire. Une partie des éléments qui le composent sont transférables : il y a un sens à parler de trajectoires des électrons/planètes, à mettre en parallèle gravitation newtonienne et force de coulomb (ne serait-ce que via leur dépendance en $1/r^2$). Par contre, l'existence de deux charges, positive et négative, dans le modèle de l'atome n'est pas interprétable en des termes astronomiques et gravitationnels. On peut donc dire que le modèle de Bohr est plus applicable à l'astronomie du système solaire que le modèle de Rutherford ne l'est au hockey.

6 Fiction, idéalisation et modèles scientifiques

Parmi les distorsions et déviations par rapport à la réalité que les modèles peuvent contenir, la notion d'idéalisation joue un grand rôle. Il n'existe pas de définition largement acceptée des idéalisation, et la frontière avec d'autres notions comme les abstractions, les approximations ou les simplifications n'est jamais clairement marquée. Par exemple, l'hypothèse de sphéricité des planètes dans le modèle du système solaire est une simplification (la sphère remplace une forme beaucoup plus complexe à décrire, notamment du point de vue mathématique), une approximation (les planètes sont néanmoins presque des sphères) et une abstraction (de l'observation de plusieurs planètes, on extrait l'essentiel de l'information qui concerne leur forme presque sphérique sous la forme d'une sphère parfaite).

Quand on lit ou entend que les modèles représentent des situations idéalisées ou des objets idéalisés, c'est à ce genre de procédures que l'on fait allusion. J'utiliserai le terme « idéalisation » pour y faire référence

sans distinguer quand ce n'est pas nécessaire ces différentes procédures. La notion d'idéalisation rassemble ce qui constitue leur point commun et explique l'essentiel de leur fonction⁷ : c'est une hypothèse pratique qui vise à simplifier un problème tout en capturant la partie importante de ce qui le caractérise.

L'idéalisation a tout d'abord une visée pratique car elle permet de plus facilement exécuter des calculs (c'est-à-dire : à découvrir de nouvelles vérités fictionnelles dans les modèles mathématisés), de manipuler le modèle et de se représenter mentalement la situation. Le lien avec le *comme si* de Vaihinger est évident⁸ : idéaliser, c'est faire *comme si* la situation était simple en sachant qu'elle est plus compliquée, mais en comptant sur le fait que cette divergence préservera l'utilité des résultats obtenus. Le parallèle tient aussi à l'aspect pragmatique de la démarche : la simplification vise un but pratique, quitte à abandonner une partie de la précision de la description. Cet équilibre entre précision et calculabilité se retrouve dans de nombreux secteurs de l'activité scientifique. Par exemple, lors de calculs numériques, il est souvent nécessaire de poser quelques hypothèses qui permettent de limiter le temps de calcul.

Une des questions largement traitée par les philosophes des sciences est la suivante : ces processus fictionnels d'idéalisation omniprésents dans la science sont-ils irréductibles ? Y a-t-il une procédure inverse, une « dé-idéalisation » possible ? C'est une question importante pour le réalisme car l'objet principal de l'optimiste épistémique est qu'il est possible de tirer une forme ou l'autre de vérité d'une adéquation em-

7. En ce sens, l'utilisation du concept d'idéalisation est une idéalisation...

8. Plusieurs entrées de la typologie des fictions qu'il propose peuvent convenir à cette description. Par exemple, on y trouve le chapitre XXVI intitulé « La méthode de correction des écarts volontaires ou méthode des erreurs antithétiques », ou encore le chapitre XI intitulé « La méthode de généralisation abstraite ».

pirique et d'une puissance prédictive des modèles et théories. Or, l'approche fictionnelle met au centre l'idée que ce sont justement des éléments qui figurent dans un jeu de faire-semblant qui permettent d'atteindre une prédictibilité expérimentale. D'où l'introduction d'une possible dé-idéalisation, et de la fiction considérée comme une simple béquille heuristique provisoire.

Par exemple, on pourrait imaginer que l'hypothèse des planètes sphériques n'est que le premier terme d'un développement dont les ordres supérieurs approchent de mieux en mieux la forme réelle des planètes. Dans ce cas, dé-idéaliser reviendrait à faire les calculs en prenant en compte plus de termes, et à se rapprocher d'une description exacte de la situation. C'est une procédure standard lorsqu'aucune solution analytique à une équation n'existe, et que l'on ne peut qu'approcher (par exemple numériquement) cette solution. Il y a bien évidemment des considérations de convergence qui interviennent dans ce genre de problème, mais ce n'est pas important pour le concept général de l'idéalisation : il y a des situations où le modèle semble en effet contenir des approximations que l'on peut compenser ou réduire. C'est un aspect important du savoir-faire scientifique de trouver le meilleur compromis entre simplicité et précision.

Une chose importante à remarquer est que le processus d'idéalisation n'est possible que si un système cible a été spécifié. En effet, une approximation porte sur une propriété attribuée à un système, une idéalisation est toujours une idéalisation de quelque chose. C'est cohérent avec la remarque de la section précédente à propos de l'aspect sémantique de la modélisation, mais cela reste valable dans le cadre d'une représentation indirecte. Si le modèle est considéré comme un système à part, intermédiaire entre une description et la cible, il peut être conçu comme un

système idéalisé. Dans l'approche directe, le modèle est lui-même une idéalisation de la cible.

L'existence d'hypothèses idéalisées (ou idéalisesatrices) ne pouvant être dé-idéalisées est reconnue par beaucoup, et cela réfute donc l'idée que toute fiction pourrait être dépourvue de ses éléments fictionnels pour se rapprocher, en termes d'une similitude représentationnelle, de sa cible. De tels exemples très discutés sont ceux des atomes de silogène (Winsberg 2006a; Winsberg 2009) et de la viscosité artificielle (Winsberg 2006b). Ces éléments sont introduits dans des modélisations informatiques et n'ont aucune contrepartie réelle. Ce sont même des cas qui contreviennent aux lois physiques les plus élémentaires. Les atomes de silogène ne sont pas sur la tableau de Mendeleïev : ce sont des entités fictionnelles qui ont certaines propriétés de l'hydrogène et certaines propriétés du silicium. La description de leur interaction permet de reproduire des résultats sur la manière dont des nanomatériaux se fracturent sous certaines contraintes :

nanomechanics models fractures in silicon by describing the point of fracture using quantum mechanics and the region immediately surrounding it using classical molecular dynamics. To bring together the descriptions of the two regions, the boundary between them is treated as if it contained "silogen" atoms, which have some properties of silicon and some of hydrogen. Silogen atoms are fictions, Winsberg argues, since they are not offered as even a "good enough" description of the atoms at the boundary ; instead, they are used so that the overall model will work. This overall model, however, is nonfiction. (Toon 2010, p. 286)

Dans ce cas, le modèle décrit le comportement d'atomes inexistantes et impossibles, mais fournit des prédictions efficaces et vérifiées.

Je suis tout de même en désaccord avec la dernière phrase (il ne s'agit pas de la position de Toon, mais de sa description de celle de Winsberg) à propos du caractère non fictionnel du modèle pris dans sa globalité. Du point de vue du *make-believe*, il s'agit effectivement d'une fiction puisque le modèle des atomes de silogène nous amène à imaginer certains processus, même si nous ne croyons pas qu'ils aient une contrepartie dans le système cible (c'est-à-dire dans les nanomatériaux dont on étudie la résistance). La réticence à qualifier ce genre de modèle de fiction chez Winsberg tient à une critique habituelle qui porte plus sur le terme de fiction que sur la notion de *comme si* ou de *make-believe*. Comme je l'ai montré plus haut, les concepts de fiction que l'on trouve chez Vaihinger ou Walton se distinguent de l'acceptation courante par leur rapport à la vérité. Qualifier un modèle scientifique de fiction peut par conséquent avoir une connotation relativiste ou anti-scientifique, voire irrationnaliste.⁹

La notion d'idéalisation amène tout de même à préciser la manière dont ces distorsions volontaires sont introduites dans le modèle. Que les modèles soient considérés comme des caricatures (Chakravartty 2001) ou des analogies (Hesse 1967) de leur cible, cela pose la question de ce sur quoi porte la caricature ou sur ce qui soutient l'analogie. Mary Hesse offre à ce sujet une distinction intéressante.

Un modèle conçu comme une analogie de sa cible entraîne une parti-

9. Ces critiques, parmi d'autres, sont abordées dans (Frigg et Nguyen 2021). Je dois avouer qu'au cours de cette thèse, je me suis progressivement détaché de l'utilisation du terme de fiction, mais je suis d'accord avec Frigg et Nguyen quand ils affirment qu'il est possible de l'utiliser sans risque si le contexte est clairement défini. La prise de conscience des connotations potentielles peut d'ailleurs avoir un effet bénéfique, en poussant à plus de précautions et de précisions lorsque le sujet est abordé.

tion des propriétés du modèle dépendant de leur relation supposée avec le système cible. L'exemple que prend Mary Hesse est celui du modèle de gaz en « boules de billard ». Pour se représenter la manière dont les molécules d'un gaz peuvent se déplacer et rebondir les unes sur les autres, on utilise l'analogie avec le jeu de billard, dont les boules sont animées de mouvements qui mènent à des collisions élastiques. L'utilisation de l'analogie ne sous-entend pas que toutes les propriétés des boules de billard sont applicables aux molécules d'un gaz :

When we take a collection of billiard balls in random motion as a model for a gas, we are not asserting that billiard balls are in all respects like gas particles, for billiard balls are red or white, and hard and shiny, and we are not intending to suggest that gas molecules have these properties. We are in fact saying that gas molecules are analogous to billiard balls, and the relation of analogy means that there are some properties of billiard balls which are not found in molecules. (Hesse 1967, p. 8)

Ces propriétés qui font partie du modèle sans faire partie de la cible, elle les baptise analogies négatives (*negative analogies*). Les propriétés que l'on souhaite attribuer à la cible sont logiquement appelées des analogies positives (*positive analogies*), et celles dont on ne sait si elles peuvent s'appliquer sont des analogies neutres (*neutral analogies*).

La remarque importante que fait Mary Hesse est que les analogies positives et les analogies négatives participent à établir le modèle, mais que les analogies neutres confèrent au modèle un pouvoir prédictif supplémentaire :

these are the interesting properties, because, as I shall argue,

they allow us to make new predictions. (Hesse 1967, p. 8)

C'est parce que l'utilisateur d'un modèle ne sait pas si certaines propriétés de ce modèle sont applicables à la cible qu'il peut étudier jusqu'où l'analogie tient la route, et cela peut amener à de nouvelles découvertes à propos de la cible.

En ce sens, une analogie peut être plus ou moins heuristiquement porteuse. Une analogie peut se révéler fructueuse en nouvelles découvertes, ou au contraire montrer très rapidement ses limites. C'est un aspect primordial car cela offre un critère supplémentaire d'évaluation des fictions scientifiques : elles n'ont pas qu'un pouvoir empirique prédictif, elles peuvent également permettre l'exploration de propriétés nouvelles qui seront elles-mêmes potentiellement l'objet de prédictions expérimentales. La fiction acquiert donc un statut qui dépasse celui de la représentation de la cible. L'exploration, via l'étude des analogies neutres, se fait dans le modèle (découverte de vérités fictionnelles) mais également à l'extérieur : on peut parier sur l'existence de nouvelles caractéristiques suggérées par l'analogie et mettre en place des tentatives de vérifications expérimentales.

L'exemple que j'ai donné plus haut rentre également dans ce cadre des analogies proposé par Hesse. Les modèles planétaires de l'atome jouent sur un parallèle avec l'astronomie, mais ils ne sont efficaces que parce que l'analogie contient elle-même des éléments qui renforcent le pouvoir heuristique du modèle.

- Dans la partie positive, on rangera la notion de trajectoire (qui, certes, se dilue en mécanique quantique, mais cela concerne moins la valeur directe de l'analogie que ses améliorations possibles) ainsi que la forme générale de l'interaction considérée.

- Dans la partie négative, on trouvera par exemple le fait qu’une étoile produit de la chaleur et pas le noyau atomique, ou qu’il est possible de trouver des formes de vie sur les planètes, pas sur les électrons.
- Les analogies neutres peuvent par exemple consister en la recherche de l’existence possible d’orbites stables dans les deux systèmes. Ces orbites sont avérées dans le cas astronomique et leur impossibilité dans le cas atomique a mené à l’élaboration de nouveaux modèles dont l’importance conceptuelle dans l’histoire de la physique n’est plus à prouver. On peut dire que l’analogie a donc été heuristiquement fructueuse.

Cette capacité de la fiction à porter la recherche en dehors de son cadre strict de définition est capitale pour formuler une position réaliste. Dans le chapitre précédent, j’avais insisté sur l’argument qui prenait la forme d’une inférence à la meilleure explication : le réalisme était la seule philosophie des sciences qui ne faisait pas du succès des théories scientifiques un miracle. C’était particulièrement marquant dans le cas des théories capables de faire des prédictions qui dépassaient le cadre initial pour lequel la théorie était appliquée (en particulier, la théorie ne se bornait pas à reproduire les phénomènes). Ce qui apparaît après l’analyse fictionnelle des modèles, c’est qu’il est tout à fait possible de faire émerger le même mécanisme en considérant que les modèles sont essentiellement des entités fictionnelles, entendues au sens du *make-believe*.

Non seulement la fiction ne se définit pas par rapport à sa fausseté, mais elle peut être le moteur de découvertes nouvelles à propos de la cible, et pas uniquement dans le cadre de ce jeu interne de *make-believe*. Cela offre également un mode de justification extérieur de ces fictions

scientifiques : elles doivent être utiles (conception de Vaihinger), suffisamment riches pour être explorées en interne (conception de Walton), mais elles se doivent également de permettre aux analogies neutres qu'elles contiennent de se développer et de donner lieu à de nouveaux tests expérimentaux. Ce dernier critère d'évaluation n'est pas réductible à l'intention représentative du sujet ou aux normes d'évaluation de la communauté scientifique.

7 Théories, lois et modèles

Avant de terminer ce chapitre dédié à la notion de fiction et à ses applications potentielles au problème épistémologique qui m'occupe dans cette thèse, il convient d'aborder une distinction centrale : celle entre théorie et lois. Plus précisément, il est important de comprendre ce que la conception fictionnelle des modèles peut entraîner comme conséquences sur la formulation de cette distinction, étant donné que modèles, théories et lois sont trois concepts proches et difficiles à définir indépendamment les uns des autres.

En effet, la différence entre théorie et loi est un problème à part entière, dont la solution dépend essentiellement de la commodité des définitions dans un contexte donné. L'abandon progressif des conceptions purement syntaxiques pour des approches sémantiques a brouillé la frontière, et comme je l'ai vu dans le chapitre précédent, les définitions des modèles plus éloignées du sens purement logique initial continuent à questionner le lien avec les théories. Il est par exemple admis que les théories aussi ont une fonction explicative et représentative, et que ces fonctions ne sont donc pas réservées à ce que l'on nomme généralement « modèles ». Que l'on considère une théorie comme un langage formel

servant à faire des déductions, comme une famille de modèles sémantiques ou comme des entités plus complexes et en perpétuel changement suivant l'évolution de la méthodologie scientifique, ce sont toujours les mêmes buts épistémologiques qu'il est nécessaire de comprendre : représentation, explication, compréhension, etc.

Je proposerai dans la suite une distinction entre modèle et théorie qui, je pense, suivra naturellement de ma proposition de défense réaliste à partir des modèles conçus comme des fictions.

La notion de loi est, quant à elle, et c'est bien connu, au cœur de nombreux problèmes philosophiques, épistémologiques et métaphysiques. Dans le contexte de la philosophie des sciences, la notion de loi est souvent remise en question sur des bases empiristes (on insiste sur la construction des lois plutôt que sur leur éventuel statut ontologique fondamental) (Harre 1993) ou utilisée dans des débats qui concernent la causalité (Mackie 1974). Qu'elles soient considérées comme fondamentales, comme des généralisations empiriques, comme des inductions ou encore comme des algorithmes de compression de données expérimentales (Wheeler 2018), leur importance dans la connaissance scientifique est quant à elle largement reconnue.

Le parallèle le plus frappant avec la notion de fiction au cœur de ce chapitre se trouve très certainement chez Nancy Cartwright, dans son fameux ouvrage intitulé « *How the Laws of Physics Lie* » (Cartwright 1983). Les lois de la physique, donc, mentent, et ce n'est pas sans rappeler l'idée que les modèles peuvent eux aussi décrire des situations fictionnelles. Cartwright développe dans ce livre une conception du rapport entre lois et modèles qui me sera très utile pour la suite.

L'idée est généralement que les lois de la physique ont une très grande puissance explicative, mais Cartwright fait remarquer qu'elles ont en

revanche une très faible capacité à décrire des situations concrètes et empiriques. Selon elle, les lois ne décrivent pas la réalité.

Cette remarque peut sembler choquante, car les lois sont souvent considérées comme les éléments principaux de la connaissance scientifique. Pourtant, une distinction très intuitive est opérée par les physiciens eux-mêmes, qui, dans la pratique, font très bien la différence entre des lois théoriques et des lois phénoménologiques. Ces dernières collent aux données bien mieux que les premières : elles sont construites dans ce but. Pourtant, cela n'empêche pas les physiciens d'accorder une importance bien plus grande aux lois théoriques, preuve que la capacité à reproduire des données et des phénomènes n'est pas le seul critère, ni même le critère principal, en physique.

Une loi phénoménologique est généralement construite directement à partir des données. Elle est obtenue en cherchant des corrélations et en les exprimant mathématiquement. On remarquera que telle et telle quantité sont reliées, et on introduira par exemple un certain facteur dont la valeur numérique maximisera une mesure de la proximité entre la fonction obtenue et les données (peu important les détails techniques). Une telle loi est la plupart du temps considérée comme un premier essai en attendant un meilleur résultat qui serait dérivé d'autres lois plus générales et déjà établies, ou comme une aide provisoire qu'il s'agira d'éliminer le plus vite possible.

Par contraste avec les lois phénoménologiques, les lois théoriques constituent le squelette d'une grande partie du travail du physicien, et leur valeur est au sommet du panthéon de la physique. Il n'est pas rare d'entendre parler des équations d'Einstein ou des lois de Newton comme des plus grands résultats de l'histoire des sciences.¹⁰ Pourtant, les cas

10. Selon des physiciens. Comme les plus grands résultats de l'histoire de la physique selon tous les

concrets doivent souvent être décrits avec des modifications très importantes de ces lois, ce qui n'entame en rien leur crédibilité, mais questionne parfois l'expression même de ces lois.

Par exemple, la gravitation newtonienne à proximité de la Terre prend la forme de ce qui est souvent appelé la « loi de la chute des corps ». Quelle est l'expression de cette loi ? Mathématiquement, la réponse est claire, mais la manière dont la loi mobilise des principes physiques ne l'est pas. Un dialogue prend typiquement la forme suivante :

Interlocuteur A : La loi de la chute des corps explique pourquoi et comment les corps tombent vers le sol quand ils sont à proximité de la Terre.

Interlocuteur B : Oui, mais un sac en plastique est soumis à la gravitation, et pourtant il s'éloigne du sol.

A : C'est parce qu'il est soumis à une autre force qui compense la gravitation, par exemple la force exercée par le vent.

B : Il est donc imprécis de dire que la loi spécifie que les corps chutent, sinon nous aurions trouvé un contre-exemple qui démontre sa fausseté.

A : Tu as raison, la loi décrit une certaine tendance à la chute, si rien ne la contrarie.

Ce genre de conversation n'est pas rare dès que l'on cherche à comprendre ce que sont les lois physiques. Assez rapidement, on arrive à un dilemme : d'une part il semble nécessaire d'adjoindre à la loi une certaine tendance (qu'elle soit portée par les corps ou non) à se comporter de telle ou telle manière, et d'autre part, il paraît plus parcimonieux (notamment du point de vue ontologique) d'adjoindre aux lois une clause

autres scientifiques.

de la forme « si rien ne contrarie », « dans des conditions normales » ou encore « toute autre chose étant égale par ailleurs ».

Je ne tiens pas à me prononcer sur la nature ontologique et logique des lois dites « fondamentales », cela m’emmènerait dans des considérations métaphysiques très éloignées du sujet de cette thèse. Je pense d’ailleurs que la conception que je détaillerai au chapitre suivant est relativement neutre dans le débat à propos des lois (j’essaierai en tout cas de spécifier les caractéristiques minimales que doivent porter les lois dans ce cadre de pensée).

Ce qui est important ici est le statut des modèles que ces considérations impliquent. Cartwright a elle-même successivement défendu les deux branches du dilemme, et cela a fortement influencé son analyse de la modélisation.

Dans « The Dappled World » (Cartwright 2008), elle développe une théorie des capacités, qui vise à rendre compte de cet aspect des explications scientifiques. Les capacités sont des caractéristiques des corps physiques qui fondent la tendance à se comporter d’une manière particulière. Elle prend l’exemple de l’aspirine (Cartwright 1994 ; Psillos 2009 ; Morrison 1995), qui a la capacité à réduire le mal de tête, ce qui prend tout aussi bien en compte les cas où le mal de tête disparaîtra effectivement que les cas où l’effet de l’aspirine sera contré par autre chose.¹¹

Plus tôt, dans « How the laws of physics lie ? », elle part du constat que les lois théoriques de la physique, par opposition aux lois phénoménologiques, ne sont pas sans exception, et cela l’amène à considérer les

11. La notion de capacité a été largement discutée et critiquée. Sans entrer dans les détails, il semble que les problèmes ontologiques vis-à-vis des lois ne soient que partiellement résolus par l’introduction des capacités. De plus, si les lois décrivent le comportement des capacités, cela explique qu’elles ne puissent être comprises qu’avec la clause *ceteris paribus*, mais cela rend par là même la notion de capacité superflue, du moins si on la compare à ce que Cartwright propose plus tôt dans « How the Laws of Physics Lie » à propos des modèles.

lois *ceteris paribus* :

Many phenomena which have perfectly good scientific explanations are not covered by any laws. No true laws, that is. They are at best covered by *ceteris paribus* generalizations – generalizations that hold only under special conditions, usually ideal conditions. The literal translation is “other things being equal” ; but it would be more apt to read “*ceteris paribus*” as “other things being right.” (Cartwright 1983, p. 45)

Ces lois, prises sans la clause *ceteris paribus* sont fausses, puisqu’elles ne décrivent que des situations sous des conditions idéales.

Par ailleurs, ces lois figurent dans nos explications des phénomènes et permettent d’organiser les observations. Les lois phénoménologiques sont vraies au sens où elles sont strictement soutenues par les résultats d’expérience, mais elles ne sont pas explicatives :

Most scientific explanations use *ceteris paribus* laws. These laws, read literally as descriptive statements, are false, not only false but deemed false even in the context of use. This is no surprise : we want laws that unify ; but what happens may well be varied and diverse. We are lucky that we can organize phenomena at all. There is no reason to think that the principles that best organize will be true, nor that the principles that are true will organize much. (Cartwright 1983, p. 52-53)

Les lois *ceteris paribus* permettent d’unifier des phénomènes, mais cela ne peut se faire qu’au prix de leur capacité à constituer des descriptions littérales des phénomènes. Une explication nécessite un niveau de généralité qui ne peut correspondre à la variabilité des observations.

C'est justement le rôle de la clause *ceteris paribus* de mettre de côté cette variabilité pour atteindre la généralité et l'unification explicative.

Cette séparation entre vérité et explication est un point d'entrée pour une conception à base de modèles, mais elle est en rupture avec les théories classiques de l'explication, et en particulier avec l'approche déductive-nomologique à la Hempel, qui requiert la vérité des prémisses pour déduire la vérité des conclusions.¹²

Les lois qui figurent dans les explications des phénomènes ne peuvent donc pas être considérées comme des descriptions d'états de fait. Les généralisations *ceteris paribus* sont vraies, selon Cartwright, dans des conditions idéales, et c'est là que la notion de modèle est utile : les lois théoriques ne décrivent pas des états de faits, mais des objets fictionnels dans des situations idéalisées. Cartwright utilise d'ailleurs le parallèle avec la fiction :

A model is a work of fiction. Some properties ascribed to objects in the model will be genuine properties of the objects modelled, but others will be merely properties of convenience. (Cartwright 1983, p. 153)

Le lien entre modèle et cible repose donc, selon Cartwright, sur l'attribution à la cible de certaines propriétés du modèle. Le modèle est comparé à une fiction car il possède également des propriétés qui ne sont pas applicables à la cible et qui n'ont pas vocation à l'être. Cet aspect est tout

12. Cela peut se combiner à la réticence empiriste à expliquer par la causalité, cette dernière étant tout aussi douteuse que les lois fondamentales, dans une perspective humienne. Cartwright défend les explications causales basées sur les entités, pas sur les lois : « What is it about explanation that guarantees truth? I think there is no plausible answer to this question when one law explains another. But when we reason about theoretical entities the situation is different. The reasoning is causal, and to accept the explanation is to admit the cause. There is water in the barrel of my lemon tree, or I have no explanation for its ailment, and if there are no electrons in the cloud chamber, I do not know why the tracks are there. » (Cartwright 1983, p. 99)

à fait comparable aux conclusions que l'on pouvait tirer de l'analyse de Walton : le cœur de la représentation se situe dans ce qui permet de générer ce jeu de *make-believe*, c'est-à-dire les propriétés du modèle que l'on imagine faire aussi partie de la cible, et l'utilisation adéquate du modèle suppose que l'on sache faire la part des choses en sélectionnant quelles propriétés sont applicable à la cible.

Cartwright explicite le lien entre modèles, idéalizations et lois dès le résumé du livre :

If we attend closely to the manner in which theoretical laws figure in the practice of science, we see that despite their great explanatory power these laws do not describe reality. Instead, fundamental laws describe highly idealized objects in models. (Cartwright 1983)

Les lois fondamentales sont donc des descriptions non pas de la réalité, mais d'objets idéalisés dans les modèles. C'est cela qui justifie la comparaison entre les modèles et la fiction. Il est important de noter que Cartwright utilise l'analogie dans son sens le plus faible : la reconnaissance que les modèles décrivent des situations fictionnelles. En aucun cas elle n'use d'une théorie de la fiction pour traiter du problème de la représentation, ni pour régler d'autres problèmes épistémologiques comme celui de l'explication. Son cheminement pourrait d'ailleurs être décrit dans le sens opposé : sa proposition du *simulacrum*¹³ pour traiter du problème de l'explication aboutit à des conclusions qui présentent un fort parallèle avec les principes qui servaient de point de départ au travail d'inspiration waltonienne de Frigg, par exemple.¹⁴

13. Un simulacre est quelque chose qui a l'apparence d'une autre chose, sans pour autant partager toutes ses propriétés : c'est exactement le cœur de la conception fictionnelle des modèles.

14. Frigg ayant été l'élève de Cartwright, il n'y a rien d'étonnant à cela.

Cette relation particulière entre les lois et les modèles sera capitale pour ma proposition du chapitre suivant. Je pense cependant que la notion de loi au travail dans les modèles doit être précisée, et que c'est en cela que l'approche fictionnelle, une fois formulée dans des termes qui permettent une défense du réalisme, peut être d'une grande utilité.

8 Conclusion

Il apparaît que l'utilisation en épistémologie de la notion de fiction, notamment grâce à une théorie comme celle de Walton, permet d'éclaircir certains aspects de la modélisation. En particulier, selon la théorie du *make-believe*, les scientifiques confrontés à un modèle s'engagent dans un jeu de faire-semblant, c'est-à-dire qu'ils acceptent certains principes de génération, certaines règles du jeu en vigueur dans le modèle, tout en sachant qu'elles n'ont pas forcément de correspondance avec la réalité.

Un physicien qui se sert d'un modèle newtonien du système solaire pour prédire la trajectoire des planètes fait, l'espace de quelques instants, semblant de croire que les planètes sont parfaitement sphériques et homogènes du point de vue de la répartition de la masse. C'est cela qui lui permet d'explorer le modèle et d'en déduire de nouvelles vérités fictionnelles, c'est-à-dire des énoncés qui sont vrais à l'intérieur la fiction.

Le seul avantage de la conception fictionnelle des modèles n'est pas d'expliquer l'usage des modèles, elle permet également d'offrir une réponse au problème de la représentation. Selon Walton, un modèle peut être qualifié de représentation s'il figure comme support dans un jeu de faire-semblant accepté par la communauté, c'est-à-dire que les utilisateurs de ce modèle se sont accordés sur les règles qui régissent le jeu.

Partant de cette définition de la représentation, Frigg et Nguyen développent leur modèle DEKI, qui constitue, selon leurs critères discutés au chapitre précédent, une réponse complète au problème.

J'ai également montré que l'approche fictionnelle permet de formuler une analyse de la notion d'idéalisation, centrale lorsqu'il s'agit de comprendre comment fonctionnent les modèles scientifiques. En effet, les modèles sont réputés mobiliser de nombreuses idéalizations, et ce mécanisme est central dans les débats qui concernent l'établissement de la connaissance scientifique ainsi que la construction des explications que nous tirons des modèles.

Cependant, les conséquences de la conception fictionnelle restent problématiques pour formuler une défense réaliste aux arguments habituels de l'antiréalisme. En particulier, si les modèles contiennent de nombreuses idéalizations, il semble impossible de les interpréter de manière littérale. Un modèle efficace dans ses prédictions ne peut être considéré comme une description d'un état de faits. Sur ce point, autant l'analyse en *comme si* de Vaihinger que le *make-believe* de Walton ou que la relation entre modèles et lois *ceteris-paribus* de Cartwright vont dans le même sens. De plus, la représentation épistémique telle qu'elle est traitée par Toon, Frigg et Nguyen n'offre pas de possibilité de construire une méthode de justification de la validité des modèles qui se libère des notions liées à l'intentionnalité ou aux références aux normes d'usage.

Dans le chapitre suivant, je partirai des constatations établies jusqu'ici pour construire une conception des modèles qui tire parti de la notion de fiction, mais dans un contexte qui puisse faire une place au réalisme.

Chapitre 4

Quel réalisme pour la fiction ? Quelle fiction pour le réalisme ?

1 Incompatibilité apparente de la conception fictionnelle et du réalisme

Le développement du chapitre précédent a fait apparaître plusieurs signes d'incompatibilité entre le réalisme scientifique et les approches fictionnelles des modèles. Ces points de tension concernent tout autant la thèse du réalisme sémantique (Sém) que la thèse épistémologique (Épi) :

- Concernant la thèse sémantique, l'interprétation du contenu des modèles ainsi que leur capacité à être évalués sur base véridique est fortement remise en cause par l'approche en termes fictionnels. En effet, à partir de l'instant où l'on accepte que la production d'un

modèle est un processus fictionnel, cela semble éliminer la possibilité d'une sémantique réaliste. Comparer un modèle à une analogie ou, pourrait-on même dire, à une caricature d'un système physique montre la nécessité de préciser la forme que pourra prendre la thèse sémantique réaliste. Puisque, dans une conception réaliste, il est impensable de défendre une interprétation littérale de modèles envisagés comme des fictions, il est nécessaire de proposer une modification de la thèse (Sém).

- Concernant la thèse épistémologique, les théories de la fiction utilisées dans le traitement du problème de la représentation souffraient de deux difficultés : elles ne permettaient ni d'offrir une explication de leur succès qui dépasse la constatation d'une adéquation empirique, ni de concevoir la relation avec la cible indépendamment de considérations intentionnelles. Un modèle efficace dans ses prédictions ne peut être évalué, dans ce cadre, qu'à hauteur de sa commodité à l'intérieur d'un jeu de *make-believe*, et la relation supposée entre ce modèle et sa cible n'est elle-même compréhensible qu'à l'intérieur d'un tel jeu dont les règles sont partagées par la communauté des utilisateurs de ce modèle. Rien ne permet dans ce mécanisme de penser une correspondance, quelle que soit sa nature, entre le modèle et le système cible qu'il est censé représenter.

La théorie du *make-believe* et ses différentes prolongations, par exemple chez Toon ou dans le modèle DEKI, visent avant tout à résoudre le problème de la représentation, c'est-à-dire à comprendre la fonction représentative des modèles. Ces derniers sont des éléments tout à fait légitimes de la connaissance scientifique, et leur utilisation en tant que

représentations est largement reconnue. Mon but principal dans cette thèse n'est pas de résoudre ce problème – je considère sur ce point le modèle DEKI comme une proposition très intéressante – mais d'examiner en quoi les outils mis en place pour le résoudre peuvent être utiles dans le cadre du débat à propos du réalisme.

L'analogie entre modèles et fictions, si elle apparaît jusqu'ici contre-productive, me semble en réalité extrêmement féconde, et c'est ce que je vais désormais démontrer. Cependant, comme j'ai déjà eu l'occasion de le dire, il faut s'attendre à voir ce cadre de pensée poser des conditions supplémentaires sur la forme de réalisme qu'il est possible de formuler et de justifier. Ces restrictions porteront à la fois sur la thèse sémantique (Sém) et sur la thèse épistémologique (Épi), comme le laissent entrevoir les deux remarques qui précèdent.

Ces limitations ne sont d'ailleurs pas indépendantes. S'il est clair qu'il faut affaiblir la thèse sémantique, puisque la fiction est par essence impossible à interpréter littéralement, cela a une conséquence immédiate sur ce que l'on peut espérer préserver de la thèse épistémologique : certaines parties, processus ou entités, du modèle sont censés avoir une contrepartie dans la cible, et d'autres non. Déduire de l'efficacité du modèle qu'il est partiellement vrai suppose que cette distinction ait été justifiée en amont. Étant donné que ces deux aspects cohabitent dans la fiction, il est nécessaire, pour pouvoir espérer construire une méthode de justification de la validité du modèle, d'expliquer ce qui sépare les parties potentiellement véridiques des parties que l'on sait fausses, et ce mécanisme ne peut être pensé que si l'interprétation du modèle est elle-même robuste. Pour être plus précis, l'utilisateur d'un modèle sait généralement faire la différence entre les analogies positives et les analogies négatives (au sens de Mary Hesse), mais une position réaliste nécessite

que cette distinction puisse être justifiée indépendamment, c'est-à-dire indépendamment de considérations méthodologiques et intentionnelles.

Je vais prendre un exemple afin d'éclaircir ce problème central, ce qui permettra de comprendre plus aisément le développement qui suivra. Le modèle du pendule simple est efficace dans ses prédictions qui concernent toute une gamme de systèmes physiques (la période des petites oscillations des pendules réels, par exemple). La question réaliste est de savoir ce qui peut être déduit de cette adéquation empirique en tant que description véridique des pendules réels. La réponse réaliste naïve est d'expliquer l'efficacité empirique du modèle par la vérité des énoncés qu'il produit. Le modèle génère certaines propositions à propos de la gamme de systèmes cibles qu'il représente, et ces énoncés sont vrais, au sens de la vérité-correspondance. Il n'y a donc rien d'étonnant à ce que le modèle soit efficace, puisqu'il génère des croyances vraies à propos des pendules réels (cette ligne argumentative n'est qu'une instance du *no miracles argument* vu au chapitre 2, section 1.5).

La conception fictionnelle des modèles réfute immédiatement cette ligne d'argument : les modèles contiennent des éléments qui non seulement ne peuvent pas être interprétés véridiquement, mais qui n'ont même pas vocation à l'être.¹ L'approche fictionnelle insiste au contraire sur l'usage non littéral des modèles dans la méthodologie scientifique : c'est exactement la fonction de l'opérateur du *comme si*.

De plus, une tentative de modifier la position réaliste naïve en introduisant une distinction dans les prétentions à la vérité du contenu du modèle mène au problème déjà soulevé à de multiples reprises dans

1. Bien entendu, ce contre-argument antiréaliste n'est pas le fait de la conception fictionnelle des modèles, il lui est bien antérieur. En revanche, la conception fictionnelle met au cœur de son analyse la présence d'éléments qui résistent à ce transfert abrupt de propriétés du modèle à la cible, et c'est tout l'objet du problème.

le chapitre précédent : s'il n'est pas possible de présenter une justification de cette distinction qui ne fasse pas appel à l'intentionnalité ou aux normes d'usage acceptées par la communauté, alors la position à laquelle on aboutit ne mérite plus le qualificatif de réaliste. En effet, cela désamorce l'optimisme épistémique de la thèse (Épi), étant donné qu'il ne sera plus possible d'associer de quelque manière que ce soit une notion de vérité-correspondance aux énoncés générés par le modèle.

La question fondamentale qui subsiste est par conséquent la suivante : est-il possible, si l'on admet que les modèles ne peuvent être interprétés littéralement (au sens de la thèse (Sém) telle que définie dans le chapitre 2), d'apprendre quoi que ce soit de non strictement instrumental grâce aux modèles ?

1.1 Apprendre, expliquer, représenter : un coup d'œil aux perspectives réalistes

Il existe une connivence évidente entre les positions associées à l'instrumentalisme et l'approche fictionnelle. J'ai montré au chapitre précédent que la définition usuelle du « fictionnalisme », telle qu'on la trouve par exemple dans (Fine 1993), est un instrumentalisme qui ne considère les fictions qu'en tant qu'outils qui visent certains buts – faire des prédictions, classer les phénomènes, générer des explications, générer de la compréhension, représenter des situations expérimentales – et qui ne peuvent être évalués qu'à leur efficacité dans cette entreprise. L'utilisation de la fiction dans l'étude de la fonction représentative des modèles a beau ne pas se réduire au fictionnalisme, les conséquences de cette approche semblent mener à la même impossibilité du réalisme : quand bien même les modèles auraient une fonction plus complexe que celle

de simples outils à propos desquels tout serait permis tant que l'adéquation empirique est au rendez-vous, il semble impossible d'accorder à leur succès la moindre parcelle de vérité correspondance.

Bien entendu, réalistes et antiréalistes s'accordent tout à fait sur les buts poursuivis par la science. Considérer que la science est incapable de formuler des explications légitimes des phénomènes ou de générer de la compréhension des processus scientifiques serait un antiréalisme aussi naïf que le réalisme qui considère que l'efficacité prédictive démontre la vérité des énoncés. L'opposition se situe dans la définition des termes utilisés, et plus précisément dans l'espace qui est accordé à la notion de vérité correspondance pour définir ces termes.

Par exemple, la compréhension des phénomènes est un mécanisme qui peut être plus ou moins orienté vers la vérité (voir par exemple (Regt 2017) pour un tour d'horizon de la question de la compréhension scientifique). On peut opposer en ce sens une notion de compréhension qui suppose la vérité de ce qui la soutient à une définition plus instrumentale qui suppose uniquement la capacité à susciter certains états mentaux et capacités chez le sujet compréhensif. Supposons que l'on dise que « le modèle de l'atome de Bohr permet de comprendre l'apparition des raies spectrales de l'hydrogène ». On peut considérer que le modèle ne constitue un support à la compréhension légitime que s'il génère des propositions au moins approximativement vraies. Par contraste, on peut considérer qu'il contribue à la compréhension dans le sens où il permet de se représenter mentalement la situation par un mécanisme simple, ce qui confère un sentiment de maîtrise d'un sujet indépendamment de la vérité.

Quand la question « comment peut-on apprendre des modèles ? » se pose, il faut donc avoir à l'esprit que les points de vue réaliste et

antiréaliste interprètent de différentes manières le type d'apprentissage que permettent les modèles, mais qu'ils s'accordent sur l'existence d'une forme d'apprentissage.

La plupart des vertus accordées aux théories et modèles scientifiques peuvent s'analyser à partir de cette opposition entre réalisme et antiréalisme. Je m'intéresserai en particulier à la notion d'explication dans le chapitre 5, et j'ai par ailleurs déjà abordé la question de la représentation. Ce qui est important de noter est que la conception fictionnelle des modèles, en tant que ressource pour l'analyse d'une certaine méthode scientifique, est indépendante de la perspective réaliste ou antiréaliste. Comparer les modèles à des fictions permet la construction d'un cadre dans lequel certaines questions sont formulées. Que les réponses qui y sont apportées suscitent des perspectives plutôt réalistes ou antiréalistes dépend des arguments qui peuvent être construits dans ce cadre, mais pas du choix initial d'introduire la notion de fiction. C'est aussi pour cette raison que la conception fictionnelle n'est pas un fictionnalisme au sens de (Fine 1993).

La connivence entre fiction et antiréalisme ne semble donc indépassable que si l'on considère que le cadre de pensée interdit la formulation des buts de la science dans des termes réalistes. C'est ce que beaucoup d'auteurs ont fait de la représentation : les conséquences qu'ils tirent d'une théorie comme celle de Walton appliquée aux modèles est que la représentation ne peut être définie comme la position réaliste le souhaiterait, c'est-à-dire indépendamment des considérations intentionnelles et liées aux normes d'usage, bref de ce que j'ai qualifié de notions épistémiques dans le chapitre 2. Je pense qu'en ce qui concerne la représentation, c'est un problème qui persistera, mais cela est moins dû à l'introduction de la notion de fiction ou à la nature des modèles scientifiques

qu'à la question elle-même.

En effet, il me semble que la notion de représentation contient déjà une touche d'intentionnalité irréductible, et que toutes les théories épistémologiques qui s'attaquent à ce concept ne peuvent écartier, sous peine d'aboutir à une définition artificielle et qui n'aurait plus grand chose à voir avec l'idée de représentation. Représenter, c'est toujours utiliser quelque chose en tant que représentation d'autre chose, et s'il est possible de dégager des critères de validité de cette relation qui soient relativement indépendants des valeurs épistémiques en vigueur, il n'en reste pas moins que c'est l'acte représentationnel d'un sujet qui est à la base de la manœuvre. La représentation et la manière de justifier son éventuelle validité sont en revanche deux choses différentes. La recherche de ce qui, dans le modèle, peut servir de base à la justification des propositions générées par le modèle à propos de sa cible est une préoccupation réaliste.

La notion d'explication scientifique est quant à elle le lieu d'une opposition plus franche entre réalistes et antiréalistes, c'est-à-dire entre des conceptions qui établissent la nécessité qu'une explication soit au moins partiellement vraie pour être légitime, et ceux qui rejettent la vérité comme critère de validité. Je traiterai plus précisément de cette problématique dans le chapitre suivant.

Je me méfie des définitions qui contorsionnent l'usage habituel des termes, et je ne vois pas d'intérêt philosophique particulier à sauvegarder à tout prix une dénomination. Proposer une définition de la représentation qui élimine toute trace d'intentionnalité aboutirait soit à une théorie qui n'aurait plus grand chose à voir avec le fait de représenter, soit à une théorie qui ne s'appliquerait finalement qu'à des cas particuliers. C'est exactement ce que le modèle DEKI de Frigg et Nguyen montre : il

y a de l'intentionnalité, notamment dans la partie « dénotation », mais aussi de nombreuses autres ressources qui enrichissent la représentation tout en gardant la généralité du concept – en particulier en ce qui concerne le problème de la démarcation (voir section 3.3 du chapitre 2), c'est-à-dire le fait que le modèle DEKI permet de comprendre dans un même cadre les représentations scientifiques et les représentations du champ artistique.

Chercher à construire une théorie purement réaliste de la représentation me semble donc un projet inutile. En revanche, les problèmes soulevés dans ce cadre, ainsi que les réponses apportées par l'approche fictionnelle, me semblent intéressantes pour aborder les oppositions entre réalisme et antiréalisme, et c'est pourquoi il était nécessaire que je prenne le temps de détailler ce contexte. C'est en ce sens que cette thèse propose une approche fictionnelle du problème du réalisme.

2 Apprendre à propos d'une fiction et apprendre d'une fiction

La notion de représentation épistémique est au cœur de la conception fictionnelle des modèles, tant chez Toon que chez Frigg et Nguyen. Une représentation épistémique permet les inférences à la cible, ce qui est un aspect central des modèles utilisés en sciences, mais cela ne se limite pas au contexte scientifique. La comparaison entre modèles et fiction paraît justement sur le fait que résoudre le problème de la représentation épistémique permettait de comprendre le fonctionnement des modèles, mais aussi de toutes les autres représentations épistémiques, par exemple les cartes, les dessins, les schémas, les modèles réduits et certaines œuvres

figuratives.

Il apparaît donc que la caractéristique principale des représentations épistémiques est qu'elles permettent d'apprendre quelque chose à propos de la cible. L'opposition entre une position réaliste et une position antiréaliste devra, au vu de l'argument que j'ai présenté dans la section précédente, se concentrer sur le type d'information qu'un modèle peut permettre d'obtenir sur la cible. Plus précisément, si réalistes et antiréalistes peuvent s'accorder sur le fait qu'un modèle peut renseigner sur la cible, tout l'enjeu est de caractériser ce qui constitue l'information obtenue, ainsi que le mécanisme qui permet de justifier sa validité. Les représentations épistémiques permettent des inférences à la cible, il s'agit de déterminer si ces inférences ont une valeur épistémologique supplémentaire ou si elles ne se conçoivent que dans un cadre de type empiriste constructiviste (voir la section 1.3 du chapitre 2) : le modèle permet des inférences empiriquement vérifiées à la cible, mais cela ne permet pas d'affirmer que les propositions générées par le modèle à propos de la cible sont vraies.

Il n'est pas étonnant dans ce contexte de constater que les auteurs qui se sont intéressés à la notion de fiction et à son utilisation en épistémologie ont accepté l'idée d'un apprentissage fictionnel. J'ai montré au chapitre précédent les difficultés que la position de Walton posaient à la construction d'une conception réaliste des modèles, mais on trouve cependant chez lui l'idée d'une utilisation de la fiction en vue d'apprendre à propos de la réalité :

it's role in make-believe is in part, however, a means to the end of conveying information and insight about historical events. (Walton 1990, p. 93)

Dans ce passage, Walton analyse la fonction du roman *Le Chant du bourreau*, de Norman Mailer, qui s'est fortement inspiré d'une histoire vraie pour construire son histoire. Le roman, comme toute représentation waltonienne, sert de soutien (*prop*) dans un jeu de faire-semblant, mais cela n'exclut pas la possibilité d'obtenir de l'information à propos d'évènements qui se sont réellement déroulés. Cet apprentissage n'est pas uniquement factuel, son contenu ne concerne pas que des choses mentionnées dans le livre et qui se trouvent avoir eu lieu. Le roman, bien qu'étant une fiction décrivant le comportement de personnages inexistants, permet par exemple d'explorer la psychologie humaine ou le système carcéral américain.²

D'autres exemples de jeux qui ont vocation à renseigner sur des situations réelles se trouvent dans le langage courant. « J'ai une faim de loup » est un cas typique de phrase qui engage l'auditeur à imaginer certaines choses inexacts à propos du locuteur, mais qui a le but de décrire une situation réelle : il a très faim. L'utilisateur de la fiction, ici l'auditeur, s'il utilise correctement les principes de génération, sait qu'il ne doit pas comprendre que celui qui a parlé souhaite chasser son repas en meute, ce serait là un mauvais usage de la fiction.

Une fiction peut donc servir de vecteur d'information à propos du monde réel. Cette observation est cependant insuffisante pour le but que je me suis donné depuis le début de cette thèse, car il manque une explication du mécanisme par lequel la relation entre le modèle et le système cible se construit et se justifie. En ce sens, le type d'apprentissage mis en avant par Walton est à comprendre au sens large : la fiction per-

2. Walton ne prend bien évidemment pas l'exemple du *Chant du bourreau* au hasard. Mailer est rattaché au mouvement du nouveau journalisme (*new journalism*), dont la caractéristique principale est l'application de techniques littéraires au travail journalistique. C'est donc un domaine privilégié pour étudier les limites de la fiction en littérature.

met éventuellement d'obtenir une information pertinente qui porte sur un sujet donné. Aucun mécanisme de vérification ou d'adéquation n'est constructible dans ce contexte, ce qui n'est pas étonnant étant donné que ce n'est pas le but de la théorie de Walton. On a vu que les tentatives de munir la théorie du *make-believe* d'un tel mécanisme ne pouvait totalement s'abstraire d'une référence à l'intentionnalité ou aux normes de la communauté des utilisateurs de la fiction.

En effet, un lecteur de Mailer, s'il possède une compréhension adéquate de la manière dont les romans sont composés et censés être lus, devine que l'histoire imaginée à l'intérieur de cette fiction fait référence aux États-Unis et à une situation bien précise (le contexte social de l'Amérique profonde des années 1970). Dans un second temps, il est possible au lecteur de vérifier un certain nombre d'énoncés, par exemple en allant consulter des livres d'histoire ou des textes légaux afin de s'assurer que telle ou telle procédure décrite dans le livre est réelle. De manière plus générale, un lecteur évaluera la validité d'un portrait psychologique romancé en le comparant à ce qui est su de la psychologie humaine et en cherchant à déterminer le contexte qui mène le personnage à agir comme il le fait.

Je pense qu'il est impossible de construire dans ce cadre une notion d'« apprentissage par la fiction » qui permette un mode de justification plus précis. C'est la principale différence entre les œuvres de fiction et les modèles scientifiques : il est attendu de ces derniers qu'ils formulent des prédictions précises, si possible quantifiables. C'est leur fonction première et l'aspect principal à la lumière duquel leur validité sera évaluée. Par contraste, les œuvres de fiction ont, comme je l'ai déjà mentionné, d'autres fonctions et suscitent d'autres attentes chez l'utilisateur. Pourtant, s'il existe une différence de degré dans la précision des prédictions

et des justifications attendues, l'analogie entre fiction et modèles scientifiques continue sur ce point d'être fructueuse.

On peut également penser au roman de Tolstoï, *Guerre et Paix*, qui contient de nombreux détails exacts sur les guerres napoléoniennes. Cet exemple montre qu'il existe une utilisation de vérités extérieures à des fins diégétiques et qu'un lecteur peut en apprendre davantage sur le monde réel en lisant le roman. Cependant, on peut difficilement parler de connaissance authentique car, en l'absence de connaissances préalables, le lecteur pourrait être tout aussi justifié de croire à l'histoire d'amour de Pierre et Natasha. Dans le cas des modèles scientifiques, en raison des hypothèses fictionnelles qui sont postulées, le même problème se pose.

Il existe cependant, je pense, un autre exemple plus adapté aux préoccupations épistémiques : les fables de La Fontaine. Il illustrera la manière dont le point de vue de la fiction peut aider à résoudre le problème. À première vue, les fables décrivent des animaux anthropomorphes dans des situations imaginaires. Les personnages eux-mêmes ressemblent moins à des humains dans leurs attitudes qu'à des archétypes de certains comportements. Pourtant, le monde fictif dépeint par la fable a pour fonction de fournir des informations sur le comportement humain réel. Je pense que l'analogie entre modèles et fictions est ici particulièrement explicite : des entités idéalisées sont postulées, certaines de leurs propriétés sont destinées à être interprétées littéralement tandis que d'autres ne le sont pas, et le but épistémique final est d'énoncer quelque chose de vrai de l'extérieur de la fiction.

L'analogie semble également suggérer un rôle important attribué aux lois. Qu'est-ce qui fait d'une fable une bonne indication du comportement humain ? Je prétends qu'il s'agit d'une sorte de conformité aux lois qui sont censées régir le comportement humain. En ce sens, la

fable fonctionne comme un modèle : les entités fictives sont intégrées dans un système décrit par certaines lois.

Il me semble que les inférences que l'on tire des modèles scientifiques fonctionnent sur le même mode : d'abord le modèle est exploré, on y apprend des vérités fictionnelles, puis ces vérités fictionnelles sont utilisées pour formuler les inférences à la cible. C'est ce mécanisme de production de propositions, ainsi que la manière de les justifier, qui est le cœur de la recherche menée dans cette thèse. La première étape s'apparente à un apprentissage à propos de la fiction, et la seconde est celle de l'apprentissage à partir de la fiction. C'est en ce sens que je prends l'approche fictionnelle comme un point de départ pour défendre une position réaliste : les modèles fonctionnent dans des jeux de *make-believe* qui portent sur une cible, l'exploration des caractéristiques du modèle s'apparente à la découverte de vérités fictionnelles au sens de Walton. L'enjeu est de pouvoir comprendre la relation entre ces vérités fictionnelles et d'éventuelles vérités externes à la fiction.

Comment un modèle produit-il des propositions qui portent sur la cible ? Que peut-on dire du modèle lorsque ses propositions se vérifient expérimentalement ? Comment la connaissance portée par le modèle peut-elle être justifiée de manière compatible avec les thèses du réalisme ? Ce sont les questions sur lesquelles je me penche dans ce chapitre. En particulier, il est nécessaire de s'intéresser plus précisément à la relation modèle-cible afin de la définir d'une manière qui soit compatible avec les conditions réalistes que j'ai présentées plus tôt. Dans le même temps, cela amènera à préciser la définition de modèle ainsi que d'explicitier le rôle des lois, comme je l'avais annoncé à la fin du chapitre précédent.

3 Modèles et réalisme sémantique : une première caractérisation de la relation modèle-cible

L'aspect des modèles soulevé au point précédent est compatible avec les deux conceptions – directe et indirecte – de la représentation (voir figures 2.1 et 2.2 du chapitre 2), mais il est nécessaire d'en réviser certains points. En effet, le problème plus général du réalisme dans le contexte fictionnel amène à élargir le questionnement au delà de la notion de représentation, mais il devient alors important de caractériser la relation entre le modèle et sa cible. C'est à cela qu'est dédiée cette section. J'y expliquerai également pourquoi il me semble important de considérer les modèles non pas comme des objets, mais comme des mondes.

Ma proposition en trois étapes prend la forme suivante :

1. Des hypothèses génèrent le modèle. Ces hypothèses, c'est-à-dire des propositions prises comme base du modèle, sont analogues aux principes de génération de Walton. Ce sont les règles du jeu qui spécifient les caractéristiques primaires du modèle, les invitations à imaginer une situation donnée. Dans la pratique scientifique, on les trouvera dans des phrases qui commencent généralement par « supposons », « considérons » ou encore « soit ». Dans l'exemple du modèle newtonien du système solaire, l'établissement de la forme de la force, du nombre de planètes, de leur densité de masse et de leur sphéricité constituent les règles du jeu. Ces hypothèses amènent à imaginer une situation où tel nombre de planètes sphériques se trouvent dans telle ou telle configuration. Tout l'intérêt de la conception fictionnelle de la modélisation

est de mettre en lumière la liberté imaginative lors de la création d'un modèle. Un scientifique est libre³ d'imaginer une situation quelconque et de poser des hypothèses qu'il sait fausses – c'est-à-dire sans les attribuer au système cible qu'il souhaite modéliser.

2. Le monde du modèle est exploré. De nouvelles vérités fictionnelles sont déduites des hypothèses de base, c'est-à-dire que des vérités internes au modèle sont découvertes. C'est le calcul qui détermine les trajectoires des sphères parfaites du modèle newtonien, ou qui permet d'obtenir la période d'oscillation du pendule simple. Toutes ces propositions sont vraies dans le cadre du jeu, car elles sont conformes aux règles. Comme pour les œuvres de fictions, le monde du modèle n'est pas complet, il est possible que certaines propositions aient une valeur de vérité indéterminée. Par exemple, la couleur des sphères n'est pas spécifiée dans les règles du jeu newtonien. Pourtant, les planètes réelles ont toutes une couleur, et l'on pourrait reprocher au modèle de passer ce fait sous silence. Bien entendu, c'est parce que la question de la couleur des planètes n'est pas considérée comme pertinente pour effectuer la prédiction des trajectoires. Cet aspect largement mis en avant par l'approche fictionnelle posera des conditions fortes sur la réponse réaliste à apporter.
3. Les propositions du modèle sont interprétées et appliquées à la cible. Le résultat de l'exploration du modèle n'est applicable à la

3. Une restriction possible de cette liberté de la modélisation peut être d'imposer une condition minimale de cohérence logique. Par exemple, il serait problématique d'introduire en même temps deux hypothèses qui entrent en contradiction car le modèle permettrait alors de déduire tout et son contraire. Les vérités fictionnelles seraient alors elles-mêmes contradictoires. Je ne m'étends pas sur ce problème d'ordre logique qui dépasse le cadre de cette thèse. Il est tout de même intéressant de noter que la possibilité d'imaginer un tel système incohérent n'est pas acceptée par tous. Voir par exemple à ce sujet (Berto 2017) et (Kind 2022) ainsi que les références qui s'y trouvent.

cible que s'il est possible de les interpréter en utilisant des termes qui ont une référence dans la cible, et c'est cette condition que j'avais baptisée applicabilité. Dans le cas newtonien, il s'agit, par exemple, de faire correspondre le paramètre des fonctions mathématiques obtenues après calcul au temps et les valeurs des fonctions elles-mêmes aux positions des planètes au temps donné. Ce n'est qu'après cette procédure qu'il est possible de donner un sens empirique à ces propositions du modèle. Bien souvent, cette étape est automatique car le modèle a été bâti avec cette interprétation définie de manière implicite. Par exemple, un physicien qui effectue les calculs de détermination des fonctions dépendant du temps pense déjà en termes de planètes et de trajectoires. C'est pour cette raison que l'applicabilité est une condition nécessaire pour garantir l'aspect sémantique des modèles : ils n'acquièrent leur sens que grâce à leur cible, et c'est en ce sens que le modèle « vise » une cible. Une fois interprétées, les propositions du modèle sont à même de faire des prédictions expérimentalement vérifiables. C'est en quelque sorte la « jointure empirique » entre le modèle et la cible.

Quelques remarques à propos de ces trois étapes sont de mise.

Tout d'abord, il est important de noter que la vérification expérimentale suite à la jointure empirique n'est pas une vérification de propositions générées par le modèle, même prises isolément. C'est tout l'enjeu du réalisme : l'efficacité prédictive et l'adéquation empirique d'un modèle ne sont pas un signe de la véracité du modèle, puisque ce dernier est conçu comme une fiction, comme un jeu de faire-semblant. Le modèle newtonien efficace dans ses prédictions de trajectoires des planètes n'est pas un argument en faveur de la sphéricité des planètes, tout au

plus du fait que cette approximation n'est pas dangereuse pour l'adéquation empirique. D'une manière encore plus claire, l'efficacité d'un modèle utilisant des atomes de silogène ne plaide pas en faveur de l'existence de tels atomes. Étant donné que la notion d'applicabilité est la seule ressource sémantique à disposition, ce qui donne leur sens aux propositions fictionnelles dépend toujours de données intentionnelles – volonté du scientifique de modéliser tel phénomène en utilisant telles et telles variables interprétées de telle manière – ou du contexte dans lequel la recherche se fait – la nature des hypothèses qui seront posées dépend fortement de la connaissance établie au moment de l'élaboration du modèle. En effet, on ne fera pas un modèle de l'atome si le concept d'atome n'existe pas.

L'un des apports de l'approche fictionnelle est à la fois de mettre en lumière cet aspect important de la modélisation – un modèle n'acquiert de sens que dans le cadre d'un jeu de faire-semblant – et de montrer précisément comment cette interprétation s'établit.

Une seconde remarque concerne plus précisément la thèse du réalisme sémantique. En effet, le réalisme exige que les propositions générées par le modèle aient une valeur de vérité, et que ce soit un état de fait, c'est-à-dire la nature de la cible et des processus qui s'y déroulent, qui détermine cette valeur. Pour l'exprimer plus brièvement, ce que le modèle dit de la cible doit être vrai ou faux en fonction de ce qui se passe effectivement dans la cible. Or, s'il est admis que c'est leur conformité aux hypothèses de base qui détermine leur valeur de vérité aux propositions internes du modèle (l'intérieur du modèle constitue en ce sens une sorte de système déductif), la situation est moins claire pour les propositions externes, celles qui portent sur la cible. Ce sont elles et non pas les propositions internes qui sont concernées par la question sémantique,

et c'est donc à ces propositions externes qu'il faut fournir une interprétation déterminée par la cible afin de remplir la condition sémantique réaliste.

Comme je l'ai vu, toutes les propositions fictionnelles n'ont pas le même statut vis-à-vis de la vérité. Certaines ont vocation à être interprétées de manière véridique, et d'autres non. C'était tout l'intérêt de la catégorisation proposée par Mary Hesse : il y a d'une part les analogies positives et d'autre part les analogies négatives. Les premières sont présentes dans le modèle et supposées avoir une contrepartie dans la cible, les dernières ne sont que dans le modèle et n'ont pas vocation à être vraies.

Est-ce là un problème pour la thèse du réalisme sémantique ? Il pourrait sembler à première vue que oui, étant donné que les valeurs de vérité des propositions générées par le modèle ne sont pas déterminées par la cible. Je pense cependant que ce problème est très facile à résoudre à partir d'une observation à la fois méthodologique et logique. L'utilisation correcte d'un modèle suppose de jouer selon les règles du *make-believe* qui ont été établies. Or, les hypothèses idéalisées, les approximations et autres introductions d'éléments fictionnels ont toujours un but premier qui est méthodologique : faciliter les calculs, permettre un raisonnement plus facile, faciliter la compréhension etc. Savoir que certaines hypothèses sont fausses dès la création du modèle n'empêche en rien de les interpréter avec des termes physiques appartenant à la cible. L'exemple des atomes de silogène est ici particulièrement éclairant : certes ce type d'atome n'existe pas, et ne peut exister, dans la nature, mais ils ont tout de même à l'intérieur du modèle le statut d'atomes assimilables à ceux qui se trouvent dans la cible. En somme, ce qui guide l'interprétation de ces entités fictionnelles en tant qu'atomes de silogène est le fait que le

modèle cherche à prédire le comportement de matériaux faits d'atomes. C'est donc bien la cible qui guide l'interprétation, le modèle reste applicable au sens où je l'ai défini plus tôt. Ensuite, la valeur de vérité des énoncés qui concernent ces atomes inexistantes peut tout à fait être attribuée à la nature de la cible : ils n'existent pas, et les propositions qui les font intervenir sont d'une manière ou d'une autre indéterminées ou trivialement fausses.^{4 5}

Du point de vue pratique et méthodologique, les hypothèses idéalisées et fictionnelles sont connues d'avance, car elles sont à la base du projet même d'établir un modèle d'un phénomène donné. Le risque de les interpréter véridiquement, c'est-à-dire de les attribuer à la cible, est par conséquent assez faible, et un scientifique qui le ferait serait très clairement en train de commettre une erreur d'interprétation. Du point de vue purement sémantique, il est simple de préserver l'idée réaliste selon laquelle c'est la cible qui détermine *in fine* la valeur de vérité des énoncés : ils sont simplement faux car ils n'ont pas de contrepartie réelle. Cette conception reproduit parfaitement bien les intuitions habituelles à propos de la modélisation : certaines hypothèses sont fausses, et connues comme telles, mais elles ne révoquent pas pour autant l'intérêt des modèles.

Il n'y a donc pas de danger immédiat pour la thèse du réalisme sé-

4. La thèse sémantique du réalisme s'intéresse plus à la dénotation des termes du modèle qu'à l'établissement de leur vérité. Introduire la possibilité que certaines propositions du modèle soient indéterminées ne représente pas, à mon avis, de danger pour la position réaliste. Même l'interprétation réaliste la plus naïve du succès empirique des modèles ne nécessite pas qu'ils soient conçus comme des systèmes complets.

5. Encore une fois, cette question est le sujet d'une vaste et abondante littérature en logique et en philosophie du langage. L'exemple le plus connu est probablement celui de Bertrand Russell et de l'actuel – et inexistant – roi de France. Le sens à donner aux propositions qui portent sur ce dernier a été longuement discuté depuis maintenant plus d'un siècle. (Russell 1905 ; Strawson 1950 ; Hochberg 1970 ; Cohen 2000 ; Jandrić 2014)

mantique, dont le cœur peut être facilement conservé dans le contexte fictionnel. En revanche, une des conséquences de cette remarque amène le même genre d'objections que celles qui ont été formulées à propos du problème de la représentation : l'utilisateur du modèle doit se conformer aux normes en vigueur pour guider son interprétation des propositions générées par le modèle. En particulier, il est nécessaire de pouvoir faire la différence entre une proposition fautive qui est de l'ordre du jeu de *make-believe* et une proposition fautive qui est susceptible de réfuter une partie du modèle. Faire la différence entre les *analogies positives* et les *analogies négatives* est de l'ordre de l'utilisation adéquate du modèle, c'est-à-dire d'un usage qui se conforme aux règles méthodologiques de la communauté.

Si la non sphéricité des planètes du système solaire n'est pas un argument pour réfuter le modèle planétaire newtonien, son incapacité à bien reproduire les trajectoires (typiquement celle de Mercure) l'est. Peut-on affirmer que la non existence des atomes de silogène réfute entièrement le modèle qui fait appel à eux ? Est-ce que ce modèle serait, de fait, condamné à ne servir que de « fiction utile » sans pouvoir prétendre à plus qu'une utilité pragmatique ? Une conception réaliste doit pouvoir proposer un critère de démarcation : c'est une des conditions imposées par la thèse épistémologique. Si le modèle peut être utilisé pour obtenir des informations véridiques à propos de la cible, il est nécessaire d'écarter les hypothèses que l'on sait fausses avec un critère indépendant, au risque de faire peser sur toute l'entreprise une menace instrumentaliste. En effet, s'il est impossible de tracer une telle frontière, la fausseté se propage potentiellement à tout le modèle. Toutes les propositions seront aussi suspectes d'être ininterprétables en des termes réalistes que les hypothèses connues pour être fausses, et le modèle dans son ensemble ne

pourra prétendre à aucun usage orienté vers la vérité. Fournir ce critère de démarcation est un des objectifs centraux de cette thèse, dans la mesure où cela constitue le principal obstacle à une compréhension réaliste de l'approche fictionnelle des modèles.

Des modèles comme des mondes : hypothèses et description

Le mécanisme de modélisation que je propose est le suivant : les hypothèses de base sont comme les règles du jeu, elles guident les raisonnements à l'intérieur du modèle et permettent de générer de nouvelles vérités fictionnelles. Une fois interprétées en termes de la cible, les propositions fictionnelles sont potentiellement attribuées à la cible.

L'ensemble des propositions fictionnelles dans un modèle peut être baptisé le « monde du modèle » (Frigg 2009, p. 262). D'une part, cela retranscrit l'idée selon laquelle une fois les règles du jeu posées, le contenu de la fiction acquiert une certaine autonomie vis-à-vis du contexte dans lequel il a été bâti, et d'autre part, cela met en exergue la nécessité d'approcher le modèle comme un système qui n'est pas figé, mais qu'il est au contraire possible d'explorer. Envisager une fiction comme un monde permet aussi d'introduire immédiatement la question des points communs entre ce monde idéalisé, contenant des objets aux propriétés parfois impossibles, et le nôtre.

La fiction peut dès lors être considérée comme une description d'un monde, mais cela uniquement une fois que l'interprétation a été fournie. Sur ce point, je suis en désaccord avec Frigg et Nguyen lorsqu'ils introduisent la différence entre *model-system* et *model-description* :

Model-descriptions are the passages in research papers, text-

books, and other scientific publications that describe a model. We here focus on the case in which this description is verbal, but in principle it can take any form (it could also be a schematic drawing or a diagram, for instance). The model-system is the object that the model-description – purportedly – introduces : spherical planets, isolated populations, and so on. (Frigg et Nguyen 2020, p. 106)

La description du modèle est, selon eux, les passages dans les livres ou articles qui décrivent le modèle. Le système-modèle est quant à lui l'objet qui est décrit. La conception fictionnelle explique par le biais du *make-believe* comment de tels objets peuvent être conceptualisés lorsque le système-modèle n'existe pas.

Mon désaccord porte sur le fait que cette distinction minimise le rôle de la description du modèle et introduit une confusion critique entre les hypothèses qui génèrent le modèle et l'interprétation des termes du modèle.

Pour rappel, la notion d'applicabilité qui permet de comprendre comment le modèle est un modèle de sa cible (comme il la « vise ») fait appel à l'interprétation du contenu de la fiction dans des termes qui sont attribués à la cible, ce qui n'implique aucunement que les propriétés du modèle sont imputées à la cible. Dans les livres, les modèles sont effectivement décrits verbalement ou via des schémas, parfois en introduisant des équations. Frigg et Nguyen appellent le système-modèle l'objet qui est décrit, mais cela suppose que la description contienne déjà l'interprétation du modèle. Quand nous lisons dans un livre « supposons que tel ou tel nombre de sphères parfaites se trouvent dans telle disposition et interagissent de telle manière » ou « considérons deux populations isolées et se comportant de telle manière », nous sommes invités à ima-

giner une situation donnée, que nous nommons modèle. Cette situation n'est imaginée de cette façon que parce que le contenu du modèle est préalablement interprété. Par conséquent, deux descriptions différentes engendrent deux modèles différents, et leur applicabilité aux systèmes cibles dépend de la description initiale.

Le problème est que, dans la pratique, plusieurs descriptions différentes peuvent générer le même modèle. Je prends pour exemple le pendule simple.

Dans la plupart des livres de cours, le modèle du pendule simple sera décrit comme un point matériel relié à un fil dont la masse est négligeable (voire nulle) et qui ne subit pas d'extension. Le tout est plongé dans un champ de gravitation, et c'est l'oscillation du point matériel qui est explorée. Typiquement, la situation sera représentée comme :

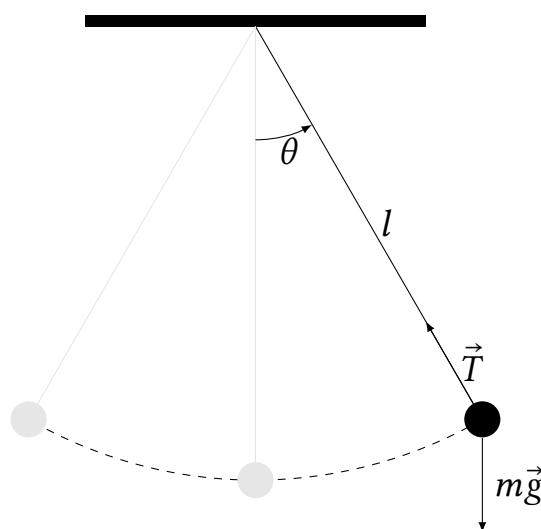


FIGURE 4.1 – Pendule simple : point matériel oscillant à l'extrémité d'un fil

Mais on pourrait fort bien donner une description tout à fait différente. Le pendule simple est l'étude du mouvement d'un point matériel

lié à une surface polie plongée dans un champ de gravitation. (Spindel 2001, p. 100) Ou schématiquement :

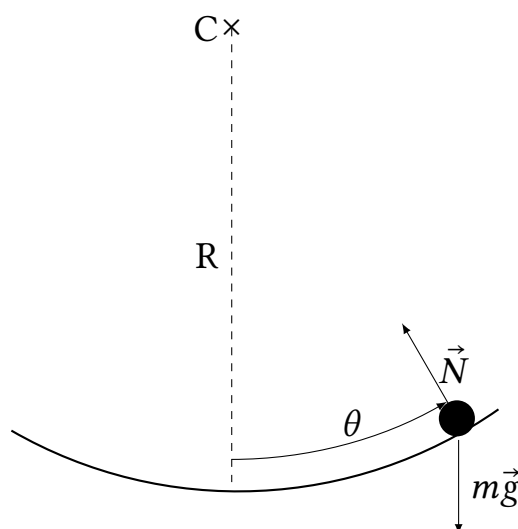


FIGURE 4.2 – Pendule simple : point matériel lié à une courbe polie

Si l'on suit le raisonnement de Frigg et Nguyen, ces deux descriptions différentes engendrent des systèmes-modèles différents et qui pourtant ont la même cible : les pendules réels. Les deux modèles font des prédictions identiques, mais l'interprétation qui leur permet de viser la cible est différente.

Pour la première description, l'interprétation du modèle en termes de la cible associera la longueur du fil dont la masse est nulle au fil réel, le point matériel au poids réel, etc. Pour la seconde, il faudra notamment considérer que la surface polie est un arc de cercle dont le rayon a la même longueur que le fil réel.

Si ces deux systèmes-modèles pointent vers la même cible, il faut expliquer ce qu'ils ont de commun. Intuitivement, et d'une manière qui colle à la façon dont ces descriptions sont utilisées dans la pratique, ces modèles sont identiques parce que leurs interprétations concordent. On

les qualifiera tous les deux de « pendules simples » sans qu'il n'y ait d'ambiguïté dans leur utilisation pour modéliser des pendules.

En revanche, dans l'approche fictionnelle, il y a tout de même un problème qui subsiste : ces différentes descriptions engendrent des mondes différents : dans le premier, des fils parfaitement inextensibles existent sans qu'il ne soit rien dit des surfaces polies, tandis que dans le second, les surfaces qui n'exercent aucun frottement existent, mais pas forcément les fils inextensibles.

Deux descriptions différentes engendrent deux jeux de *make-believe* différents, et ce qui unifie ces deux fictions est l'interprétation qui en est proposée. Dès lors, il semble que l'on ne puisse échapper à l'intentionnalité sémantique des modèles envisagés comme des fictions. En première analyse, le modèle vise sa cible parce que son contenu peut être interprété en des termes qui sont attribués à la cible (applicabilité), mais cela laisse une grande latitude dans le contenu du modèle. Frigg et Nguyen prennent l'exemple du modèle de l'atome utilisé pour représenter un palet de hockey sur la glace (Frigg et Nguyen 2020, p. 101) : même si, dans ce cas précis, l'interprétation semble tirée par les cheveux, rien ne peut la réfuter si l'on ne considère que l'aspect sémantique du modèle. Il est parfaitement applicable puisqu'une interprétation en termes de la cible existe : le palet est représenté par l'électron du modèle, etc.

Pour comprendre les modèles en tant que fictions de manière réaliste, la condition d'applicabilité est donc nécessaire mais non suffisante. L'applicabilité offre une explication de la façon dont le modèle vise sa cible, mais cela reste trop sensible aux normes épistémiques et à l'intentionnalité de la modélisation. Étant donné qu'un modèle est toujours bâti avec une ébauche d'interprétation en tête, il n'est pas étonnant de constater la facilité avec laquelle cette interprétation peut s'adapter à différentes

situations. De manière analogue, il est facile de plaquer une interprétation sur un modèle existant en changeant la dénotation de certains objets pour l'appliquer à d'autres situations. Frigg et Nguyen, dans leur analyse de la représentation, arrivaient à la conclusion qu'il existe une étape de dénotation ou d'imputation irréductiblement intentionnelle.

Il est impossible de proposer une conception fictionnelle des modèles qui décrive le mécanisme de modélisation indépendamment des normes épistémiques, et donc en accord avec les thèses réalistes. Quelle que soit l'approche, il semble que l'introduction du concept de fiction entraîne fatalement l'obligation de considérer l'usage de la fiction, et particulièrement le bon usage au regard des normes épistémiques. Cette introduction de l'intentionnalité se faisait dans le problème de la représentation au niveau même du concept de représentation : représenter, c'est toujours avoir l'intention de représenter. L'analyse ne pouvait donc pas se passer du sujet qui représente. Quant à évaluer la validité d'une représentation, la question ne pouvait se formuler que dans un cadre où le sujet connaît les normes de ce qui fait une représentation acceptable, de ce qui constitue une utilisation acceptable du modèle. Il est donc impossible de caractériser la validité d'une représentation sur base d'une correspondance entre le modèle et la réalité.

Lorsque la question plus générale du réalisme se pose, le problème subsiste : étant donné que les modèles sont des fictions, qu'ils sont des mondes imaginés selon certaines règles, ils n'ont la capacité à être mis en contact avec la réalité qu'après avoir été interprétés en termes de la cible. C'est ce que j'appelle la jointure empirique. Même si l'idée d'obtenir une information véridique – c'est-à-dire qui peut être qualifiée de vraie sur base du succès empirique – peut a priori se formuler sans référence au sujet, c'est l'interprétation qui est sujette à l'intentionnalité.

Les modèles n'ont de sens qu'interprétés, et cette interprétation dépend essentiellement du but que le scientifique se donne lorsqu'il construit le modèle ainsi que ce qu'il imagine à propos de la cible.

Par conséquent, s'il y a une possibilité de conserver la thèse du réalisme sémantique dans le contexte fictionnel, je viens de le montrer, la thèse épistémologique continue à poser problème : si nous voulons interpréter de façon véridique ce que le modèle dit de la réalité, il faut pouvoir justifier l'interprétation du modèle sans référence au sujet.

4 Vers une nouvelle stratégie *divide et impera*

Une autre des grandes leçons de l'approche fictionnelle est qu'il existe une part irréductible d'intentionnalité dès lors que le contenu du modèle est interprété. En première analyse, cela semble mener à une impasse, comme je l'ai montré à la fin de la section précédente. Pourtant, en deuxième analyse, il me semble que l'échec découle d'une confusion tenace : celle entre le contexte de découverte et le contexte de justification.

En effet, je suis tout à fait prêt à accepter l'intentionnalité de l'interprétation des modèles, ainsi que leur caractère relatif à certaines normes lorsqu'ils sont analysés en termes de jeux de *make-believe*. Le principal enjeu réaliste de la thèse épistémologique est de caractériser leur succès sans faire appel à des notions épistémiques (liées à la connaissance disponible) ou intentionnelles. Le traitement de la thèse sémantique de la section précédente s'intègre assez naturellement au contexte de découverte : le contexte épistémique dans lequel le modèle est formulé

influence sa construction, son interprétation et son utilisation. Or, le réaliste est avant tout intéressé par la caractérisation de la justification, c'est-à-dire par l'évaluation de la validité du modèle une fois qu'il a été bâti.

L'approche fictionnelle met en avant la liberté du scientifique d'imaginer des situations idéalisées et irréalistes, pour ne pas dire alambiquées et impossibles. Toute liberté heuristique est a priori acceptable, pour autant qu'elle mène aux vertus habituellement acceptées en sciences : pouvoir explicatif, unification, production de compréhension (*understanding* (Regt 2017)), puissance prédictive, adéquation empirique. En ce sens, les fictions scientifiques sont tout d'abord évaluées selon leur capacité à remplir de tels buts. C'est dans un second temps que se pose la question du sens à accorder à ces vertus épistémologiques.

La fiction se propage-t-elle à tout le modèle ? C'est la question qui préoccupe le plus le réaliste : si l'on accepte que les modèles sont des fictions, les conséquences antiréalistes sont-elles irrémédiables et partout présentes ? C'est ce que les exemples précédents semblent avoir montré.

Deux remarques s'imposent cependant.

La première concerne la liberté heuristique mise en avant par l'approche fictionnelle. Cette liberté intervient lors de la création du modèle, c'est le geste créatif initial qui engendre le monde du modèle. De toute évidence, cette étape fait partie du contexte de découverte, non du contexte de justification. Un scientifique est libre d'imaginer tout ce qui lui plaît, de prétendre que tel ou tel objet idéalisé dans son modèle réfère à telle ou telle entité supposée appartenir au système cible. Lors d'une expérience de pensée, autre type de modèle fictionnel tout aussi important dans la pratique scientifique, une situation est imaginée, des forces introduites ou des objets imaginés sans aucune contrainte prag-

matique ou technique. Dans ce dernier cas, le scientifique entre de plain-pied dans une fiction dont le lien avec de potentiels systèmes cibles peut parfois sembler ténu.

Beaucoup de paramètres peuvent influencer la production des expériences de pensées, des modèles, des entités qui y sont introduites, ainsi que le choix des variables qui seront approximées. Il est certain que ces paramètres seront d'ordre épistémique : connaissance scientifique et théories préalablement disponibles, présupposés métaphysiques, normes méthodologiques, préjugés et biais de tous types. En revanche, le problème du réaliste se pose essentiellement dans un second temps : il consiste à trouver une justification de la validité du modèle indépendante de ces paramètres épistémiques. C'est la raison pour laquelle les réalistes insistent souvent (parfois de manière détournée, parfois de manière explicite) sur la distinction entre contexte de découverte et contexte de justification. Si l'on accepte que beaucoup de choses interviennent dans la création d'un modèle, une fois qu'il est bâti, il acquiert un statut indépendant. Il devient évaluable en tant qu'objet scientifique sur base empirique et théorique. Le cœur du réalisme est à la fois de montrer la pertinence de cette distinction et de fournir la justification épistémologique.

La seconde remarque concerne le contenu des exemples comme ceux du pendule simple. Les éléments qui sont considérés comme idéalisés ou dont les propriétés sont approximées sont les objets décrits par le modèle. Cela se retrouve également dans l'analyse des modèles que proposait Nancy Cartwright, et qui se retrouve dans le passage que j'ai déjà cité au chapitre précédent :

fundamental laws describe highly idealized objects in models. (Cartwright 1983)

Dans le cas d'un modèle comme celui du pendule simple, il paraît évident que les points matériels qui ne subissent aucun frottements n'existent pas dans les systèmes cibles associés. En revanche, les lois qui y figurent et font office de principes de génération n'ont pas le même statut. Elles forment en quelque sorte le squelette du modèle, sur lequel les objets idéalisés peuvent venir se greffer. Ces objets ne sont pas idéalisés parce que les lois sont elles-mêmes des descriptions approximatives de lois réelles, ils sont idéalisés parce qu'ils prennent part à des jeux de *make-believe* décrits par des règles – c'est-à-dire des lois – qui décrivent des situations idéalisées.

Pour le formuler de manière plus compacte : les lois qui génèrent les modèles ne sont pas elles-mêmes des idéalizations, mais elles décrivent des situations idéales.

C'est cela qui explique la distance entre les lois fondamentales et le monde empirique, ainsi que la distinction que Cartwright introduit entre lois fondamentales et lois phénoménologiques : les premières sont vraies des modèles, mais ne décrivent pas directement la réalité, seulement des situations idéalisées éloignées des systèmes réels où de nombreux autres paramètres peuvent interférer. (Cartwright 2008)

Se pose alors la question de la nature d'une stratégie *divide et impera* dans le cadre des modèles. En effet, le réalisme structural (voir section 1.6 du chapitre 2) opérait une division entre objets et structures afin de faire reposer le pari épistémologique et ontologique réaliste sur les dernières, contrairement à la vision généralement acceptée et qui est fortement orientée vers les objets. Il semble que dans les modèles, une distinction naturelle apparaisse entre lois – associées aux principes de génération, dans l'approche *make-believe* fictionnel – et le contenu des modèles – le cœur de la fiction, le monde fictionnel peuplé d'objets aux propriétés

qui ne sont pas toutes attribuables à la cible.

Suivant cette ligne de pensée, une proposition de défense réaliste analogue au réalisme structural serait : il est admis que le contenu des modèles n'est pas une description véridique des systèmes physiques ciblés, les mondes fictionnels sont au mieux heuristiquement porteurs, utiles dans la compréhension qu'ils permettent d'obtenir, dans la manipulation du monde physique, dans le développement de technologies. En revanche, les lois qui génèrent ces modèles peuvent être considérées comme approximativement vraies : ce qui explique le succès de certains modèles est qu'ils mobilisent en tant que principes de génération des lois qui correspondent suffisamment bien aux lois réelles.

Une telle défense est de fait réaliste au sens où je l'ai défini dans le chapitre 2. En effet, elle explique le succès empirique et prédictif du modèle par une correspondance entre le modèle et la cible.

Par contre, une remarque importante s'impose immédiatement. Si ce mécanisme rend compte de la clause *ceteris paribus* qui figure dans l'expression des lois fondamentales – car les lois fondamentales ne s'appliquent qu'aux situations idéalisées des modèles –, cela ne permet pas d'éclaircir le mécanisme de justification dont le réaliste a besoin, bien au contraire. La ligne de défense réaliste semble faire une présupposition assez forte : il existe des lois réelles dont la forme est suffisamment semblable à celles que l'on trouve dans les modèles pour que cela ait du sens d'y chercher une correspondance. Comme je l'ai déjà mentionné, il existe une littérature extensive à propos des problèmes ontologiques et épistémologiques des lois scientifiques, et cela dépasse le sujet de cette thèse. Il est tout de même important de noter que cette hypothèse est d'ordre métaphysique et qu'elle fait intervenir de nouvelles entités au statut très difficile à cerner. En ce sens, les lois ne sont pas des objets

physiques, mais elles réglementent le monde physique. Elles décrivent des régularités qui ne sont jamais pleinement observées, puisqu'elles ne s'appliquent qu'à des situations idéalisées. En tant que lois dites fondamentales, elles sont supposées s'appliquer à des cas qui dépassent ceux pris en compte par les modèles. Il paraît par conséquent douteux que l'efficacité du modèle puisse plaider en leur faveur sur l'ensemble de leur domaine d'application. Par exemple, l'efficacité du modèle du pendule simple ne pourrait que très difficilement être considéré comme une preuve de la véracité des lois newtoniennes dans toute leur généralité.

Et ce ne sont là que quelques uns des nombreux problèmes liés au statut des lois.

Si la recherche d'une formulation réaliste du succès empirique des modèles prend, dans le cadre fictionnel, la direction d'une stratégie de division du domaine des modèles en question, elle ne peut en tout cas pas s'appuyer sur une correspondance supposée avec des lois extérieures, sous peine de soulever davantage de problèmes que d'apporter de solutions.

Une fois de plus, je pense qu'une piste de réponse se trouve dans le point de vue motivé par l'approche fictionnelle, en ce qui concerne notamment le statut dynamique que les modèles acquièrent dans cette optique.

5 Idéalisations et manipulabilité des modèles

La conception fictionnelle met en avant un aspect des modèles qui est souvent passé sous silence : ils doivent être manipulés pour être utiles.

C'est une conséquence de la notion même de jeu de *make-believe* : il existe un support matériel qui soutient les choses imaginées (ce sont les *props*, comme les appelle Walton), mais la découverte de nouvelles vérités fictionnelles nécessite un travail d'exploration. En ce sens, un modèle n'est pas un objet d'étude monolithique, mais un monde à parcourir, une entité dynamique que l'on utilise en la manipulant.

C'est aussi pour cette raison que j'ai adopté la terminologie de Walton en ce qui concerne la construction des modèles : les principes de génération ne servent pas que de squelette fixe au modèle, ils guident les raisonnements, l'exploration, et donc participent à la génération des propositions fictionnelles. On peut donc voir un modèle comme un outil que l'on utilise pour générer des propositions à propos de la cible.

La citation de Ian Hacking à laquelle j'ai déjà fait référence prend un sens nouveau dans le contexte fictionnel :

a model in physics is something you hold in your head rather than your hands. (Hacking 1983, p. 216)

Les modèles, qu'ils soient mathématiques ou non, fonctionnent au cœur d'un jeu d'exploration qui génère des prescriptions à imaginer certaines choses, qui ensuite sont appliquées à la cible du modèle. Manipuler un modèle, avec les mains ou dans sa tête, c'est générer ces propositions, faire cracher au modèle ce qu'il cachait, découvrir des vérités fictionnelles.

Je pense que cette idée est primordiale pour comprendre à la fois le rôle des idéalizations et le statut des lois dans les modèles.

En effet, je soutiens que le rôle principal des idéalizations est de mettre en lumière certaines relations. Sur ce point, je prends le chemin inverse de Cartwright. Là où elle considère que les lois décrivent

le comportement d'objets idéalisés dans les modèles, je mets en avant le rôle déjà idéalisé des lois : elles ne sont pas elles-mêmes idéalisées – ce qui soulève le problème d'une correspondance avec des supposées lois réelles que j'ai mis en avant dans les sections précédentes – mais elles décrivent des situations idéalisées, c'est-à-dire partielles, approximatives, dans lesquelles seuls des objets idéalisés peuvent évoluer. C'est donc moins, dans la perspective que je propose, des objets idéalisés qui sont supposés et imaginés que des objets mis dans de telles situations qu'ils ne peuvent que se comporter de manière idéalisée.

Cela renverse l'analyse que l'on peut faire des modèles une fois qu'ils ont été bâtis et explorés. Par exemple, le modèle planétaire newtonien n'est pas la description de planètes parfaitement sphériques qui interagiraient au moyen d'une loi en $1/r^2$. C'est le monde que l'on décrit une fois une certaine loi posée, loi qui se comporte par hypothèse en $1/r^2$. Cette loi, pour être utile en pratique, doit être accompagnée d'autres hypothèses, dont celle de sphéricité et de densité de masse homogène ou d'assimilation possible des planètes à des points matériels. Ces hypothèses rendent le modèle calculable, c'est-à-dire explorable, et elles en spécifient certaines propriétés, mais c'est la loi qui reste la description centrale de ce monde.

Ce renversement de l'analyse est, il me semble, la façon la plus naturelle de prolonger le raisonnement fictionnel, car il met en avant le rôle des lois, tout comme l'analyse waltonienne du *make-believe* insistait sur le caractère central des principes de génération.

Cette façon d'envisager le modèle rend aussi parfaitement compte de l'aspect dynamique des modèles. Il y a d'une part la possibilité d'explorer le modèle en cherchant à dériver de nouvelles vérités fictionnelles, mais ces dernières ne sont valables qu'à l'intérieur du monde du modèle. Ce

qui opère la jointure empirique, c'est la génération par le modèle de propositions expérimentalement testables, ce qui nécessite dans un second temps d'adjoindre au modèle toute une série d'hypothèses de calculabilité ainsi que des conditions initiales.

En effet, ce qui fait que le modèle newtonien du système solaire est effectivement un modèle de notre système solaire, ce n'est pas uniquement la présence de la loi en $1/r^2$ et la possibilité d'une interprétation des termes du modèle en fonction de la cible. Comme je l'ai affirmé plus tôt, c'est une condition nécessaire, mais non suffisante. Une fois le modèle exploré, les propositions générées sont en quelque sorte dénuées de valeur empirique, du moins jusqu'à ce que l'on incorpore des valeurs initiales qui permettent de générer des valeurs expérimentalement testables, ou qui ont du moins un sens empirique précis.

Le modèle newtonien du système solaire est donc un monde possiblement semblable au nôtre, dans lequel des points matériels interagissent gravitationnellement selon une loi en $1/r^2$. Cette description amène l'utilisateur du modèle à imaginer une certaine situation et lui permet de l'explorer. Le modèle est en quelque sorte incomplet : il est nécessaire de spécifier bien d'autres choses pour le rapprocher de la situation que l'on souhaite effectivement modéliser. Ainsi, il faut y incorporer la valeur des masses en question, leurs distances, leur nombre.

Le modèle est capable de décrire avec la même facilité (sous réserve de calculabilité mathématique, dans ce cas précis) un système stellaire à vingt-huit planètes disposées sur une seule orbite disposées au sommet d'un polygone régulier à vingt-huit faces. Le résultat ne méritera pas le nom de « modèle du système solaire » car la plupart des éléments que contient cette fiction ne sont pas interprétables en termes du système solaire. Par contre, si l'on incorpore huit planètes disposées d'une ma-

nière qui correspond aux connaissances astronomiques disponibles (le nombre huit fait déjà partie de cette connaissance), le résultat sera légitimement qualifié de « modèle du système solaire ». C'est seulement une fois toutes ces hypothèses incorporées à la fiction que le modèle est capable de générer des propositions empiriques, par exemple en permettant de prédire la date d'une éclipse.

Dans ce raisonnement, une chose pourrait cependant soulever une objection : il semble que j'aie utilisé le terme « modèle » pour qualifier deux choses différentes.

D'une part, j'ai appelé « modèle » la fiction décrivant des points matériels qui interagissent avec une force donnée et, d'autre part, j'ai attribué la qualification de « modèle » au résultat obtenu lorsque les hypothèses ont été adjointes afin d'obtenir des propositions testables expérimentalement.

Cet usage n'est pas contradictoire. Il m'amène au contraire à préciser la définition que l'on peut donner d'un modèle en généralisant le principe fondamental qui fait des modèles un élément primordial de la connaissance et de la pratique scientifiques.

Ce point est également une conséquence naturelle de l'approche fictionnelle. Les mondes fictionnels sont, comme on l'a vu, incomplets. Il convient donc de les compléter afin d'opérer la possible jonction avec le monde empirique. Dans l'approche que je propose, les modèles sont comme je l'ai déjà mentionné des entités dynamiques, ils sont perpétuellement engagés dans des jeux de questions-réponses qui dépendent des spécifications qui leur sont imposées, c'est-à-dire de la manière dont on les complète.

Le modèle de gravitation newtonienne le plus général spécifie une force avec une forme donnée. Cela décrit un monde, cela génère une

fiction, un jeu de *make-believe*. Il est ensuite possible de le compléter de plusieurs manières, en jouant sur le nombre de planètes, leurs positions, leurs masses respectives.

Les modèles s'organisent naturellement selon ce principe en une arborescence. Verticalement, cet arbre décrit différents niveaux de généralité, c'est-à-dire différents niveaux de spécifications imposées au modèle. Horizontalement, l'arbre décrit différentes possibilités de spécifications. Chacun des éléments de l'arbre mérite le nom de modèle au sens où ils sont le lieu d'un jeu de *make-believe*. Chacun décrit un monde différent, peuplé d'entités aux propriétés plus ou moins semblables.

Dans cette arborescence, les lois, en tant que principes de génération du *make-believe*, sont au plus haut niveau de généralité. Cela explique leur statut généralement considéré comme fondamental dans la pratique scientifique sans faire peser la charge ontologique sur les lois : le cadre du *make-believe* reste d'application, ces lois sont à la base de la génération des différentes fictions, cela n'implique pas leur existence sous une forme ou sous une autre. Je traiterai spécifiquement ce problème dans la section suivante.

L'arborescence des modèles peut se représenter comme suit :

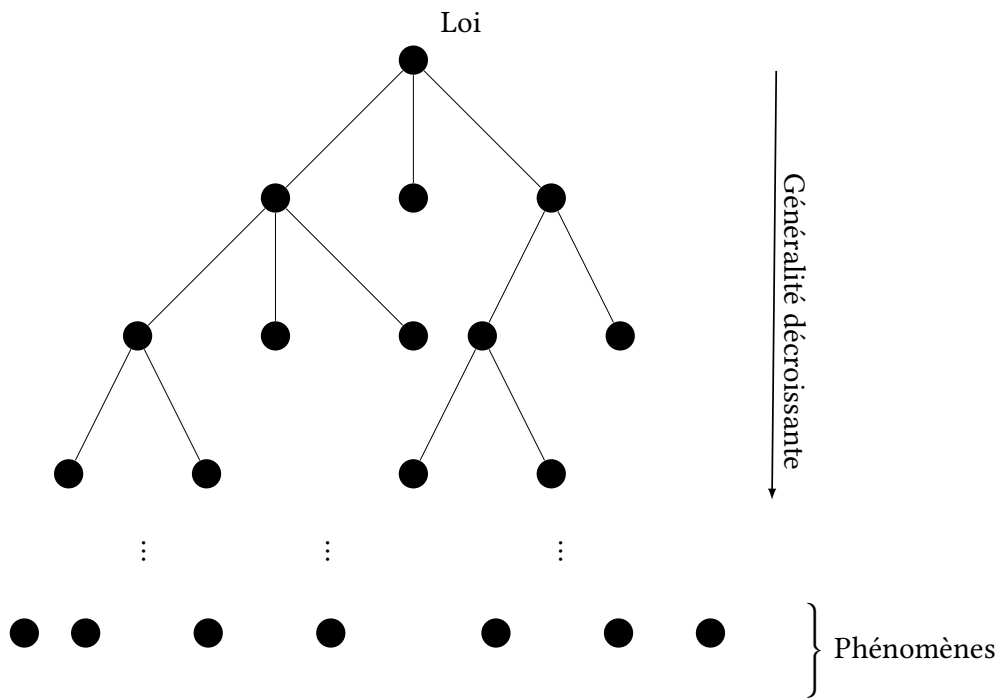


FIGURE 4.3 – Hiérarchie des modèles

Par exemple, dans le cas du modèle newtonien du pendule simple, on pourrait trouver :

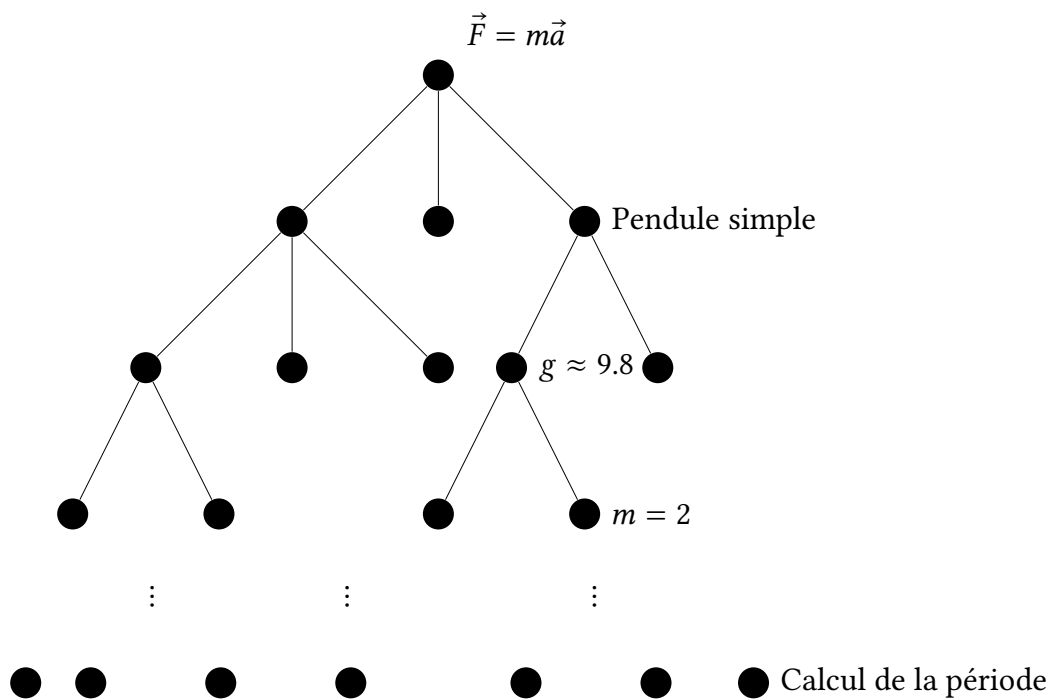


FIGURE 4.4 – Le pendule simple en tant que modèle newtonien

Un modèle est donc le support d'un jeu de questions-réponses, une entité dynamique à même de générer des propositions de plusieurs niveaux de généralité à propos d'une cible. La loi, au plus haut niveau de généralité, met en relation certaines variables qui peuvent être spécifiées de plusieurs manières, ce qui donne naissance à plusieurs modèles moins généraux qui, à leur tour, entrent dans le même genre de jeu. Au niveau de généralité le plus faible se trouvent les phénomènes, que je conçois comme des modèles de données.

Les modèles, une fois spécifiés, sont comme des boîtes noires auxquelles on peut fournir des conditions initiales afin qu'elles fournissent des propositions testables expérimentalement. Cela rend compte de l'idée répandue, et tout à fait juste il me semble, selon laquelle les modèles servent d'intermédiaire entre les lois et le monde empirique.

Le modèle du pendule simple une fois exploré, met en évidence une formule qui, une fois interprétée en termes de la cible, permet de calculer la période d'oscillation : $T = 2\pi \sqrt{\frac{L}{g}}$, avec L la longueur du fil et g l'accélération de la pesanteur. Ces deux paramètres a priori libres doivent être spécifiés à leur tour afin de faire générer au modèle une valeur numérique comparable à des mesures expérimentales. Le modèle le plus général du pendule génère donc une proposition susceptible d'être particularisée à des situations empiriques précises. L'ensemble de ces propositions peut en toute légitimité être considéré comme un modèle : il est le support d'actes d'imagination, et il est capable de générer des propositions à propos d'un système donné. Parmi ces modèles moins généraux, fixer $g \approx 9.81$ particularise à des situations potentiellement terriennes de pendules de longueur L , longueur qui sera spécifiée à son tour. Libre au scientifique d'imaginer un pendule lunaire en fixant $g = 1.61$, ou sur une planète quelconque en choisissant arbitrairement la valeur de g .⁶

Il est dès lors possible de compléter la description de la manipulation des modèles. Il s'agit d'une part de les explorer pour en découvrir de nouveaux aspects, de nouvelles vérités fictionnelles et d'autre part, il s'agit d'y incorporer différentes valeurs, de les spécifier afin d'obtenir des propositions différentes, des fictions aux caractéristiques proches mais cependant différentes.

À l'échelle locale, c'est-à-dire pour un modèle précis plutôt que pour l'arborescence dans laquelle il se situe, un modèle peut être vu comme une machine de génération de propositions dépendant de plusieurs facteurs que l'on peut régler pour voir l'effet de ces variations sur les valeurs

6. Il est important de noter que la valeur de g dépend elle-même de la géométrie de la planète, ainsi que de sa distribution de masse, des forces de marée, etc. Cela montre l'omniprésence d'hypothèses idéalisatrices dans l'arborescence des modèles.

que l'on obtient en sortie.

C'est évident dans le cas simplifié d'une formule mathématique interprétée afin qu'elle vise une cible physique : changer les valeurs des variables change le résultat. Mais la situation n'est pas fondamentalement différente dans le cas des modèles matériels. Ce que l'on peut entendre par la manipulation du modèle réduit d'un pont est évident : il s'agit là d'une manipulation physique (pour suivre Hacking, c'est bel et bien un modèle que l'on tient avec les mains). La théorie du *make-believe* prend en compte ces exemples et la proposition de Frigg et Nguyen pour traiter le problème de la représentation les met sur le même pied que les autres représentations épistémiques. Qu'en est-il de ma définition de la manipulabilité des modèles ? Il me semble que l'analyse peut-être transférée sans problème à ce genre de modèles : un modèle matériel comme un pont miniature prend en entrée (*input*) certaines forces, déformations ou contraintes et renvoie en sortie (*output*) des valeurs de solidité ou de résistance. Le lien entre les entrées et sorties dépend des propriétés du modèle, elles-mêmes dérivées des principes de génération.

Applicabilité et manipulabilité sont donc deux aspects essentiels des modèles. Si l'applicabilité permet de conceptualiser la manière dont le modèle vise sa cible, ce qui lui confère une interprétation et donc une portée sémantique compatible avec les thèses réalistes, la manipulabilité rend compte de la manière dont le modèle acquiert une portée empirique. Ces deux éléments sont évidemment importants pour la construction d'une théorie réaliste des modèles dans le cadre fictionnel, mais il manque toujours un ingrédient : la capacité à construire une procédure de justification.

Dans la section précédente, j'ai esquissé une stratégie de type *divide et impera* structuraliste qui montrait assez rapidement ses limites.

La notion de manipulabilité participe à faire émerger une stratégie qui permettra – c’est le but de cette thèse – de formuler une position réaliste tenable dans le contexte fictionnel. Cette stratégie passera également par l’examen plus précise du rôle des lois dans les modèles : c’est le sujet de la section suivante.

6 Idéalisations, lois, *comme si* : vertus épistémiques de la modélisation

Dans la perspective que j’ai développée au long de la section précédente, les idéalisations, approximations et abstractions apparaissent comme des procédures qui permettent de mettre en évidence certaines relations et de les rendre manipulables.

Il faut à ce propos rappeler la distinction introduite par Nancy Cartwright entre lois fondamentales et lois phénoménologiques : les lois fondamentales sont générales et explicatives, mais fausses – car elles ne décrivent que des situations idéalisées, elles ne sont vraies que dans les modèles –, tandis que les lois phénoménologiques sont vraies, mais ne s’appliquent qu’à des situations particulières, et ne sont donc pas explicatives.

Le rôle des idéalisations peut sembler paradoxal : elles permettent une isolation par la pensée des systèmes en particularisant, par exemple, une seule interaction isolée des autres qui interfèrent avec elle dans le monde empirique, mais dans le but d’atteindre une plus grande généralité. C’est là que se situe le cœur de l’argument de Cartwright : le caractère général des lois ne s’obtient que par le dépouillement, rendu visible par la clause *ceteris paribus*. C’est en supprimant toutes les sources d’in-

terférences que l'on peut obtenir un résultat aussi général qu'une loi fondamentale, mais le prix à payer est l'éloignement du monde empirique où ces interférences opèrent leur confusion en permanence. Cette observation n'est évidemment pas nouvelle : la possibilité de trouver de la régularité derrière la variabilité que l'on observe est un très ancien thème philosophique, et pas uniquement en philosophie des sciences, d'ailleurs.

Les modèles tels que je les conçois établissent la jointure empirique entre les lois fondamentales et le monde de l'expérience, comme je l'ai montré précédemment. C'est en cela que mon argument renverse le rôle attribué aux idéalizations : elles ne sont pas un prix à payer pour pouvoir décrire le monde au moyen d'outils aussi simples que des lois fondamentales au prix de la vérité empirique, elles sont au contraire le meilleur moyen par lequel les modèles peuvent toucher le monde empirique.

Les régularités encodées mathématiquement dans les lois physiques ne peuvent l'être que s'il y a préalablement une idéalisation : rassembler des données, les catégoriser, les ramener à un même phénomène est déjà de l'ordre du jeu fictionnel : on considère un ensemble d'observations *comme si* elles découlaient du même processus.

Dans une perspective à la Vaihinger, on peut tout à fait considérer l'apport des idéalizations comme une vertu méthodologique dont la perte de précision dans la description des systèmes physiques serait un prix à payer. En ce sens, les fictions sont utiles, mais ne sont pas un guide vers la vérité, et l'on retrouve l'argumentation antiréaliste de type fictionnaliste. Au contraire, l'omniprésence des idéalizations dans la pratique scientifique serait même le signe qu'il est nécessaire de s'éloigner de la vérité pour réussir à capturer la variabilité de la réalité au moyen de lois simples.

La conception fictionnelle rend justice à l'importance de ces idéalizations : elles ne figurent pas dans la connaissance scientifique « faute de mieux », elles en constituent au contraire un élément central et irréductible. Je reviendrai sur ce point dans la section 8 de ce chapitre.

Le prix à payer semble cependant être celui de la possibilité d'une conception réaliste du succès des sciences. Comme le fait remarquer Kwame Anthony Appiah dans son *As If: Idealization and Ideals* publié en 2017 (Appiah 2017) :

there is a range of virtues that an idealization might exhibit. A model might provide a unified account of a variety of phenomena or it might capture the essential causes of an effect with elegant simplicity. There's genuine insight in noticing a mathematical regularity in some natural phenomenon, even if the regularity becomes legible only through a loss of exactitude. (Appiah 2017, p. 54)

La capacité des modèles à capturer un aspect des phénomènes avec une *simplicité élégante* est en soi une vertu. Les régularités que l'on trouve dans les modèles, même si elles s'obtiennent au prix d'une perte d'exactitude, font bel et bien partie de la connaissance que la science procure.

Bien entendu, comme je l'ai déjà mentionné, accorder ce statut aux idéalizations ne permet pas d'interpréter cette vertu de manière réaliste, mais c'est un pas dans la direction que je souhaite suivre dans cette thèse : Appiah dépasse en ce sens la conception à tendance instrumentaliste de Vaihinger, qu'il cite explicitement et dont il reprend le concept de *comme si*. Ce que les idéalizations permettent d'obtenir a une place légitime dans la connaissance scientifique et ne se réduit pas à des ou-

tils qui ne sont évaluables qu'à la hauteur de leur aspect plus ou moins pratique.

Dans cet extrait, Appiah insiste sur deux vertus en particulier : l'unification et la capture de *causes essentielles*. L'unification est souvent une vertu épistémique mise en avant par les réalistes, mais qui est également critiquée par de nombreux arguments antiréalistes. D'une manière générale, les réalistes mettent en avant le fait qu'une théorie capable d'englober plus de phénomènes avec un nombre plus restreint de principes de base est un signe de maturité scientifique⁷, tandis que les antiréalistes insistent sur le fait que l'unification offre des avantages méthodologiques sans pour autant garantir la moindre correspondance ou vérité. De la même manière, l'élégance mathématique d'une théorie ou d'un modèle ne saurait être considérée comme un signe de vérité des énoncés empiriques que l'on obtient de la théorie ou du modèle.

La capture de *causes essentielles* d'un effet donné souffre des problèmes habituels lorsque la causalité est introduite. Il y a d'une part les nombreux arguments d'inspiration humienne à propos de l'impossibilité de vérifier expérimentalement un lien de cause à effet et, d'autre part, les difficultés à cerner le concept de causalité de manière suffisamment fine pour permettre de mettre en évidence des liens clairs entre causes et conséquences (Mackie 1974).

Pourtant, je pense qu'il y a tout de même un point intéressant dans ce que soulève Appiah avec l'idée que le modèle est capable de capturer quelque chose qui est de l'ordre de la cause d'un effet, et c'est également un élément que l'approche fictionnelle permet de comprendre avec une plus grande clarté.

7. J'entends maturité au sens défini dans le chapitre 2 dans le cadre du *no miracles argument* mis en avant par les réalistes.

7 Modèles comme générateurs de propositions contrefactuelles

7.1 Très brève introduction à la logique contrefactuelle

J'ai montré précédemment dans quelle mesure il est profitable d'envisager les modèles comme des entités dynamiques capables de générer des propositions à propos de leur cible. Je n'ai jusqu'ici pas précisé la nature de ces propositions : je prétends qu'elles sont contrefactuelles.

Une proposition contrefactuelle⁸ a une forme conditionnelle classique « si a alors b », mais dont l'antécédent a est factuellement faux. De exemples possibles sont : « Si j'étais président de la république, je prendrais telle ou telle mesure », « Si j'étais capable de voler, j'arriverais plus vite à mon rendez-vous » ou encore « S'il se mettait à pleuvoir, je sortirais mon parapluie ».⁹

La difficulté à traiter les contrefactuels avec les outils de la logique propositionnelle habituelle vient du problème qui se manifeste lorsqu'on essaie de déterminer leur valeur de vérité. En effet, envisagés comme de simples conditionnels de type « si a alors b », ils sont trivialement vrais, puisque l'antécédent est faux.

Pour rappel, la table de vérité de l'implication logique \rightarrow est la suivante :

8. Pour alléger le propos, j'utiliserai parfois le terme de « contrefactuel » pour désigner une proposition contrefactuelle ou un raisonnement contrefactuel.

9. Bien évidemment, ces exemples sont des conditionnels contrefactuels dans la mesure où le locuteur n'est pas président de la république et n'est pas capable de voler, ou s'il prononce la troisième lors d'un jour ensoleillé.

a	b	$a \rightarrow b$
F	F	T
F	T	T
T	F	F
T	T	T

Les propositions contrefactuelles nécessitent d'introduire une logique qui puisse rendre compte de l'aspect non trivial de ce genre de conditionnels. En effet, certains contrefactuels semblent tout à fait vrais, mais de manière non triviale, tandis que d'autres semblent faux. Par exemple, un jour de Soleil, « s'il se mettait à pleuvoir, le sol serait mouillé » est vrai, mais non trivial : il semble que ce soit lié à une certaine situation physique (par exemple l'absence de toit au-dessus du sol) et à une propriété de la pluie. En revanche, « si je lâchais cette tasse de café, elle rebondirait sur le sol » semble, dans des conditions physiques de la vie de tous les jours, être faux.

Le but de cette thèse n'est pas de proposer une analyse logique des contrefactuels, ce qui a été le travail de nombreux logiciens et philosophes, dont les deux plus connus sur ce sujet sont probablement Saul Kripke (Kripke 1959 ; Kripke 1963) et David K. Lewis (Lewis 1986 ; Lewis 2001) avec leurs propositions de logiques modales. Je souhaite par contre faire remarquer comment la notion de contrefactuels et de modèles fictionnels se marient bien ensemble.¹⁰

Pour l'exprimer succinctement, les modèles intégrant des hypothèses fausses, une partie des antécédents des propositions générées par ces modèles seront fausses. La menace de rendre les modèles triviaux est donc la même qu'avec la logique des propositions conditionnelles de

10. Pour une analyse de la modalité en physique théorique, voir (Butterfield 2002)

type contrefactuel. Par exemple, une proposition générée par le modèle newtonien du système solaire prendra la forme : « si les corps massifs s'attirent selon une force proportionnelle à la distance qui les sépare au carré et si les planètes sont des sphères ou sont assimilables à des points matériels alors telle planète aura telle trajectoire ». Cette dernière proposition est trivialement vraie si l'on utilise les outils de la logique non modale.

Le passage à une logique modale, c'est-à-dire une logique de la possibilité et de la nécessité, s'opère de manière assez naturelle à travers la notion de monde possible. L'idée principale est que la modification des conditionnels habituels peut se faire au moyen de l'introduction de deux opérateurs : \Box exprimant la nécessité et \Diamond exprimant la possibilité. Chacun peut être défini à partir de l'autre étant donné que : $\Box p = \neg \Diamond \neg p$ et $\Diamond p = \neg \Box \neg p$.

Afin d'interpréter les propositions modales, la notion de monde possible¹¹ est introduite. Prétendre qu'une proposition p est nécessaire (affirmer que $\Box p$) revient à affirmer que p est vraie dans tous les mondes possibles. La possibilité de p , notée $\Diamond p$, stipule quant à elle que dans certains mondes, mais pas tous¹², p est vraie. L'impossibilité de p , $\Box \neg p$, revient à affirmer que dans tous les mondes possibles, p est fautive, c'est-à-dire qu'il n'y a aucun monde possible où p pourrait avoir lieu.

Dans ce contexte, affirmer un contrefactuel du type « si je lâchais ma tasse de café, elle tomberait au sol et se casserait » se formalise en $\Box(p \rightarrow q)$ où p est « je lâche ma tasse » et q est « elle tombe et se casse ». Ce qui justifie la validité du contrefactuel énoncé est la nécessité de l'implication non contrefactuelle « lâcher la tasse implique qu'elle tombe et

11. David Lewis est connu pour avoir défendu une forme de réalisme des mondes possibles.

12. On a bien évidemment l'identité $\Box p \rightarrow \Diamond p$.

se casse ». En termes de mondes possibles, cela veut dire que dans tous les mondes où la tasse est lâchée, elle se casse. D'où la justification de la vérité du contrefactuel : même sans lâcher ma tasse, je sais qu'elle se cassera si je la lâche car j'ai connaissance de la nécessité de l'implication.

7.2 Modèles et contrefactuels

L'introduction de contrefactuels n'a pas seulement l'avantage de rendre compte du statut fictionnel des modèles. Les contrefactuels permettent également d'éclaircir le rôle et le statut des lois.

En effet, il est reconnu que les contrefactuels et les lois sont fortement liés : justifier la vérité d'un conditionnel contrefactuel prend assez naturellement la forme d'une loi. L'exemple de la pluie et du sol mouillé le montrait déjà : même s'il ne pleut pas, le conditionnel semble vrai, et la justification comporte une référence à un principe physique, c'est-à-dire à une loi (qui pourrait d'ailleurs elle-même prendre une forme conditionnelle « s'il pleut alors le sol est mouillé »).

Les modèles sont donc à même de générer ce type de proposition, c'est leur nature fictionnelle qui l'indique. Les lois, c'est-à-dire les principes de génération du *make-believe*, sont, comme on l'a vu précédemment, à la base de la création des modèles. Elles sont par conséquent aussi le support de la justification des propositions générées par le modèle.

Cela permet de proposer une définition minimale des lois dans le contexte des modèles fictionnels, définition qui évitera les problèmes ontologiques que j'ai abordés précédemment. Les lois ne sont pas des entités externes aux modèles, et la validité empirique et prédictive de ces derniers ne permet pas de justifier une correspondance entre ces lois réelles

et les lois mobilisées par le modèle. Les lois apparaissent avec l'utilisation des modèles car ces derniers permettent, par leur manipulation, de générer des propositions qui sont d'ordre contrefactuel. Cela explique pourquoi, malgré la reconnaissance de la position centrale des modèles dans la construction de la connaissance scientifique, les lois sont considérées comme le vecteur principal d'explications et de compréhension des phénomènes naturels.

Considérer que les contrefactuels mis en évidence par les modèles sont responsables de la justification sous forme de loi ne résout pas le problème posé au réaliste que j'ai soulevé précédemment. Aucune adéquation directe avec de supposées lois réelles ne peut être construite dans ce cadre : la loi est la forme logique que prend la justification des propositions générées par le modèle, et l'adéquation empirique que permettent ces dernières ne saurait justifier l'existence de lois extérieures au modèle. Le problème ontologique s'en trouve même obscurci. Étant donné qu'aucun sens physique n'est attaché aux lois du modèle, l'éventuelle correspondance avec des lois physiques doit répondre à une objection supplémentaire et expliquer la nature d'une telle correspondance.

Plutôt que de faire reposer la charge ontologique sur les lois elles-mêmes, il peut être tentant de leur accorder le pouvoir explicatif du modèle – qui serait alors essentiellement d'ordre logique – tout en détournant la recherche d'une correspondance sur un autre aspect des modèles. En particulier, si l'on interprète le conditionnel contrefactuel en termes causaux, le modèle semble capable de mettre en évidence des liens de causalité, d'une manière qui fait écho à la citation d'Anthony Appiah que j'ai commentée à la fin de la section précédente.

L'introduction des contrefactuels dans la discussion n'aide pourtant pas à reconforter les inquiétudes réalistes. On peut distinguer deux pro-

blèmes.

Le premier est que les théories contrefactuelles de la causalité conceptualisent la causalité au moyen des dépendances contrefactuelles, mais que considérer la présence de contrefactuels comme un signe de la présence de causalité est erroné.

We think of a cause as something that makes a difference, and the difference it makes must be a difference from what would have happened without it. Had it been absent, its effects – some of them, at least, and usually all – would have been absent as well. (Lewis 1973, p. 557)

La cause est ce dont la modification entraîne la modification de l'effet. Il y a donc bien entre cause et effet une dépendance contrefactuelle : « si la cause avait été modifiée de telle façon (ce qui n'est pas le cas), nous verrions telle et telle changement dans l'effet ».

Cependant, il serait incorrect de croire en la contraposée : toute dépendance contrefactuelle n'est pas un signe de causalité, et déduire des modèles générant des propositions contrefactuelles qu'ils mettent en évidence la causalité est une erreur d'analyse :

Suppose that two wholly distinct events occur, C and E; and if C had not occurred, E would not have occurred either. I say that if one event depends counterfactually on another in this way (and if it's the right sort of counterfactual, governed by the right sort of closeness of worlds) then E depends causally on C, and C is a cause of E. To be sure, this is only the beginning of a counterfactual analysis of causation. Not all counterfactuals are of the right sort, and it is a good question how to distinguish the ones that are from the ones

that aren't. We need an account of eventhood, and of distinctness of events. And not all effects depend counterfactually on their causes; for instance, we may have causation by a chain of stepwise dependence, in which E depends on D which depends on C, and thereby C causes E, yet E does not depend directly on C because of some alternate cause waiting in reserve. (Lewis 1986, p. 23)

Lewis met ici en évidence une double limitation : tous les effets ne dépendent pas contrefactuellement de leurs causes, et tous les contrefactuels ne sont pas le signe d'une relation causale. Cela impose de minimiser l'affirmation faite à propos des modèles : ils mettent en évidence des relations contrefactuelles interprétables causalement, mais cela suppose que l'on accorde déjà une valeur à cette interprétation. Faire reposer la validité des modèles sur leur adéquation avec des processus causaux reste ouvert aux critiques empiristes.

L'existence d'une dépendance contrefactuelle ne peut donc pas être considérée comme une preuve de l'existence d'un lien causal, au plus comme un signe de l'efficacité d'une telle description, mais les arguments antiréalistes qui visent l'impossibilité d'établir empiriquement la causalité gardent ici toute leur force.

Le second problème auquel fait face cette analyse est que les modèles étant des objets fictionnels décrivant parfois des entités inexistantes, il semble particulièrement douteux d'attacher à ce genre d'entités des propriétés causales.

On voit que la proximité entre les notions de lois, de causalité et de dépendances contrefactuelles se retrouvent au cœur de l'analyse fictionnelle des modèles que je propose, mais l'agencement exact de ces concepts reste encore à être mis en lumière.

Pour éclaircir ce point, il est utile de s'intéresser à la notion de schémas (ou motifs) de causes (*causal patterns*). Pour ce faire, je fais appel à la notion de manipulabilité des modèles que j'ai définie précédemment.

8 Contrefactuels et *Causal patterns* : la manipulabilité des modèles au service de l'exemplification de structures contrefactuelles

L'idée que les idéalizations jouent un rôle positif dans la construction de la connaissance scientifique est reconnue par de nombreux auteurs qui se sont intéressés à la question (voir par exemple (Bokulich 2009; Bokulich 2012; Grüne-Yanoff 2013; Sugden 2013)). Dans cette optique, les idéalizations ne sont pas un prix à payer afin d'obtenir en échange de la généralité explicative ou représentationnelle, elles participent activement à la pratique scientifique. Une conséquence de cette conception est qu'il n'est pas désirable de les éliminer du contenu de la science : cela ne ferait qu'affaiblir l'édifice tout entier en lui ôtant un important outil heuristique. En particulier, chercher à réduire l'importance des idéalizations pour se rapprocher de la vérité ou d'une description véridique des phénomènes naturels semble une voie sans issue :

epistemic aim of science is not truth but human understanding. Understanding is a cognitive achievement, and unlike truth, it can be directly furthered by idealizations. (Potochnik 2017, p. x)

Il semble que la reconnaissance de l'importance et de l'omniprésence

des idéalizations et des fictions remet en cause l'idée que la vérité est le but de la science. Si les idéalizations participent si intensément à la science, c'est parce que ce qu'elles permettent d'atteindre est d'un autre ordre : compréhension, explication, progrès technologique, bref des buts cognitifs ou techniques.

Depuis le début de cette thèse, j'ai tenté de résister à cette idée en montrant que l'on peut tout à fait aller jusqu'à reconnaître l'importance des fictions scientifiques sans pour autant abandonner l'optimisme épistémique du réalisme et continuer à considérer la vérité comme un but légitime (mais pas le seul) de la science.

Alisa Bokulich (Bokulich 2016) défend l'idée que la fiction a un rôle explicatif tout à fait légitime¹³ en sciences, mais elle n'exclut pas entièrement la vérité des buts que se donne la recherche.

À partir de l'exemple du modèle en vortex de l'éther de Maxwell, elle met en évidence un possible rôle à portée véridique des fictions. Maxwell, cherchant à expliquer le comportement de l'éther, établit un modèle dans lequel l'éther est un fluide agité par des tourbillons. Cela lui permet d'arriver à une description adéquate des phénomènes électromagnétiques, mais sans accorder le moindre statut ontologique à ces vortex d'éther :

we have now shown in what way electro-magnetic phenomena may be imitated by an *imaginary* system of molecular vortices. ((Maxwell 1861, p. 347) cité dans (Bokulich 2016, p. 269), les italiques sont ajoutés par Bokulich)

Maxwell qualifie lui-même ces entités d'imaginaires. Son modèle, bien qu'efficace dans sa description des phénomènes électromagnétiques

13. Je m'intéresserai plus longuement au problème de l'explication en sciences dans le chapitre suivant.

ne peut être interprété littéralement. Pourtant, l'analyse qu'en propose Bokulich ne se limite pas à la reconnaissance de son utilité pour la compréhension d'un phénomène physique. Le notion de vérité, malgré la tension qui l'oppose aux idéalizations, y garde une place :

Maxwell is quite explicit that this is a fictional representation; nonetheless it is by reasoning with this fiction that he comes to the most important discovery of his career : that electromagnetic waves travel at the same speed as light, a coincidence that could only be justified by their identity. In other words, the fiction of vortices functioned as a vehicle for the truth that light is electromagnetic radiation. (Bokulich 2016, p. 269)

La fiction est un guide vers la vérité (*vehicle for truth*), et la découverte de Maxwell, bien que faite à partir d'une entité totalement imaginaire¹⁴, est pleinement justifiée véridiquement.

Il y a donc une place pour la vérité et les idéalizations dans les modèles. La position de Bokulich me semble compatible avec une position réaliste structurale du genre de celles que l'on trouve chez Poincaré. L'exemple de l'éther est d'ailleurs analogue à celui dont se sert John Worrall pour établir son *best of both worlds* (voir chapitre 2). Le modèle de Maxwell se trompe sur les entités, mais il met en évidence de bonnes structures, et ce serait cela qui fonde son adéquation empirique et son pouvoir prédictif.

L'apport de l'approche fictionnelle que je défends consiste à montrer comment un modèle capture cette structure, quel est son lien avec les lois ainsi que de proposer une définition plus stricte de la structure, ce qui

14. Deux, en réalité, puisque l'éther n'existe pas plus que les vortex en son sein.

est une critique récurrente du réalisme structural : c'est une notion qui semble extrêmement vague et que les tenants de la conception ont du mal à définir précisément.

Les structures qu'un modèle met en évidence sont contrefactuelles, comme je l'ai déjà montré, et j'ai également explicité le rôle des lois dans la construction des modèles. La question qu'il me reste à éclaircir avant de formuler la stratégie *divide et impera* réaliste dans le cadre fictionnel est celle de l'exemplification : comment les modèles peuvent-ils faire émerger des structures contrefactuelles ?

Je nomme cette étape exemplification car il y a une sémantique attachée au modèle, c'est son applicabilité à la cible, qui n'est pas univoque du point de vue empirique. En ce sens, le modèle exemplifie des phénomènes en proposant une interprétation possible des dépendances contrefactuelles qu'il contient. L'interprétation des processus à l'intérieur du modèle sont désignés par des termes qui reçoivent leur sens des systèmes physiques auxquels le modèle est rattaché, mais cela n'est pas vérifiable, dans le sens où aucun lien ne saurait être fait entre les images dont nous nous servons pour parler du modèle et l'adéquation empirique de ce dernier. En revanche, ce sont les dépendances contrefactuelles elles-mêmes qui fondent la capacité du modèle à générer des propositions vérifiables empiriquement.

Pour utiliser des termes plus proches de Poincaré et de Worrall : tout ce qui dans le modèle tient de l'interprétation, du jeu de *make-believe* et du *comme si* fictionnel se rattache aux images que la nature gardera toujours hors de notre portée. Nous pouvons en revanche espérer avoir un accès aux structures contrefactuelles que le modèle met en évidence, qu'il exemplifie en nous les présentant sous une forme qui convoque des images c'est-à-dire une interprétation qui facilite le raisonnement,

la représentation mentale des phénomènes et le discours à leur propos.

Le fait que les modèles sont manipulables, qu'ils sont des entités dynamiques que l'on peut utiliser afin de générer des propositions qui portent sur des systèmes-cible, est ce qui leur permet d'exemplifier des relations contrefactuelles. Spécifier un modèle revient à établir des relations entre des variables, comme je l'ai montré dans les sections précédentes. Pris ainsi, le modèle offre un avantage très restreint : c'est toute une machinerie dont le seul aspect nouveau est d'offrir un cadre pour interpréter cette relation.

Ce n'est que par la manipulabilité que le modèle fait véritablement émerger la structure contrefactuelle : c'est en variant les entrées qu'on lui donne et en examinant ce qui en ressort que le modèle acquiert un pouvoir prédictif non trivial.

L'exemple du pendule simple est particulièrement parlant. On peut découvrir à son propos, et grâce aux idéalizations que le modèle contient, une vérité fictionnelle qui concerne la fréquence des petites oscillations. Cette proposition, une fois interprétée en termes de la cible acquiert un pouvoir prédictif. En faisant varier contrefactuellement ce que l'on donne au modèle en entrée, on peut vérifier par l'observation la robustesse de la relation contrefactuelle dans toute une gamme de situations.

On changera par exemple la longueur du fil et l'accélération de la pesanteur¹⁵ et si le modèle résiste à cette manipulation contrefactuelle, il acquiert une puissance prédictive forte.

Il est possible d'interpréter cette robustesse de la variation contrefactuelle de manière causale, d'une manière analogue à David Lewis, ou même dans une veine humienne :

15. Ce qui est nettement plus difficile que de couper un morceau de corde !

Suitably to this experience, therefore, we may define a cause to be an object followed by another, and where all the objects, similar to the first, are followed by objects similar to the second. Or, in other words, where, if the first object had not been, the second never had existed ¹⁶ (Hume 1748, p. 208)

La proposition d'Angela Potochnik insiste sur la possibilité pour les idéalizations de mettre en évidence des *causal patterns*, et cela en relation directe avec cette conception de la causalité :

Scientific laws, theories, and models typically only hold approximately, in some range of circumstances, and they liberally employ idealizations to accomplish this partial fit. I argue that they depict introduction causal patterns. These patterns qualify as causal according to a manipulability approach to causation. (Potochnik 2017, pp. 18-19)

De plus, la mise en évidence de ces *causal patterns* est la raison d'être principale des idéalizations :

depicting causal patterns regularly motivates departures from accuracy of any given phenomenon ; this is why idealizations are used to represent as-if. Put in these terms, the present idea is that idealizations positively contribute to generating understanding by revealing causal patterns and thereby enabling insights about these patterns that would otherwise be inaccessible to us. (Potochnik 2017, p. 95)

16. « Conformément à cette expérience, nous pouvons donc définir la cause comme un objet qui est suivi d'un autre, et de telle sorte que tous les objets semblables au premier sont suivis par des objets semblables au second. Ou, en d'autres mots : de telle sorte que le second objet n'aurait jamais existé sans l'existence du premier. » (Hume 1748, p. 209)

La découverte de ces motifs causaux est la raison de l'utilisation d'idéalisations, et c'est cela qui génère la compréhension des phénomènes. On peut se demander dans quelle mesure ces *causal patterns* sont réelles, au sens où elles seraient de véritables portions causales de la réalité. Potochnik est peu claire sur ce point : elle les qualifie de réelles, mais, comme je l'ai vu en début de section, elle nie également la possibilité de concevoir la vérité comme un des buts de la science. De plus, elle accorde un statut privilégié aux *causal patterns* par rapport à la causalité :

I employ a manipulability approach to causation to articulate the nature of causal patterns, but I remain agnostic on the status of manipulability relations in the metaphysics of causation (Potochnik 2017, p. 23)

Il semble donc que ces motifs causaux puissent être acceptés sans forcément les adjoindre à une théorie particulière de la causalité. C'est la raison pour laquelle il me semble intéressant de faire reposer la question de la causalité sur les lois plutôt que sur la structure contrefactuelle : les lois servent de justification aux propositions contrefactuelles générées par le modèle. Dans un second temps, il est possible d'attribuer un statut causal à ces lois, mais cela n'interfère en rien avec la manière dont le modèle exemplifie la structure contrefactuelle. Potochnik ne considère pas l'importance active des lois dans la génération des modèles, et la causalité, dans le cas où on déciderait de ne pas rester agnostique à son sujet, doit être imposée au niveau des motifs eux-mêmes. Cela me semble poser un problème si on les considère d'une manière ou l'autre comme réels, car cela entraîne de devoir affronter plus directement encore la question de l'ontologie de la causalité.

Le rôle et la nature exacts de ces *causal patterns* restent donc flous.

Potochnik écrit que son point de vue ne repose sur aucune métaphysique particulière de la causalité, mais elle reconnaît que les schémas de causalité doivent être réels :

How, then, can we tell if understanding is actual and not merely apparent? For this, the causal pattern apparently grasped must be real. [...] Briefly, for a causal pattern to be real, it must be embodied (to some degree or other) in some range of phenomena. (Potochnik 2017, p. 115)

Les motifs de cause sont incarnés dans des phénomènes, mais Potochnik rejette, comme je l'ai déjà montré, la vérité comme principal objectif épistémique de la science et lui substitue d'autres buts :

The clearest illustrations of this are idealizations themselves, which are quite far from the truth but, in the right circumstances, are epistemically acceptable nonetheless. So, in my view, science simply is not after the truth. There are some important ways in which truth still may be involved in the scientific enterprise, but in each case, it is only a means to other ends. (Potochnik 2017, p. 117)

Ainsi, lorsque la vérité est effectivement mise en avant par les scientifiques, ce serait toujours avec d'autres objectifs, plus importants, à l'esprit : la compréhension des phénomènes, qui n'est pas axée sur la nécessité de la vérité. Mais qu'est-ce que cela signifierait de mettre en évidence des modèles de causalité réels d'une manière qui ne soit pas centrée sur la vérité ? Il me semble évident que qualifier quelque chose de réel dans un modèle, même fortement idéalisé, implique une certaine forme de correspondance et, en fin de compte, une vérité (au moins partielle). Il

Il y a une tension en jeu ici : l'exactitude est une condition de l'acceptabilité épistémique, mais nous devons refuser de l'aligner sur la vérité, car l'idéalisation aide à générer des explications et de la compréhension. L'énigme initiale n'est toujours pas résolue, la tension entre idéalisation et vérité persiste.

Appiah est quant à lui conscient du pouvoir des modèles à représenter structurellement. À propos d'un modèle fonctionnel du cerveau à partir de neurones dont le comportement est très fortement idéalisé, il écrit :

So a highly idealized model of the brain acquires independent utility because its simplifying idealizations ended up providing techniques for mimicking the functions rather than the material substrate of the mind. The McCulloch-Pitts neuron idealized radically the structural components of the nervous system and helped us understand how a brain composed of actual neurons might work. (Appiah 2017, p. 34)

Le modèle est donc à même de reproduire un comportement observé à partir d'entités idéalisées plongées dans une structure donnée. Selon l'analyse que j'ai proposée plus tôt dans ce chapitre, cette structure idéalisée est d'ordre contrefactuel, et ce qu'Appiah met en avant peut donc s'apparenter aux *causal patterns* de Potochnik. Ce modèle du cerveau se trompe radicalement sur la nature des neurones, mais il reproduit avec une grande précision le comportement fonctionnel du système, ce qui permet de comprendre comment un cerveau composé de véritables neurones peut fonctionner.

La manipulabilité du modèle permet non seulement de mettre en évidence des dépendances contrefactuelles, mais aussi de tester leur robus-

tesse en variant les entrées que l'on donne au modèle. Le statut dynamique du modèle est à même de produire une gamme de situations et, si cela reproduit le comportement du système cible, le modèle acquiert une grande adéquation empirique et un grand pouvoir prédictif.

Cette robustesse des relations contrefactuelles exemplifiées par le modèle à représenter le comportement d'un système cible s'apparente à ce que l'on pouvait qualifier de mature dans les théories : les modèles qui sont capables d'une grande robustesse dans une large gamme de situations sont ceux qui jouissent du statut épistémique le plus sérieux. Il convient donc maintenant de s'interroger spécifiquement sur le rôle de cette approche fictionnelle des modèles dans la recherche d'une formulation réaliste de la modélisation.

C'est le sujet de la section suivante.

9 Une stratégie contrefactuelle du *divide et impera* structural

En résumé, je propose de concevoir le fonctionnement des modèles selon un processus inspiré de l'approche fictionnelle du *make-believe* tel que théorisé par Kendall Walton. Selon les tenants de l'application de ce concept aux questions scientifiques, un modèle sert de support dans un jeu de *make-believe*, c'est-à-dire que le modèle acquiert un pouvoir de représentation d'un système physique en poussant l'utilisateur à imaginer certaines choses à propos de la cible.

En m'écartant du problème de la représentation scientifique par les modèles, je me suis concentré sur la question plus générale du réalisme scientifique. En particulier, je souhaitais montrer qu'il était possible de

construire une position véritablement réaliste dans le contexte fictionnel, malgré l'apparente contradiction entre la notion de fiction et le réalisme. En effet, le réalisme vise principalement à interpréter le succès empirique et prédictif des théories et modèles comme une justification d'une forme de correspondance ou de vérité, alors que l'approche fictionnelle met au contraire en évidence le caractère imaginaire du contenu des modèles.

Au cours de ce chapitre, j'ai montré que les conséquences de l'approche fictionnelle ne mènent pas nécessairement à concevoir les modèles comme des outils dont la seule évaluation se mesure à leur utilité pratique et instrumentale. En particulier, je me suis attelé à formuler les deux thèses réalistes centrales que j'avais définies dans le chapitre 2 : la thèse sémantique et la thèse épistémologique.

Du point de vue sémantique, j'ai éclairci ce que « être un modèle de » veut dire ainsi que la manière dont on confère au contenu du modèle une interprétation. Le réaliste a besoin que la sémantique du modèle soit essentiellement déterminée par la cible, et cela se fait, comme je l'ai montré, au moyen du concept d'applicabilité. La notion d'application permet de fournir une sémantique au modèle en définissant ce que « viser la cible » veut dire : un modèle donné est un modèle d'une cible particulière parce qu'il lui est applicable. J'ai également montré que les limitations que la conception fictionnelle impose à cette relation sémantique modèle-cible ne posaient pas un véritable problème au réaliste.

Du point de vue épistémologique, il apparaît que les modèles, qui sont souvent définis en tant que structures (voir à ce sujet la discussion de la sous-section 3.4 du chapitre 2), amène naturellement à imaginer une défense de type *divide et impera* structuraliste. Il se trouve que la conception fictionnelle met en avant le rôle des lois dans la génération

des modèles, et l'idée intuitive est par conséquent d'attribuer aux lois ou aux structures (généralement interprétées comme causales) qu'elles décrivent le statut de fondement de la justification réaliste du modèle. Ce qui explique l'efficacité empirique et prédictive du modèle est qu'il a correctement capturé (au moins approximativement) les lois ou les structures qu'elles décrivent. J'ai montré que la première option n'était pas viable à moins de résoudre préalablement les problèmes liés à l'ontologie et la métaphysique des lois. Quant à la seconde, elle nécessite de rendre plus explicite ce que l'on entend par structure au sein des modèles.

C'est la manipulabilité du modèle, c'est-à-dire sa capacité à fonctionner comme une entité dynamique avec laquelle l'utilisateur interagit, qui permet de faire émerger une structure de dépendances contrefactuelles. La nature contrefactuelle des propositions générées par le modèle est une conséquence directe de l'approche fictionnelle et de la reconnaissance de la présence d'idéalisations dans le processus de modélisation. Le modèle exemplifie ensuite ces dépendances contrefactuelles. La robustesse du modèle, sa capacité à reproduire de manière cohérente les dépendances observées permet de leur attacher une interprétation qui elle aussi vise la cible : le modèle met en scène des interactions et leur donne un sens qui dépend de la cible.

Je suis maintenant en position de formuler une interprétation réaliste du succès empirique et prédictif des modèles : elle concerne directement la structure de dépendances contrefactuelles exemplifiée par le modèle.

Prétendre qu'un modèle est empiriquement adéquat revient à dire que les propositions contrefactuelles qu'il génère sont vérifiées par l'expérience. Ces propositions proviennent de la structure exemplifiée par le modèle, et la manipulation lui confère une capacité prédictive : la nature contrefactuelle de ces dépendances exemplifiées pousse à varier les situations initiales et à

tester la capacité du modèle à reproduire correctement le phénomène modélisé.

Ma proposition est véritablement agnostique quant au statut causal de ces dépendances : il n'est pas nécessaire de les comprendre comme des relations de cause à effet dans la cible. De même qu'il n'est pas nécessaire de considérer les lois fondamentales comme des descriptions de chaînes causales. En revanche, la loi étant la forme que prend la justification d'une relation contrefactuelle, il est naturel d'accorder aux lois fondamentales un statut explicatif privilégié.

Le modèle permet donc une justification indirecte des lois décrivant la structure contrefactuelle exemplifiée grâce à la manipulation du modèle. Cela se fait selon un processus en trois étapes : application, manipulation, exemplification. Les lois peuvent donc être prises comme la base de la justification du succès du modèle, mais uniquement dans leur rôle de descriptions de relations contrefactuelles. Ce sont ces dernières qui sont l'élément du pari épistémologique réaliste : c'est la capacité du modèle à être robuste dans sa reproduction de dépendances contrefactuelles dans le comportement de la cible qui justifie sa valeur empirique.¹⁷ Cela offre un argument de type *no miracles* au réaliste : ce qui explique le mieux l'efficacité des modèles les plus robustes est l'adéquation de la structure de dépendances contrefactuelles qu'ils exemplifient avec celles

17. Ce cadre d'analyse permet également d'accommoder la distinction entre observables et inobservables. En effet, les arguments avancés par van Fraassen et de nombreux autres antiréalistes visent explicitement le discours théorique à propos d'inobservables : ce qui est nié n'est pas l'existence des objets observables – tables, chaises et moteurs thermiques –, mais la possibilité de déduire de l'adéquation empirique d'un modèle l'existence des entités inobservables – électrons, forces ou points matériels – qu'il mobilise. Il est largement reconnu que la frontière qui sépare observables et inobservables est floue, et l'approche fictionnelle a l'avantage de ne pas dépendre de la manière dont on choisit de tracer la ligne de division. Les objets inobservables introduits dans le modèle font en effet partie du jeu de *make-believe*, ils participent à la manipulabilité du modèle ainsi qu'à l'exemplification des dépendances contrefactuelles. En revanche, la justification réaliste du modèle ne dépend pas d'une éventuelle correspondance de ces entités avec des objets réels.

des phénomènes modélisés.

Chapitre 5

Véritisme et explication scientifique

Jusqu'ici, j'ai développé un cadre de pensée général pour formuler une stratégie réaliste à partir de la conception fictionnelle des modèles. Il me semble intéressant de chercher à l'appliquer à des problèmes connexes à celui du débat entre réalisme et antiréalisme. Ce chapitre dédié à la notion d'explication est l'occasion d'illustrer les différents concepts abordés dans ce qui précède et de montrer dans quelle mesure la conception fictionnelle telle que je la défends permet de résoudre le problème du véritisme de l'explication, proche de celui du réalisme.

Dans le chapitre 2, j'ai insisté sur le caractère très général du terme « réalisme ». La définition que j'utilise, basée sur trois thèses principales, permet d'aborder le versant épistémologique du réalisme. Toutefois, comme l'écrivait Stathis Psillos, le réalisme est une constellation de positions philosophiques qui partagent certains principes de base, comme une tendance matérialiste et une approche naturaliste des phénomènes et des produits de la science.

La question de l'explication fait partie des problèmes de philosophie des sciences qui peuvent tout à fait être abordés avec une perspective réaliste ou antiréaliste. Il s'agit, d'une part, de définir ce qu'est une explication en sciences et, d'autre part, de montrer comment l'explication d'un phénomène scientifique peut trouver une justification. On peut voir immédiatement comment les deux types de réponse peuvent s'opposer sur ce terrain : d'un côté les réalistes insisteront sur la nécessité d'une forme de vérité ou de correspondance pour obtenir une explication scientifique légitime¹, tandis que les antiréalistes insisteront sur l'impossibilité d'une telle démarche.

Or, l'épistémologie des modèles est confrontée à une énigme : les modèles sont souvent décrits comme hautement idéalisés, alors qu'ils sont considérés comme des vecteurs d'explications scientifiques. Les conceptions de l'explication orientées vers la vérité, que l'on peut qualifier de « véritistes », semblent ainsi minées par cette contradiction. Le point commun entre l'analyse de la fonction explicative des modèles et le problème général du réalisme sont évidents.

Comme lors du développement des chapitres précédents, je montre dans celui-ci comment ce paradoxe apparent peut être évité en faisant appel à la notion de fiction et à la conception qui en découle. Je ferai également appel aux concepts de modalité et de mondes crédibles, très fortement liés à la notion de propositions contrefactuelles. Au final, véritisme de l'explication et fiction peuvent effectivement aller de pair, mais, comme on peut s'y attendre, la portée des explications que l'on tire des modèles doit être plus précisément délimitée.

1. L'exemple le plus connu est le modèle déductif-nomologique (DN) proposé notamment par Hempel et dans lequel une explication se trouve au sein d'un système déductif, dont la vérité des prémisses garantit la vérité des conclusions.

1 Définition du problème

Les modèles sont omniprésents en science et sont supposés remplir de nombreuses fonctions, parmi lesquelles la représentation de systèmes physiques, que j'ai déjà abordée précédemment. L'explication en fait également partie : il est largement reconnu que les modèles peuvent participer à expliquer des phénomènes. La question qui se pose est celle du mécanisme explicatif : quel est le rôle des modèles dans les explications scientifiques ? Les modèles constituent-ils les explications ou y participent-ils d'une manière qui reste à déterminer ?

La reconnaissance de la large utilisation des modèles entraîne le même questionnement face à leur contenu. Étant donné que l'une des principales caractéristiques des pratiques de modélisation est l'idéalisation – les modèles sont toujours plus simples que les systèmes ou les situations qu'ils décrivent ou expliquent (« Models are generally caricatures of the natural world » (Chakravartty 2001)) – la question de la nécessité d'une forme de vérité pour justifier la valeur des explications tirées des modèles se pose.

Cela semble à première vue problématique, voire paradoxal : les modèles sont des véhicules d'explication, et pourtant ils contiennent des idéalizations, des distorsions, des objets purement fictifs ou même des impossibilités. À la suite d'Elgin (voir par exemple (Elgin 2017)), face à l'utilisation de ces *felicitous falsehoods* (faussetés heureuses), on pourrait plaider pour l'assouplissement de la vérité comme principal objectif épistémique de la science, d'où sa critique du *véritisme*.

Dans le contexte de l'explication scientifique, le véritisme est le point de vue qui considère la vérité comme une condition nécessaire à l'explication (Pincock 2021). Nous sommes confrontés à un dilemme qui a un

air de déjà-vu : soit les modèles idéalisés réfutent la position vériste, soit les modèles idéalisés ne sont pas épistémiquement légitimes et les explications fournies par ces modèles doivent être rejetées. La seconde possibilité n'est manifestement pas conforme à la pratique scientifique : des modèles hautement idéalisés sont souvent utilisés pour expliquer les phénomènes observés, et les explications basées sur les modèles sont considérées comme une partie pleinement légitime de la connaissance scientifique. Le véritisme semble alors réfuté.

Comme dans le cas du problème de la représentation, les inquiétudes portent sur la conception d'une relation modèle-monde comme base de la connaissance. Le point commun que je mets en avant entre les deux problèmes ne force pourtant pas à adopter la même stratégie de réponse dans les deux cas :

veritist need not endorse any specific account of how models represent. (Pincock 2021)

Le véritisme est compatible avec plusieurs visions de la manière dont les modèles représentent les systèmes physiques, et il est possible de critiquer la position vériste à partir de plusieurs conceptions de la représentation. Je suis tout à fait d'accord avec cette remarque, mais la proposition des chapitres précédents forme un cadre de pensée plus général que pour traiter spécifiquement de la représentation. Réalisme et véritisme sont deux positions proches. Je veux montrer ici comment il est possible d'utiliser les ressources fournies par le cadre fictionnel pour clarifier l'énigme du véritisme.

Une courte remarque sur l'agencement des problèmes philosophiques

Il existe une constellation des conceptions philosophiques orientées vers une forme de vérité-correspondance, et la défense de chacune d'entre elles nécessite un traitement spécifique, même si les arguments qui leur sont opposés ont souvent des points communs.

J'ai déjà eu l'occasion dans ce qui précède de mettre en évidence une caractéristique particulière de la notion de représentation qui empêchait un traitement purement réaliste : l'intentionnalité irréductible de ce concept. Cela limitait les aspirations orientées vers la vérité que l'on pouvait construire à propos de la représentation de systèmes physiques par les modèles. C'est en cela que je prétendais me servir d'un outil – l'approche fictionnelle – développé pour traiter un problème précis afin de construire un cadre de pensée plus général.

La stratégie réaliste suggérée par l'approche fictionnelle établit que c'est la structure de dépendances contrefactuelles exemplifiée par le modèle qui doit être prise comme base de la justification de la validité du modèle en question. Une fois ce cadre accepté, il convient d'examiner comment il peut s'appliquer à des problèmes plus précis, dont celui d'explication qui m'occupe ici.

Cependant, si l'objet d'étude influe sur la stratégie à adopter, si une forme d'adaptation est requise, il faut s'attendre à ce que le cadre de pensée que j'ai développé tout au long de cette thèse ne puisse s'appliquer brutalement au problème de l'explication.

En effet, ce qui est attendu d'une explication scientifique est assez nébuleux. La notion elle-même est sujette à de nombreux débats, et la délimitation du concept d'explication dépasse le cadre de cette thèse.

Une remarque s'impose tout de même : la conception de la fiction que je défends apporte un argument de défense du véritisme sans chercher à restreindre la définition de ce qu'est une explication. Je défends que la position véritiste est tenable, qu'il est donc possible d'obtenir d'un modèle des explications qui tirent leur validité d'une forme de correspondance, pas que toute explication doit se conformer à ce schéma.

Le problème du réalisme concerne de manière très générale le lien entre la force empirique d'une théorie ou d'un modèle et la potentielle véracité de la théorie ou du modèle. La représentation concerne plus précisément la capacité du modèle à soutenir certains types d'inférences à la cible, que Frigg et Nguyen, parmi d'autres, comparent à la représentation en art. L'explication scientifique est plus complexe à cerner, car elle semble à la fois contenir une notion logique – qui à trait à la manière dont un énoncé trouve sa place dans un système plus large (comme c'est par exemple le cas dans le modèle déductif-nomologique à la Hempel ou à la Popper) – et un potentiel d'éclaircissement de certains processus ou propositions. Une explication génère aussi un sentiment de compréhension, que la seule dérivabilité logique ne peut recouvrir.

Je nomme dans la suite « explication » cette notion générale et nébuleuse, sans plus de précision dans la définition, mon but étant de montrer qu'une base véritiste peut de toute façon être préservée.

2 Fiction et explication

La critique du véritisme dans le cadre du problème de l'explication prend une forme qui, dans le prolongement de tout ce que j'ai développé au cours des chapitres précédents, n'a rien d'étonnant : les modèles procurent des explications légitimes grâce aux idéalizations qu'ils

mobilisent, et cet éloignement d'une description véridique des processus naturels rend le concept de vérité impossible à concilier (du moins en apparence) avec la reconnaissance de la nature fictionnelle des modèles.

Dans son livre de 2017 *True Enough*, Catherine Elgin établit cette incompatibilité dès la première page :

Modern science is one of humanity's greatest cognitive achievements. To think that this achievement is a fluke would be mad. So epistemology has the task of accounting for science's success. A truth-centered, or veritistic, epistemology must treat models, idealizations, and thought experiments as mere heuristics, or forecast their disappearance with the advancement of scientific understanding. Neither approach is plausible. We should not cavalierly assume that the inaccuracy of models and idealizations constitutes an inadequacy; quite the opposite. I suggest that their divergence from truth or representational accuracy fosters their epistemic functioning. When effective, models and idealizations are, I contend, felicitous falsehoods. They are more than heuristics. They are ineliminable and epistemically valuable components of the understanding science supplies. (Elgin 2017, p. 1)

Les idéalizations ont donc, comme je l'ai déjà montré, une vertu épistémique qui leur procure une place légitime dans la connaissance scientifique. C'est même, selon Elgin, qui est sur ce point en total accord avec Bokulich (voir la section 8 du chapitre précédent), le principal apport des idéalizations : elles renforcent le pouvoir explicatif des modèles.

Comme le point de vue du *make-believe* parie sur l'importance centrale des aspects fictifs de la modélisation, des explications fondées sur

la vérité ne peuvent pas du tout être tirées des modèles, semble-t-il, et la conclusion est la même que celle d'Elgin : les *felicitous falsehoods* ne peuvent pas être éliminées du succès de la science, et le véritisme doit par conséquent être abandonné.

On retrouve ici la conception à tendance instrumentaliste de la fiction, mais appliquée au problème des explications scientifiques. Puisque la fausseté s'insinue d'une manière ou d'une autre dans les modèles, cela remet en question le statut de ce qui est tiré des modèles, et la vertu épistémique des idéalizations ne peut être basée sur une forme de vérité ou de correspondance.

Je pense qu'il existe néanmoins un moyen de défendre une forme modifiée de véritisme dans le contexte de la vision fictionnelle. Une telle défense doit se faire en deux étapes. La première concerne les vertus épistémiques. Elgin reconnaît la vertu des idéalizations, mais rejette celle de la vérité. De son côté, le véritiste affirme que la principale vertu épistémique est la vérité, mais sans pour autant rejeter l'idée que les idéalizations peuvent également avoir une certaine valeur. L'approche fictionnelle permet de réfuter l'affirmation selon laquelle ces deux vertus épistémiques ne peuvent pas être articulées dans un cadre commun.

Il existe une abondante littérature sur les vertus épistémiques, et plusieurs auteurs ont défendu l'idée que le véritisme et l'idéalisation ne sont peut-être pas aussi incompatibles qu'on pourrait le penser à première vue. Par exemple, Nawar affirme que :

in grasping and idealizing claim *as an idealizing claim*, it seems that one in fact grasps a truth. (Nawar 2019)

Le fait de savoir qu'une idéalisation est une idéalisation consiste déjà en une forme de connaissance qui touche à une forme de vérité. Il reste

évidemment à caractériser cette vérité que le modèle idéalisé permet d'atteindre, mais cela montre déjà une compatibilité a priori.

Sullivan et Khalifa (Sullivan et Khalifa 2019) admettent quant à eux que l'idéalisation a des vertus, mais qu'elles ne sont pas épistémiques. Les idéalizations sont utilisées pour des raisons de commodité, de simplicité ou de calculabilité. En ce sens, elles sont des faussetés heureuses, mais seules leurs composantes véridiques peuvent participer à une explication légitime. Dans le même ordre d'idées, Lawler affirme :

falsehoods can play an epistemic enabling role in the process of obtaining understanding but are not elements of the explanations or analyses that constitute the content of understanding. (Lawler 2019)

Cela revient à reconnaître une vertu aux idéalizations : elles permettent de mettre en avant une explication, mais elles ne constituent pas elles-mêmes les explications. Cette remarque est analogue à la conclusion obtenue plus tôt dans le cadre fictionnel : les idéalizations permettent ce que j'ai appelé la manipulabilité du modèle, mais elles ne remettent pas en cause l'entière vérité de la potentielle véracité du modèle. Ce que Lawler esquisse dans ce passage est une forme de stratégie *divide et impera* appliquée aux explications scientifiques. La distinction se fait entre le processus d'explication que les idéalizations facilitent et le contenu de l'explication qui – c'est ma proposition – peut faire l'objet d'une analyse véritiste, c'est-à-dire orientée vers la vérité.

L'avantage de l'approche fictionnelle trouve ici une nouvelle formulation : il est possible d'admettre une place à la vertu des idéalizations tout en refusant d'abandonner la vérité comme l'un des buts principaux de la science et des explications qu'elle vise à fournir.

La distinction processus/contenu de l'explication, appelée *extraction view* (la conception de l'extraction), est intéressante pour l'objectif que je me suis donné dans ce chapitre. En ce qui concerne les processus, le point de vue de la fiction permet toutes les vertus habituelles accordées aux idéalizations en considérant ces procédures de fictionnalisation comme la caractéristique centrale des modèles. En ce qui concerne le contenu, pour être pleinement adéquat dans le contexte de la fiction, il est nécessaire de clarifier en termes de *make-believe* ce qu'est exactement le contenu des explications fournies par les modèles.

Il est maintenant nécessaire de se tourner vers la question des explications elles-mêmes. Comment les modèles fictifs peuvent-ils générer des explications vertueuses ? Comment le faux peut-il expliquer le vrai ?

3 Les contrefactuels au travail

3.1 Le véritisme face aux arguments antiréalistes

J'ai développé dans le chapitre précédent ce qui constitue à mon sens la principale contribution de la conception fictionnelle des modèles : elle conduit à une sorte de stratégie structurale *divide et impera* qui permet de fournir une interprétation réaliste du succès empirique d'un modèle.

Dans ce cadre, les modèles sont considérés comme des structures modales, des mécanismes utilisés pour exemplifier un réseau de relations contrefactuelles. La structure modale est le squelette du modèle, les hypothèses fictives et les idéalizations sont la chair qui rend le modèle plus facile à manipuler ou à interpréter. Plus important encore : la structure modale illustrée est partiellement indépendante des hypothèses fictives dans le sens où elle permet de fournir la base véridique à la justification

réaliste du succès empirique.

En manipulant le pendule simple, on trouve une relation mathématique pour la période qui est indépendante de la masse du pendule, et ce malgré des hypothèses sciemment fausses. C'est cette structure contre-factuelle qui est primordiale. Lorsqu'un modèle est empiriquement adéquat, il faut se demander : qu'est-ce qui fait le travail de prédiction ? Lorsqu'on explique avec des modèles, la question est : qu'est-ce qui fait le travail explicatif ? Expliquer un phénomène en termes de validité des principes de génération (les règles du jeu fictionnel) n'est pas possible : ces principes sont faux en toute connaissance de cause. Alors si quelque chose fonde l'explication, c'est la structure modale elle-même, héritée des lois.

Si l'on demande une explication de l'indépendance de la période du pendule par rapport à la masse, on peut présenter le modèle du pendule simple, montrer comment l'équation $T = 2\pi\sqrt{l/g}$ est dérivée et expliquer pourquoi elle s'applique au pendule cible. Dans ce contexte, les idéalizations, les abstractions et tous les processus fictifs ont une valeur épistémique non pas parce qu'ils constituent l'explication, mais parce qu'ils aident à rendre le modèle calculable en expliquant sa structure contre-factuelle. Comme le formule Bokulich :

Fictional models can succeed in offering genuine explanations by correctly capturing relevant patterns of counterfactual dependence and licensing correct inferences. (Bokulich 2016).

Le modèle permet des inférences à la cible, et c'est ce qui fonde son pouvoir explicatif. Cependant, cela n'est pas suffisant pour obtenir une position véritiste, car la valeur empirique des inférences à la cible ne jus-

tifie pas véridiquement le modèle : l'efficacité des inférences ne plaide pas pour leur vérité. C'est le même problème qui émerge avec la représentation : une représentation épistémique permet des inférences à la cible, mais ces inférences ne justifient pas le modèle en tant que description fidèle du système cible.

Explorer le réel par le biais de faussetés semble toujours une stratégie dangereuse si nous ne sommes pas en mesure de distinguer entre les explications valides et invalides, car il existe toujours une menace de méta-induction pessimiste (PMI) à l'œuvre : dans l'incapacité de justifier plus précisément le succès explicatif des modèles, il est possible – et même probable – qu'ils soient remplacés par des modèles qui offrent de meilleures explications d'une nature tout à fait différente. Le parallèle se prolonge : le véritisme, comme forme particulière de conception orientée vers la vérité, se trouve face aux mêmes types d'arguments que le réalisme.

Pour illustrer cela, Bokulich prend l'exemple d'une fiction non explicative : le modèle des épicycles du système solaire. D'un point de vue purement fictionnel et contrefactuel, ce modèle et le modèle héliocentrique newtonien sont sur un pied d'égalité, mais le premier ne parvient pas à capturer les bonnes dépendances contrefactuelles. Il s'agit bien sûr d'un raisonnement circulaire si nous ne pouvons pas justifier ce que « bonnes » veut dire, c'est-à-dire s'il est impossible de formuler une distinction entre ce qui dans le modèle fonde son succès et ce qui est de l'ordre de l'idéalisation. C'est ce que Bokulich appelle *justificatory step* (étape justificative), mais elle affirme qu'en rendre compte de manière générale est impossible, sapant ainsi les espoirs véritistes :

However, what does it mean to say that a fictional representation is adequate? It has to be more than mere empirical

adequacy. Unfortunately, here is where I think abstract philosophical generalizations purporting to hold across all model explanations give out, and one needs to turn to the nitty-gritty details of the science in question. What is to count as an adequate fictional representation is something that has to be negotiated by the relevant scientific community and will depend on the details of the particular science, the nature of the target system, and the purposes for which the scientists are deploying the model. (Bokulich 2012, p. 734)

Il serait donc impossible de déterminer un critère général d'adéquation d'un modèle, les particularités de chaque cas, de chaque modèle et de chaque phénomène modélisé, ainsi que les normes méthodologiques de la communauté interférant avec l'ambition de rassembler tous ces éléments dans une unique description.

Je pense cependant que cette objection ne constitue pas une menace fondamentale pour ma proposition, et ce pour deux raisons. Premièrement, trouver le terrain commun interprétable de manière véridique pour les explications et fournir un critère général d'acceptabilité sont deux choses différentes. En ce sens, mon exposé est minimal, puisqu'il propose une condition nécessaire mais non suffisante, et il n'est bien entendu pas exclu que d'autres critères interviennent dans la sélection d'une explication. Je reconnais tout à fait que l'établissement d'une condition suffisante peut dépendre du contexte de la découverte. Deuxièmement, je ne vois pas de raison supplémentaire de douter du fait que les ressources fournies par la conception fictionnelle peuvent aider dans l'entreprise véritiste. Au contraire, l'objection de Bokulich soulève un argument que je me suis déjà proposé de contourner dans les chapitres qui précèdent. La réponse aux inquiétudes de type méta-induction pes-

simiste soulevées par Bokulich peut suivre le même chemin que celui que j'ai présenté précédemment.

La base vériste de l'explication du modèle se trouve dans la manière dont les lois génèrent le squelette du modèle, et le modèle est évalué en manipulant cette structure pour qu'elle génère des propositions testables empiriquement. La liberté heuristique rendue possible par l'approche fictionnelle met en avant les procédures d'idéalisation et de postulation d'hypothèses simplifiées ou impossibles : cela permet la manipulation et l'exemplification de la structure contrefactuelle au cœur du modèle. Fournir une interprétation des dépendances contrefactuelles est le fondement de toute explication : ce qui explique l'indépendance de la période du pendule par rapport à la masse est l'existence d'une interprétation de la dépendance contrefactuelle exemplifiée par le modèle. À la question « qu'est-ce qui explique ce phénomène qui concerne les pendules ? », une réponse aura la forme : « ce modèle est très efficace à reproduire le phénomène et met en avant cette formule qui elle-même est confirmée pour une large gamme d'expérience différentes ». Or, la dépendance contrefactuelle qui sert de support à l'explication est dérivée de lois plus générales qui sont à leur tour utilisées comme justifications : c'est la mécanique newtonienne qui explique le comportement du pendule.

Je rappelle qu'il n'est pas nécessaire d'interpréter ces dépendances comme des causes et des conséquences. Ma proposition reste agnostique et laisse place à différentes interprétations et à différents niveaux d'engagements ontologiques.

3.2 Explications contrefactuelles

Selon le développement des chapitres précédents, les modèles sont des machines qui génèrent des propositions contrefactuelles à propos de leurs cibles. Quelles explications peut-on espérer formuler à partir de tels objets ?

Verreault-Julien soutient que les modèles peuvent fournir des *how-possibly explanations*, qui sont des propositions qui prennent la forme $\diamond(p \text{ because } q)$ (Verreault-Julien 2019). Face à la question de savoir comment des modèles hautement idéalisés comportant des impossibilités peuvent fournir une telle explication, Verreault-Julien insiste sur l'importance de faire une distinction claire entre les propositions-modèle (ce que j'ai appelé propositions ficitonnelles, c'est-à-dire vraies à l'intérieur du modèle) et les propositions-monde (les propositions extérieures, qui portent sur la cible). Il écrit à ce propos :

What model propositions (e.g., unrealistic assumptions) do is to give reasons to believe in the truth of the possibility claim (Verreault-Julien 2019, p. 22)

Selon lui, les modèles peuvent représenter des impossibilités tout en soutenant des affirmations de possibilité, qui sont des propositions non fictives sur le monde. Cela est dû à la nature contrefactuelle des modèles : les explications qu'ils fournissent se servent d'hypothèses fausses, et les propositions qui portent sur le monde en dépendent évidemment. Le modèle du pendule simple affirme l'indépendance de la période par rapport à la masse, ce qui ne peut se traduire dans la réalité par une possibilité. Le modèle exemplifie en ce sens un monde possible, dont il s'agit de déterminer la proximité avec le nôtre. Dans ma conception, c'est exactement le rôle que joue la manipulation du modèle, qui vise à établir la

robustesse non pas d'une proposition contrefactuelle, mais d'une structure à même de reproduire le comportement d'un phénomène. L'aspect dynamique des modèles est capital.

Les modèles génèrent donc des propositions contrefactuelles qui établissent la possibilité que les choses se passent de telle et telle façon dans la cible. C'est effectivement possible, puisque cela a lieu à l'intérieur du modèle.

Je suis d'accord avec la proposition de formalisation de Verreault-Julien, mais seulement si on ne considère pas des propositions de la forme $\diamond(p \text{ because } q)$ pour p et q donnés, mais une fonction contrefactuelle plus générale qui attribue une valeur de q à une valeur de p . Le modèle n'est pas utilisé pour étayer un ensemble d'affirmations précises sur la cible, mais une structure contrefactuelle censée être intégrée dans les phénomènes que l'on souhaite expliquer. Ces fonctions ont souvent une arité plus élevée, car elles relient plusieurs valeurs physiques. Par exemple, la formule que l'on obtient en manipulant le pendule simple n'est pas réductible à la forme que propose Verreault-Julien. De plus, l'utilisation de *because* comme relation semble suggérer une relation causale or j'ai montré dans le chapitre précédent la difficulté à assimiler les contrefactuels à la causalité, en particulier quand de nombreux facteurs entrent en jeu (Mackie 1974).

Plus généralement, et pour éviter tout engagement métaphysique à l'égard du terme *because*, je pense que les modèles convenablement applicables soutiennent des affirmations de la forme $\diamond(p \sim q)$, où \sim est une relation entre les quantités illustrées par le modèle. Le pendule simple représente un objet impossible mais soutient néanmoins l'affirmation modale qui relie un certain nombre de quantités applicables aux pendules du monde réel lorsqu'elles sont interprétées comme étant la masse,

la longueur du fil, etc.

La nécessité d'une distinction claire entre les propositions fictives et les propositions du monde est également étayée par le processus fictif lui-même. Comme Verrault-Julien l'explique clairement, il est possible pour un modèle de soutenir non-trivialement des revendications de possibilité seulement si on suppose au départ $\diamond p$ et $\diamond q$. Ce sont les hypothèses qui sont à la base du modèle et qui lui permettent de formuler des propositions. Mais bien sûr, ces hypothèses fictives peuvent être impossibles, entendues comme des propositions à propos du monde, d'où la nécessité de distinguer les affirmations internes au modèle des affirmations externes, comme je l'ai suggéré plusieurs fois dans ce qui précède. Le modèle est construit à partir d'hypothèses fausses du monde, mais qui établissent une possibilité : dans le modèle, tout se passe *comme si* p et q , c'est-à-dire que l'on affirme la possibilité d'avoir p et q .

Dans la même veine, Sugden suggère que les similitudes posées entre le modèle et la réalité peuvent autoriser des inférences inductives (voir par exemple ses (Sugden 2000) et (Sugden 2013, p. 240)). Il donne un argument qui prend la forme suivante : connaître $p \rightarrow q$ dans le modèle et p et q dans le monde donne une bonne raison d'inférer que $p \rightarrow q$ dans le monde. Comme dans tout argument inductif, une similitude entre des observations spécifiques est posée et, dans le cas de l'inférence basée sur un modèle, c'est la similitude entre le modèle et le monde qui soutient l'induction. La deuxième étape de l'argument de Sugden est analogue à ce que j'appelle l'applicabilité, et l'accent mis sur la structure contrefactuelle peut soutenir l'étape inférentielle en fournissant un terrain pour la justification de la similarité.

Sugden considère également que la modalité est une caractéristique importante des modèles :

So what might increase our confidence in such inferences? I want to suggest that we can have more confidence in them, the greater the extent to which we can understand the relevant model as a description of how the world could be. (Sugden 2000, p. 24)

Les inférences à la cible sont justifiées par une ressemblance entre le modèle et sa cible. D'une possibilité exemplifiée par le modèle, on cherche à expliquer un phénomène réel. Or, tout l'enjeu est de justifier cette inférence, de fonder la ressemblance.

Il existe de nombreuses façons dont le monde pourrait être, et Sugden propose la *crédibilité* comme moyen de les trier, mais, selon lui, il n'est pas évident de déterminer comment les mondes crédibles (c'est-à-dire les mondes qui pourraient être réels) traiteraient les impossibilités postulées. Les mondes impossibles sont-ils crédibles?

Je ne suis pas d'accord avec Sugden lorsqu'il écrit :

To put this another way, the real world is equivalent to an immensely complicated model : it is the limiting case of the process of replacing the simplifying assumptions of the original model with increasingly realistic specifications. (Sugden 2000, p. 23)

Dans la conception fictionnelle, il est impossible de concevoir le monde comme un modèle très complexe dont on pourrait se rapprocher en intégrant à nos modèles des hypothèses plus réalistes. La fiction permet des impossibilités qui ne sont pas seulement des simplifications ou des idéalizations.

La crédibilité d'un monde ne peut donc pas être définie à partir de la possibilité de le complexifier pour se rapprocher du monde supposé réel.

Cette procédure n'a aucun sens dans le contexte fictionnel. En revanche, un modèle peut être qualifié de crédible s'il est compatible avec une certaine loi, qui régit et décrit la structure contrefactuelle qui constitue son squelette. La crédibilité n'empêche donc pas les impossibilités pour peu que l'on introduise la distinction adéquate entre la loi qui génère le modèle et le contenu fictionnel du modèle. Or, cette distinction apparaît naturellement dans l'approche fictionnelle.

Les lois expriment une structure de dépendances contrefactuelles. Ces relations contrefactuelles sont exemplifiées dans des modèles, et l'exemplification peut impliquer toutes sortes de processus fictionnels tels que, mais sans s'y limiter, des idéalizations. C'est de là que les *felicitous falsehoods* d'Elgin tirent leurs vertus épistémiques. Postuler des impossibilités (c'est-à-dire des incompatibilités avec les lois) est la principale liberté de modélisation offerte par la vision fictionnelle, mais comme les lois générant les modèles expriment un réseau de dépendances contrefactuelles, seules ces relations doivent être compatibles pour que le modèle soit acceptable. Pour utiliser les termes de Sudgen : les mondes impossibles peuvent être crédibles, à condition que le modèle soit robuste à la manipulation et que la structure contrefactuelle reste applicable aux phénomènes modélisés.

Manipuler un modèle fictif ne signifie pas seulement dériver de nouveaux énoncés fictifs à partir d'anciens, mais aussi être capable de donner une interprétation de ces propositions en termes de cible. Utiliser le modèle de manière adéquate, c'est aussi savoir ce qui ne doit pas être supposé vrai, ou même ce qui n'est pas possible du tout. Si un utilisateur de pendule simple prétend que l'adéquation empirique du modèle soutient l'existence de masses ponctuelles, il est évident qu'il n'utilise pas le modèle correctement, la non-existence de masses ponctuelles n'étant

pas un argument contre le modèle lui-même. Il est possible de se tromper sur certains aspects ou propriétés des entités décrites, voire sur leur existence, mais la conservation de la structure contrefactuelle par la manipulation du modèle est le point central de la compréhension et des explications que l'on tire des modèles.

Les propositions du modèle ne doivent pas être prises pour argent comptant. Un modèle est une entité dynamique, un moteur contrefactuel qui génère des ensembles de propositions sur un système physique et fournit une justification pour les inférences à la cible. Ce sont des « descriptions of how the world could be » (Sugden 2013, p. 241).

4 Le retour du véritisme ?

Comme dans de nombreux sous-domaines de l'épistémologie et de la philosophie des sciences, les positions réalistes et orientées vers la vérité ont été confrontées à de solides arguments de toutes parts. Si, comme je le prétends, le véritisme de l'explication peut être conservé dans la vision fictionnelle des modèles, quelles sont les limites que ce cadre de pensée lui impose ?

De manière analogue au réalisme scientifique, je pense qu'une position doit remplir deux conditions pour être considérée comme véritiste :

1. Il doit y avoir une certaine forme de correspondance ou de similarité entre le monde et le modèle en jeu, et cette relation doit être, au moins partiellement, le fondement de l'explication. En particulier, l'adéquation empirique pure, l'acceptation par la communauté scientifique ou la conformité aux normes scientifiques sont insuffisantes si elles ne sont pas fondées sur la relation de corres-

pondance.

2. La correspondance doit être dotée d'un mécanisme de démarcation : il est nécessaire d'être en mesure de lier la validité de l'explication basée sur le modèle aux autres vertus épistémiques, telles que l'adéquation empirique, et la différence entre une explication acceptable et une explication inacceptable doit être constructible dans le cadre.

Premièrement, comme je l'ai expliqué dans les sections précédentes, je considère que la structure contrefactuelle exemplifiée par le modèle est le fondement de l'explication. Il est clair que le modèle doit reproduire et prédire les données empiriques afin d'être évalué de manière véridique. La robustesse pour une gamme de valeurs d'entrée, c'est-à-dire la manipulabilité du modèle, offre un terrain pour justifier les inférences à la cible. C'est là que la relation de correspondance entre en jeu : comme dans l'exemple d'argument inférentiel de Sugden, une similitude donne à l'argument son squelette. Je prétends que cette similarité est une similarité de la structure contrefactuelle héritée des lois dans le modèle.

Deuxièmement, la leçon à tirer du point de vue de la fiction est que la démarcation entre les explications valides et invalides est possible, mais que nous devons nous abstenir d'interpréter de manière véridique tout ce qui ne fait pas partie de la structure contrefactuelle de nos explications. Par exemple, en utilisant un modèle d'atomes de silogène pour expliquer des phénomènes dans la physique des solides, il est possible parler de ces atomes *comme* s'ils étaient réels, mais la fondation vériste de l'explication se situe dans les processus de mécanique quantique en jeu dans la physique de l'état solide. Expliquer des phénomènes de

la physique des solides avec le silogène serait une mauvaise utilisation de la fiction du silogène, tout comme affirmer que la morale d'une fable n'est valable que pour les animaux anthropomorphes que l'histoire décrit.

En fin de compte, y a-t-il une place pour le véritisme de l'explication dans la vision fictionnelle des modèles ? Je pense que oui, mais l'exploration des aspects des modèles en tant que jeux fictionnels dans des mondes crédibles impose des limites au type de véritisme que l'on peut espérer atteindre.

Tout d'abord, la vision fictionnelle peut s'accommoder de la vertu épistémique des idéalizations et des approximations : celles-ci contribuent épistémiquement aux explications et à la compréhension en simplifiant les choses, et la liberté fictionnelle dans la construction des modèles offre toutes les stratégies heuristiques aux scientifiques pour ce faire.

L'énigme centrale des approches fictionnelles véritistes peut être clarifiée en portant une attention particulière aux aspects modaux des modèles. Les modèles sont des descriptions de mondes crédibles que nous pouvons manipuler pour générer, lorsqu'elles sont convenablement interprétées, des propositions sur des systèmes physiques cibles. La manipulation d'un modèle signifie que nous explorons la robustesse de la structure contrefactuelle à travers une gamme de valeurs d'entrée. Des processus fictifs, comme la postulation d'entités inexistantes, peuvent aider à cette exploration, mais la structure contrefactuelle elle-même est immunisée contre la fiction et reste interprétable de manière véridique.

La relation entre les lois et les modèles apparaît également clairement dans le contexte fictionnel. Les lois génèrent des modèles dans le sens où elles sont l'expression des dépendances contrefactuelles que le

modèle contient. La structure est le squelette, les hypothèses fictives sont la chair qui facilite le raisonnement et l'interprétation. Lorsque le modèle s'avère robuste, on peut dire que la structure contrefactuelle reproduit celle du système physique. C'est là le terrain privilégié des explications, et il reste véritiste dans le sens où c'est l'accord entre les deux structures qui renforce les inférences inductives que nous tirons du modèle pour construire des explications. Cela explique aussi pourquoi les lois sont si importantes pour les explications. L'observation de régularités doit être expliquée, la loi est l'explication, et les modèles font le lien entre les lois et le monde en les rendant manipulables.

Les modèles ne sont donc pas l'explication, mais ils sont nécessaires pour pouvoir exprimer les explications.

Chapitre 6

Conclusion et perspectives

Ce chapitre de conclusion est l'occasion de mettre en évidence les étapes les plus importantes du développement des chapitres qui précèdent afin d'en dégager les thèses principales. Cela me permettra ensuite de présenter quelques prolongements possibles et perspectives de cette recherche.

1 Le réalisme scientifique à l'épreuve de la fiction

Cette thèse s'inscrit dans le cadre du débat entre réalisme et antiréalisme. Plus précisément, j'ai défendu une position réaliste à partir de la conception fictionnelle des modèles, qui se base sur une analogie entre modèles scientifiques et œuvres de fiction.

C'est une stratégie que l'on pourrait penser à première vue contre

productive, voire paradoxale.

En effet, le réalisme scientifique est la thèse épistémologique qui tend à interpréter le succès des meilleures théories ou modèles comme un signe de la vérité des énoncés scientifiques qui se voient confirmés par l'expérience. De nombreux arguments antiréalistes, parmi lesquels la méta-induction pessimiste, visent à saper l'optimisme épistémique qui caractérise le réalisme : il serait impossible d'établir un lien suffisamment solide entre l'adéquation empirique dont jouit une théorie ou un modèle et la vérité d'énoncés, particulièrement quand ils portent sur des entités inobservables.

Dès lors, baser une défense réaliste sur le concept de fiction semble voué à l'échec, dans la mesure où la notion même de fiction désamorce immédiatement la possibilité de construire une approche qui soit orientée vers la vérité. C'est moins dû à l'association habituelle entre fiction et fausseté – rapidement remise en cause dans les différentes théories de la fiction que j'ai abordées – qu'à la mise en avant d'hypothèses délibérément fausses dans la construction des modèles scientifiques.

L'argument antiréaliste prend typiquement la forme suivante :

1. Les modèles scientifiques peuvent être évalués sur des critères empiriques : adéquation empirique, pouvoir prédictif, etc.
2. Les modèles contiennent des idéalizations, des approximations et des abstractions qui simplifient radicalement les systèmes physiques qu'ils modélisent.

Il s'en suit donc :

3. Les modèles ne peuvent être évalués sur base véridique.

La stratégie réaliste standard est de réfuter, ou tout du moins d'affaiblir, la deuxième étape de ce raisonnement en minimisant le rôle des

idéalisations, ou en affirmant qu'elles n'ont qu'un apport marginal dans la production de la connaissance et qu'elles sont *in fine* éliminables de la pratique scientifique. Les idéalisations ne constitueraient qu'une étape de la procédure de modélisation, elles seraient vouées à être « dé-idéalisées » pour permettre au modèle de se rapprocher d'une description véridique des systèmes physiques modélisés.

Pourtant, un exemple comme celui des atomes de silogène résiste à cette analyse. Ce type d'atomes inexistant participe activement au succès empirique du modèle dans lequel ils figurent, mais ils ne sont pas dé-idéalisables, dans le sens où ils ne peuvent absolument pas être considérés comme une approximation d'atomes qui existent.

L'approche fictionnelle prend à bras le corps ce problème en considérant ces inventions, ces objets imaginaires et ces processus impossibles comme le cœur de ce qui fait l'intérêt des modèles scientifiques.

L'opposition entre la notion de fiction et les approches orientées vers la vérité se retrouve dans les conceptions fictionnalistes ou instrumentalistes des sciences. Pour les tenants de ces positions, il y a des « fictions utiles » dont le pouvoir heuristique est parfaitement légitime, mais dont les espoirs réalistes sont vains. En ce sens, les fictions participent au travail scientifique car elles permettent de nouvelles découvertes et facilitent la manipulation de systèmes physiques ou le développement de nouvelles technologies, mais elles ne sont que des outils. En tant qu'outils, elles ne peuvent qu'être jugées plus ou moins utiles pour la tâche qui leur incombe, mais certainement pas plus ou moins vraies.

Ma première étape dans cette thèse, après avoir défini les critères qu'une position philosophique doit remplir pour être qualifiée de réaliste, a été de montrer que la conception fictionnelle des modèles, entendue comme l'approche de la modélisation qui compare les modèles

scientifiques à des jeux de *make-believe* tels que définis par Kendall Walton, ne se réduit pas à un fictionnalisme instrumentaliste.

Dans un premier temps, il s'agissait d'examiner les conséquences de l'usage de la fiction que propose Vaihinger dans sa théorie du *comme si*. J'ai montré que l'on pouvait voir la théorie de Walton comme un prolongement du *comme si* qui lui confère suffisamment de ressources pour aborder plus finement les problèmes épistémologiques que pose l'utilisation de modèles en sciences.

Roman Frigg et James Nguyen ont développé, à partir de Walton, une théorie de la représentation par les modèles, dans laquelle les modèles sont largement conçus comme des *représentations épistémiques*. Leur utilisation du concept de *make-believe* vise essentiellement à comprendre la manière dont les modèles représentent certains systèmes physiques. Si le problème du réalisme et celui de la représentation ont des points communs, mon but n'est pas de traiter directement de la représentation, mais d'examiner dans quelle mesure l'outil fictionnel peut être utile au problème général du réalisme.

L'idée centrale est que les modèles ne se contentent pas de simplifier les choses en isolant les variables, en idéalisant les processus ou en abstrayant certaines propriétés. Ils invitent à penser *comme si* les choses se passaient de telle ou telle manière, tout comme une œuvre de fiction est une incitation à imaginer des situations, des personnes et des lieux.

Le *make-believe* de Walton fonctionne en deux étapes : des principes de génération sont d'abord stipulés, puis de nouvelles choses sont imaginées à partir de *props*, qui guident et soutiennent ces actes imaginatifs. L'exemple qu'utilisent Frigg et Nguyen est celui du jeu où l'on stipule au départ qu'un morceau de bois représente un ours. Une fois cette fiction générée, chaque morceau de bois observé pousse à imaginer que

c'est un ours qui se trouve là. De la même manière, le pendule simple est une fiction dans le sens où, en lisant sa description théorique (qui sert de *prop*), nous imaginons une masse ponctuelle oscillant au bout d'une corde non extensible et sans masse, même si nous savons que les masses ponctuelles et les cordes sans masse n'existent pas dans la réalité.

Une fois stipulé, le modèle est ensuite étudié, investigué et manipulé. De nouvelles vérités fictionnelles (c'est-à-dire des propositions vraies dans le modèle) sont dérivées des principes de génération. Le pouvoir représentationnel du modèle vient du fait que ces propositions fictionnelles permettent des inférences à la cible, c'est-à-dire que certaines vérités du modèle sont attribuables au système qu'il modélise.

Dès lors, le problème du réalisme se pose en ces termes : si les modèles sont conçus comme des fictions au sens du jeu de *make-believe*, peut-on les considérer comme des descriptions véridiques des systèmes physiques ? Et si oui, par quel mécanisme ?

Conformément aux deux thèses réalistes principales, les thèses sémantique et épistémologique, j'ai développé une réponse en deux temps.

Tout d'abord, le réalisme requiert que ce soit la cible qui confère au modèle son aspect sémantique. Cela se fait au moyen de la notion d'applicabilité : un modèle est applicable à la cible s'il existe une interprétation du contenu du modèle en termes de la cible. Cela permet de définir ce qu'« être un modèle de » veut dire. Le modèle du pendule simple est un modèle de pendule parce que le contenu du modèle peut être interprété avec des termes qui concernent les modèles réels. Les termes abstraits m , l et g du modèle sont respectivement compris comme une masse, une longueur de fil et l'accélération de la pesanteur.

La notion d'applicabilité permet de contrer les arguments selon lesquels face à un modèle vu en tant que fiction, l'utilisateur est libre de son

interprétation, ou du moins il n'est contraint que par certaines normes méthodologiques ou par les habitudes d'une certaine communauté de chercheurs. L'applicabilité permet de préciser quels aspects du modèle doivent être interprétés de manière véridique, et comment ils doivent l'être. J'ai pour illustrer ce point pris l'exemple d'une fourchette utilisée pour représenter la situation géographique d'une route. Il semblait qu'il n'y ait là qu'un acte de stipulation totalement libre, et que par conséquent n'importe quel objet pouvait représenter n'importe quel système dans n'importe quel contexte, pour autant que l'on se mette d'accord sur certaines règles. Cela introduit dans l'analyse un aspect contextuel qui sape la nécessité réaliste de fonder l'interprétation du modèle sur la nature de sa cible. Avec la notion d'applicabilité à disposition, il est possible d'affirmer que la fourchette constitue un modèle car elle peut être interprétée comme une route (route qui appartient au système cible), ce qui entraîne que c'est sa disposition géométrique qui compte dans le modèle, et non pas le fait qu'elle soit en métal ou qu'elle ait quatre dents. La disposition géométrique entre la fourchette représente donc l'orientation de la route réelle, ce qui permet de connaître la direction à prendre.

Une fois le problème sémantique réglé – l'interprétation du modèle se fait bien sur base de sa cible –, reste le problème épistémologique. Ma proposition fonctionne en deux étapes : d'une part, établir un lien entre l'efficacité empirique du modèle et une forme de correspondance avec la réalité, et, d'autre part, offrir un critère de justification de la validité du modèle, c'est-à-dire fournir une démarcation entre les modèles acceptables et ceux qui doivent être rejetés.

Ma proposition de réponse prend la forme suivante : les idéalizations que contiennent les modèles permettent de mettre en évidence certaines

relations entre des quantités données. C'est là le rôle premier des idéalizations, qui permettent de simplifier les systèmes réels pour mettre en évidence ces relations simples. Le schéma de Walton appliqué aux modèles scientifiques identifie les principes de génération aux lois et hypothèses générales qui décrivent le modèle et permettent son exploration, sa manipulation, c'est-à-dire la génération de propositions qui portent sur la cible (si le modèle est applicable).

Ces propositions générées par le modèle sont contrefactuelles, ce qui est une conséquence immédiate de la conception fictionnelle et de la possibilité d'introduire des hypothèses fausses dans la génération du modèle. Il en suit que les dépendances mises en évidence par le modèle sont des dépendances contrefactuelles. La manipulation du modèle revient à chercher dans quelle mesure il est capable de s'adapter à différentes situations en reproduisant le comportement de la cible. Cette manipulabilité du modèle constitue son aspect dynamique. Un modèle n'est pas un objet que l'on étudie, c'est une machine dont on se sert pour générer certaines propositions qui portent sur la cible.

La robustesse du modèle, qui correspond à sa capacité à reproduire le comportement de la cible pour une gamme variée de conditions initiales qu'on lui donne en entrée, permet de s'assurer de la capacité des dépendances contrefactuelles à reproduire le comportement de la cible. Les modèles qui ont du succès dans cette entreprise exemplifient avec succès une structure contrefactuelle semblable à celle du phénomène modélisé.

La nature contrefactuelle du modèle permet également d'explicitier le rôle des lois, qui sont souvent mises en avant, pour leur grande généralité et simplicité, dans la pratique scientifique. Les lois génèrent un modèle parce qu'elles sont associées dans le schéma waltonien aux prin-

cipes de génération de la fiction. Dans le contexte contrefactuel, elles apparaissent naturellement dans la justification des dépendances au cœur du modèle.

La base véridique des modèles se situe donc dans la structure de dépendances contrefactuelles exemplifiée par le modèle, une fois que sa robustesse a été vérifiée par manipulation. Il est alors possible de fournir une réponse aux questions réalistes.

Ce qui explique le succès empirique du modèle est la correspondance des dépendances contrefactuelles qu'il exemplifie avec celles du phénomène modélisé. Un modèle incapable de reproduire le comportement contrefactuel d'une cible, c'est-à-dire un modèle qui n'est pas robuste, sera rejeté (sur base véridique, ce qui ne veut pas dire qu'il ne peut pas avoir une autre utilité).

Ces deux points répondent aux demandes réalistes spécifiées plus haut et sont le cœur de la position que j'ai développée au cours de cette thèse.

2 Perspectives

La position que j'ai défendue, une conception réaliste et fictionnelle à propos des modèles, constitue un cadre de pensée assez large pour traiter de plusieurs problèmes épistémologiques. J'ai pris en exemple la notion d'explication et le problème associé du véritisme, mais toutes les conceptions philosophiques orientées vers la vérité peuvent a priori être approchées par ma conception fictionnelle et modale.

C'est un des prolongements possibles de ce travail. Je pense notamment à la notion d'objectivité en sciences, sujette à un très large et vivant débat, et dont une partie des sujets mobilisés par mon raisonnement, ty-

piquement la correspondance entre une construction théorique et un système réel, serait aisément applicable. Aborder l'objectivité à partir de l'aspect modal de la connaissance scientifique me semble en tout cas être une piste attrayante.

Le problème de la représentation pourrait également gagner à intégrer une approche contrefactuelle. La notion de représentation épistémique telle qu'on la trouve par exemple chez Toon, Frigg et Nguyen met en avant son caractère fictionnel, mais il serait intéressant de pousser l'analyse en insistant sur l'aspect modal des stratégies représentationnelles.

Je ne vois pas non plus d'objection à l'idée d'aborder la représentation artistique par le biais contrefactuel, bien que je ne sois absolument pas certain de la pertinence des conclusions que l'on pourrait tirer d'une telle analyse. Une première étape pourrait être de débusquer le raisonnement contrefactuel là où il se cache déjà, par exemple dans l'analyse que propose Walton. Cela constituerait, si l'approche se révèle judicieuse, un intéressant et nouveau point de contact entre l'épistémologie et des considérations issues des débats en esthétique.

Si j'ai principalement pris des exemples issus de la mécanique newtonienne, c'est avant tout pour leur simplicité, mon but principal étant de rendre le plus explicite le mécanisme fictionnel de modélisation. L'approche fictionnelle ne se limite pourtant pas à ces cas, et les exemples en physique abondent. Un prolongement possible du travail effectué ici serait de multiplier les études de cas qui emploient l'approche fictionnelle. J'ai brièvement abordé le cas de la physique des solides ou de l'électromagnétisme, mais tant les interprétations de la mécanique quantique que les modèles de cosmologie, pour ne citer qu'eux, peuvent être ap-

prochés en tant que fictions.¹

De la même manière, si mon parcours me porte à traiter d'exemples venant de la physique, je ne vois aucune raison pour laquelle l'approche fictionnelle ne serait pas généralisable aux autres disciplines scientifiques. On trouve de toute évidence des hypothèses idéalisatrices en biologie – par exemple dans le modèle proie-prédateur qui pousse à imaginer deux espèces isolées du reste de l'écosystème – ou en sociologie – le modèle de la ségrégation de Schelling qui cherche à reproduire des dynamiques de population en formulant des hypothèses très simples sur les préférences individuelles de voisinage. La question de la nature des explications fournies par ces modèles, lorsqu'ils sont expérimentalement confirmés, se pose de la même manière qu'en physique. De plus, s'il est souvent avancé que la biologie et les sciences sociales ne sont pas réductibles à des lois comme c'est le cas en physique, il me semble que la définition minimale des lois qui émerge de l'approche fictionnelle s'applique aussi bien à ces cas. En effet, rien dans mon utilisation de ce terme ne nécessite de formalisation mathématique. Les lois figurent dans les explications en tant que justification des dépendances contrefactuelles. Ainsi, un sociologue qui cherche à expliquer les résultats scolaires construit, consciemment ou non, un modèle analogue à ce que j'ai décrit dans cette thèse. L'identification de certains paramètres pertinents (classe sociale, niveau d'études des parents, etc.) amène à imaginer une situation et un mécanisme de sélection, puis à chercher à en déduire les conséquences,

1. Il est important de noter que l'approche fictionnelle des modèles ne permet pas seulement de capturer les mécanismes explicatifs ou de justification des modèles. L'hypothèse de l'atome primitif de Lemaître est une fiction qui ne cherche pas directement à expliquer ou reproduire des phénomènes, mais elle est pourtant heuristiquement fructueuse, car l'image qu'elle porte suggère des pistes de réflexion et participe à produire de nouvelles idées, indépendamment de toute velléité représentationnelle. De la même manière que les modèles scientifiques ont d'autres fonctions que la représentation et l'explication, l'analogie avec la fiction peut se prolonger à ces autres fonctions. Je remercie Dominique Lambert pour avoir attiré mon attention sur cet exemple.

et enfin à examiner si les vérités fictionnelles de cette école imaginaire correspondent à ce qui est observé. Si le sociologue se refuse à édifier des lois, il cherche cependant à dégager des tendances et à les expliquer par un mécanisme exemplifié par son modèle. S'il détecte des dépendances contrefactuelles robustes (par exemple une forte corrélation entre le niveau d'études des parents et le niveau d'étude qu'atteignent leurs enfants), leur justification faisant appel au mécanisme décrit dans le modèle prendra la forme minimale de ce que j'ai appelé loi.² Une future exploration de la fécondité de l'approche fictionnelle dans ces domaines est un des principaux prolongements du présent travail

J'ai esquissé une ébauche de définition plus précise des modèles et de leur rapport avec les théories et les lois qui gagnerait à être développée. Ce sont trois concepts souvent mobilisés en philosophie des sciences, et il me semble important de les utiliser dans un cadre qui rende explicite leurs relations étroites. Au début de cette thèse, j'utilisais la définition la plus générale et intuitive de modèle pour ensuite la préciser au fil des chapitres, selon ce que la conception fictionnelle, puis l'introduction de la notion de propositions contrefactuelles, induisaient. J'en suis arrivé à concevoir les modèles comme des entités dynamiques capables de générer des propositions contrefactuelles, et à les situer dans une arborescence selon leur niveau de généralité, les lois fondamentales ayant le degré de généralité le plus haut et les phénomènes, conçus comme des modèles de données, le plus faible. Cette image doit être précisée et complétée, notamment en spécifiant le rôle des théories dans ce contexte et peut-être en développant un formalisme qui permette de décrire plus

2. La nature statistique de ces corrélations et la forme probabiliste que prendraient les explications sont tout à fait compatibles avec le cadre fictionnel et ne contredisent en rien la définition des lois qui découlent de l'approche fictionnelle. L'aspect statistique intervient dans la forme des lois et dans ce qui est attendu d'une vérification expérimentale, mais pas dans la justification elle-même du modèle.

efficacement les relations entre modèles et la relation modèle-cible.

J'ai également fait appel à la distinction entre contexte de découverte et contexte de justification, en insistant sur le fait que le problème du réalisme concerne essentiellement le second, et il serait intéressant de questionner à nouveau la valeur de cette distinction avec l'approche fictionnelle et modale à l'esprit. En particulier, la notion de contrefactuels joue un rôle dans la justification de la validité d'un modèle, mais il faudrait examiner dans quelle mesure la méthodologie et la pratique scientifiques intègrent – ou non – cet aspect. C'est un problème distinct, mais complémentaire, de celui du réalisme, qui intéressera cependant peut-être plus le sociologue des sciences que le philosophe.

Bibliographie

- Ankeny, Rachel A. et Sabina Leonelli (juin 2011). « What's so special about model organisms? » In : *Studies in History and Philosophy of Science Part A* 42.2, p. 313-323. DOI : 10 . 1016/ j . shpsa . 2010 . 11 . 039.
- Appiah, Anthony (2017). *As if idealization and ideals. idealization and ideals*, p. 218. ISBN : 9780674975002.
- Armatte, Michel (déc. 2005). « La notion de modèle dans les sciences sociales : anciennes et nouvelles significations ». In : *Mathématiques et sciences humaines* 172. DOI : 10 . 4000/ msh . 2962.
- Barberousse, Anouk et Pascal Ludwig (2009). « Models as Fictions ». In : *Fictions in science : philosophical essays on modeling and idealization*. Sous la dir. de Mauricio Suárez. New York : Routledge.
- Barwich, Ann-Sophie (déc. 2013). « Science and Fiction : Analysing the Concept of Fiction in Science and its Limits ». In : *Journal for General Philosophy of Science* 44.2, p. 357-373. DOI : 10 . 1007 / s10838 - 013 - 9228 - 2.
- (oct. 2014). « Fiction in Science ? Exploring the Reality of Theoretical Entities ». In : *Defending Realism*. De Gruyter, p. 291-310. DOI : 10 . 1515/9781614516651 . 291.

- Bayard, Pierre (2008). *Qui a tué Roger Ackroyd?* Éditions de Minuit, p. 192. ISBN : 9782707320438.
- Bell, J. S. (oct. 1992). « Six possible worlds of quantum mechanics ». In : *Foundations of Physics* 22.10, p. 1201-1215. DOI : 10 . 1007 / bf01889711.
- Berto, Francesco (jan. 2017). « Impossible Worlds and the Logic of Imagination ». In : *Erkenntnis* 82.6, p. 1277-1297. DOI : 10 . 1007 / s10670-017-9875-5.
- Black, Max (jan. 1962). *Models and Metaphors*. Cornell University Press. DOI : 10 . 7591/9781501741326.
- Bokulich, Alisa (2009). « Explanatory Fictions ». In : *Fictions in science : philosophical essays on modeling and idealization*. Sous la dir. de Mauricio Suárez. New York : Routledge.
- (déc. 2012). « Distinguishing Explanatory from Nonexplanatory Fictions ». In : *Philosophy of Science* 79.5, p. 725-737. DOI : 10 . 1086 / 667991.
- (2016). « Fiction As a Vehicle for Truth : Moving Beyond the Ontic Conception ». In : *The Monist* 99.3, p. 260-279. DOI : 10 . 1093 / monist / onw004.
- Butterfield, Jeremy (2002). « Some Aspects of Modality in Analytical Mechanics ». In : *Symposium of the Philosophy of Science Association*.
- Cartwright, Nancy (1983). *How the Laws of Physics Lie*. Oxford University Press. DOI : 10 . 1093/0198247044 . 001 . 0001.
- (1994). *Nature's Capacities and Their Measurements*. Oxford University Press, USA, p. 280. ISBN : 9780198235071.
- (2008). *The Dappled World*. Cambridge University Press. 260 p. ISBN : 0521644119.

- Cassini, Alejandro (oct. 2018). « Modelos sin target ». In : *ArtefaCToS. Revista de estudios sobre la ciencia y la tecnología* 7.2, p. 185. DOI : 10.14201/art201872185209.
- Chakravartty, Anjan (2001). « The Semantic or Model-Theoretic View of Theories and Scientific Realism ». In : *Synthese* 127.3, p. 325-345.
- Cohen, Ariel (2000). « The King Of France Is, In Fact, Bald ». In : *Natural Language Semantics* 8.4, p. 291-295. ISSN : 0925854X, 1572865X.
- Contessa, Gabriele (jan. 2007). « Scientific Representation, Interpretation, and Surrogate Reasoning ». In : *Philosophy of Science* 74.1, p. 48-68. DOI : 10.1086/519478.
- Currie, Gregory (2016). « Models As Fictions, Fictions As Models ». In : *The Monist* 99.3, p. 296-310. DOI : 10.1093/monist/onw006.
- Elgin, Catherine Z. (sept. 2017). *True Enough*. MIT Press Ltd. 352 p. ISBN : 0262036533.
- Engel, Pascal (20 juin 2015). *Le réalisme kitsch*. Sous la dir. de Zilsel. URL : <https://zilsel.hypotheses.org/2103> (visité le 26/01/2023).
- Fine, Arthur (1993). « Fictionalism ». In : *Midwest Studies in Philosophy* 18, p. 1-18. DOI : 10.1111/j.1475-4975.1993.tb00254.x.
- (1998). « Fictionalism ». In : *Routledge Encyclopedia of Philosophy*. Sous la dir. d'E. Craig. T. 3. Routledge, p. 667-668. DOI : 10.4324/9780415249126-q035-1.
- (2009). « Fictionalism ». In : *Fictions in science : philosophical essays on modeling and idealization*. Sous la dir. de Mauricio Suárez. New York : Routledge.
- Frappier, Mélanie, Letitia Meynell et James Robert Brown (2012). *Thought experiments in philosophy, science, and the arts*. Routledge.

- French, Steven (déc. 2003). « A Model-Theoretic Account of Representation (Or, I Don't Know Much about Art...but I Know It Involves Isomorphism) ». In : *Philosophy of Science* 70.5, p. 1472-1483. DOI : 10.1086/377423.
- (1^{er} fév. 2017). *The Structure of the World : Metaphysics and Representation*. Oxford University Press. 416 p. ISBN : 0198776667.
- French, Steven et James Ladyman (2003). « Remodelling Structural Realism : Quantum Physics and the Metaphysics of Structure ». In : *Synthese* 136.1, p. 31-56. DOI : 10.1023/a:1024156116636.
- Frigg, Roman (2009). « Models and fiction ». In : *Synthese* 172.2, p. 251-268. DOI : 10.1007/s11229-009-9505-0.
- Frigg, Roman et James Nguyen (2016). « The Fiction View of Models Reloaded ». In : *The Monist* 99.3, p. 225-242. DOI : 10.1093/monist/onw002.
- (3 sept. 2020). *Modelling Nature : An Opinionated Introduction to Scientific Representation*. Springer International Publishing. 260 p. ISBN : 3030451526.
- (2021). « Seven Myths About the Fiction View of Models ». In : *Logic, Epistemology, and the Unity of Science*. Springer International Publishing, p. 133-157.
- Frigg, Roman et Fiora Salis (2017). « Of Rabbits and Men : Fiction and Scientific Modelling ». In : *Fictionalism in Philosophy*. Sous la dir. de Bradley Armour-Garb et Fred Kroon. Oxford University Press.
- Frigg, Roman et Ioannis Votsis (2011). « Everything you always wanted to know about structural realism but were afraid to ask ». In : *European Journal for Philosophy of Science* 1.2, p. 227-276. DOI : 10.1007/s13194-011-0025-7.

- Gallais, Matthieu (2012). « Modèles scientifiques comme fictions : Propriétés et fonctions hintikkiennes ». In : *Klesis - Revue philosophique*. Concept(s) et fiction(s) 23.
- Giere, Ronald N. (déc. 2004). « How Models Are Used to Represent Reality ». In : *Philosophy of Science* 71.5, p. 742-752. DOI : 10 . 1086 / 425063.
- (2009). « Why Scientific Models Should Not Be Regarded as Works of Fiction ». In : *Fictions in science : philosophical essays on modeling and idealization*. Sous la dir. de Mauricio Suárez. New York : Routledge.
- Godfrey-Smith, Peter (déc. 2006). « The strategy of model-based science ». In : *Biology and Philosophy* 21.5, p. 725-740.
- Grüne-Yanoff, Till (déc. 2013). « Appraising Models Nonrepresentationally ». In : *Philosophy of Science* 80.5, p. 850-861. DOI : 10 . 1086 / 673893.
- Hacking, Ian (oct. 1983). *Representing and Intervening*. Cambridge University Press. DOI : 10 . 1017 / cbo9780511814563.
- Harre, Rom (1993). *Laws of nature*. Duckworth, p. 128. ISBN : 0715624644.
- Held, Carsten (2009). « When Does a Scientific Theory Describe Reality? » In : *Fictions in science : philosophical essays on modeling and idealization*. Sous la dir. de Mauricio Suárez. New York : Routledge.
- Hesse, Mary B. (1967). *Models and Analogies in Science*. University of Notre Dame Press. ISBN : 9780268001827.
- Hochberg, Herbert (1970). « Strawson, Russell, and the King of France ». In : *Philosophy of Science* 37.3, p. 363-384. ISSN : 00318248, 1539767X. URL : <http://www.jstor.org/stable/186465> (visité le 18/04/2023).
- Hume, David (1748). *Enquete sur l'entendement humain*. Trad. par Michel Malherbe. Paris : Vrin.

- Ichikawa, Jonathan et Benjamin Jarvis (nov. 2007). « Thought-experiment intuitions and truth in fiction ». In : *Philosophical Studies* 142.2, p. 221-246. DOI : 10 . 1007 / s11098 - 007 - 9184 - y.
- Jandrić, Andrej (2014). « "The king of France is bald" reconsidered : a case against Yablo ». In : *Philosophical Studies : An International Journal for Philosophy in the Analytic Tradition* 169.2, p. 173-181. ISSN : 00318116, 15730883.
- Jebeile, Julie et Ashley Graham Kennedy (oct. 2015). « Explaining with Models : The Role of Idealizations ». In : *International Studies in the Philosophy of Science* 29.4, p. 383-392. DOI : 10 . 1080 / 02698595 . 2015 . 1195143.
- Kind, Amy (avr. 2022). « Introduction : exploring the limits of imagination ». In : *Synthese* 200.2. DOI : 10 . 1007 / s11229 - 022 - 03669 - z.
- Klein, Étienne (2020). « Théorisation et images : le « chemin de pensée » de Boltzmann ». In : *Raison présente* N°214-215.2, p. 47. DOI : 10 . 3917 / rpre . 214 . 0047.
- Knuuttila, Tarja (déc. 2005). « Models, Representation, and Mediation ». In : *Philosophy of Science* 72.5, p. 1260-1271. DOI : 10 . 1086 / 508124.
- Kripke, Saul (1963). « Semantical Considerations on Modal Logic ». In : *Acta Philosophica Fennica* 16, p. 83-94.
- Kripke, Saul A. (mars 1959). « A completeness theorem in modal logic ». In : *Journal of Symbolic Logic* 24.1, p. 1-14. DOI : 10 . 2307 / 2964568.
- Ladyman, James (1998). « What is structural realism? » In : *Studies in History and Philosophy of Science Part A* 29.3, p. 409-424. DOI : 10 . 1016 / s0039 - 3681 (98) 80129 - 5.

- Ladyman, James et al. (2009). *Every Thing Must Go Metaphysics Naturalized. Metaphysics Naturalized*. Oxford University Press, p. 358. ISBN : 9780199573097.
- Lange, Marc (déc. 1996). « Laws of nature, cosmic coincidences and scientific realism ». In : *Australasian Journal of Philosophy* 74.4, p. 614-638. DOI : 10 . 1080/00048409612347561.
- Laudan, Larry (1981). « A Confutation of Convergent Realism ». In : *Philosophy of Science* 48.1, p. 19-49. ISSN : 00318248, 1539767X. URL : <http://www.jstor.org/stable/187066>.
- (1984). « Realism without the Real ». In : *Philosophy of Science* 51.1, p. 156-162. DOI : 10 . 1086/289171.
- Lawler, Insa (déc. 2019). « Scientific understanding and felicitous legitimate falsehoods ». In : *Synthese* 198.7, p. 6859-6887. DOI : 10 . 1007/s11229-019-02495-0.
- Levy, Arnon (déc. 2020). « Models and Fictions : Not So Similar after All? » In : *Philosophy of Science* 87.5, p. 819-828. DOI : 10 . 1086/710616.
- Lewis, David (1973). « Causation ». In : *Journal of Philosophy* 70.17, p. 556-567. DOI : 10 . 2307/2025310.
- (1986). *On the Plurality of Worlds*. Blackwell Publishing Limited, p. 276.
- (2001). *Counterfactuals*. Blackwell Publishing Limited, p. 168. ISBN : 9780631224952.
- Mackie, John L. (1974). *The Cement of the Universe. A Study of Causation (Clarendon Library of Logic and Philosophy)*. Oxford University Press, USA, p. 340. ISBN : 9780198246428.
- Mäki, Uskali (2011). « The Truth of False Idealizations in Modeling ». In : *Models, Simulations, and Representations*. Sous la dir. de Paul Humphreys et Cyrille Imbert. Routledge.

- Maxwell, James Clerk (1861). « On Physical Lines of Force ». In : *The Scientific Papers of James Clerk Maxwell*. Sous la dir. de W. D. Editor Niven. T. 1. Cambridge Library Collection - Physical Sciences. Cambridge University Press, 451–513.
- Morrison, Margaret (1995). « Capacities, Tendencies and the Problem of Singular Causes ». In : *Philosophy and Phenomenological Research* 55.1, p. 163. DOI : 10 . 2307 / 2108315.
- (2009). « Fictions, Representations, and Reality ». In : *Fictions in science : philosophical essays on modeling and idealization*. Sous la dir. de Mauricio Suárez. New York : Routledge.
- Nawar, Tamer (oct. 2019). « Veritism refuted? Understanding, idealization, and the facts ». In : *Synthese* 198.5, p. 4295-4313. DOI : 10 . 1007 / s11229 - 019 - 02342 - 2.
- Pincock, Christopher (jan. 2021). « A Defense of Truth as a Necessary Condition on Scientific Explanation ». In : *Erkenntnis*. DOI : 10 . 1007 / s10670 - 020 - 00371 - 9.
- Poincaré, Henri (1905). *La science et l'hypothèse*. Paris : Flammarion.
- Potochnik, Angela (17 nov. 2017). *Idealization and the Aims of Science*. University of Chicago Press. 288 p.
- Poznic, Michael (2016). « Make-Believe and Model-Based Representation in Science : The Epistemology of Frigg's and Toon's Fictionalist Views of Modeling ». In : *Teorema : International Journal of Philosophy* 35.3, p. 201-218.
- Psillos, Stathis (1999). *Scientific realism : how science tracks truth*. London New York : Routledge.
- (1^{er} juin 2009). *Knowing the Structure of Nature : Essays on Realism and Explanation*. Springer Nature. 230 p. ISBN : 0230007112.

- Purves, Gordon Michael (déc. 2012). « Finding truth in fictions : identifying non-fictions in imaginary cracks ». In : *Synthese* 190.2, p. 235-251. DOI : 10 . 1007 / s11229 - 012 - 0144 - 5.
- Putnam, Hilary (déc. 1981). *Reason, Truth and History*. Cambridge University Press. DOI : 10 . 1017 / cbo9780511625398.
- Rahman, Shahid (2011). « Idealizations as Prescriptions and the Role of Fiction in Science ». In : *Cadernos de Filosofia* 2011/2, p. 123-160. URL : <https://shs.hal.science/halshs-00713268>.
- Regt, Henk W. de (2017). *Understanding Scientific Understanding*. Oxford University Press, Incorporated, p. 304. ISBN : 9780190652913.
- Rouse, Joseph (2009). « Laboratory Fictions ». In : *Fictions in science : philosophical essays on modeling and idealization*. Sous la dir. de Mauricio Suárez. New York : Routledge.
- Russell, Bertrand (1905). « On Denoting ». In : *Mind* XIV.4, p. 479-493. DOI : 10 . 1093 / mind / xiv . 4 . 479.
- Ruyant, Quentin (fév. 2020). « Semantic realism in the semantic conception of theories ». In : *Synthese* 198.8, p. 7965-7983. DOI : 10 . 1007 / s11229 - 020 - 02557 - 8.
- Saatsi, Juha (2016). « Models, Idealisations, and Realism ». In : *Models and Inferences in Science*. Sous la dir. de F. Sterpetti, E. Ippoloti et T. Nickles. Springer.
- Salis, Fiora (2016). « The Nature of Model-World Comparisons ». In : *The Monist* 99.3, p. 243-259. DOI : 10 . 1093 / monist / onw003.
- Salis, Fiora et Roman Frigg (2016). « Capturing the Scientific Imagination ». In : *The Scientific Imagination*. Sous la dir. de Peter Godfrey-Smith et Arnon Levy. New York, USA : Oxford University Press.

- Schuppert, Guillaume (2021). *Retour à la mimêsis : La philosophie de Kendall L. Walton. La philosophie de Kendall L. Walton*. PU Rennes, p. 384. ISBN : 9782753580978.
- Smart, J. J. C. (juill. 1985). « Laws of Nature and Cosmic Coincidences ». In : *The Philosophical Quarterly* 35.140, p. 272. DOI : 10 . 2307 / 2218906.
- Soler, Lena (2013). « Qu'est-ce qu'un modèle scientifique ? Des caractéristiques du modèle qui importent du point de vue de l'enseignement intégré de science et de technologie ». In : *Spirale. Revue de recherches en éducation* 52.1, p. 177-214. DOI : 10 . 3406 / spira . 2013 . 1067.
- Spindel, Philippe (2001). *Mécanique tome 1. mécanique newtonienne*. Gordon & Breach, p. 234. ISBN : 9782847030044.
- Strawson, P. F. (1950). « On Referring ». In : *Mind* 59.235, p. 320-344. ISSN : 00264423, 14602113. URL : <http://www.jstor.org/stable/2251176> (visité le 18/04/2023).
- Suárez, Mauricio (déc. 2004). « An Inferential Conception of Scientific Representation ». In : *Philosophy of Science* 71.5, p. 767-779. DOI : 10 . 1086/421415.
- Sugden, Robert (jan. 2000). « Credible worlds : the status of theoretical models in economics ». In : *Journal of Economic Methodology* 7.1, p. 1-31. DOI : 10 . 1080/135017800362220.
- (sept. 2013). « How fictional accounts can explain ». In : *Journal of Economic Methodology* 20.3, p. 237-243. DOI : 10 . 1080/1350178x . 2013 . 828872.
- Sullivan, Emily et Kareem Khalifa (jan. 2019). « Idealizations and Understanding : Much Ado About Nothing ? » In : *Australasian Jour-*

- nal of Philosophy* 97.4, p. 673-689. DOI : 10 . 1080 / 00048402 . 2018 . 1564337.
- Suppe, Frederick (1977). *The Structure of scientific theories*. University of Illinois Press, p. 818. ISBN : 0252006550.
- (1989). *The semantic conception of theories and scientific realism*. University of Illinois Press, p. 475. ISBN : 025201605X.
- Suppes, Patrick (sept. 1960). « A comparison of the meaning and uses of models in mathematics and the empirical sciences ». In : *Synthese* 12.2-3, p. 287-301. DOI : 10 . 1007 / bf00485107.
- Suárez, Mauricio (2009). « Scientific Fictions as Rules of Inference ». In : *Fictions in science : philosophical essays on modeling and idealization*. Sous la dir. de Mauricio Suárez. New York : Routledge.
- (2010). « Fictions, Inference and Realism ». In : *Fictions and Models : New Essays*. Sous la dir. de John Woods. Munich : Philosophia Verlag.
- Teller, Paul (2009). « Fictions, Fictionalization, and Truth in Science ». In : *Fictions in science : philosophical essays on modeling and idealization*. Sous la dir. de Mauricio Suárez. New York : Routledge.
- Thomson-Jones, Martin (mai 2009). « Missing systems and the face value practice ». In : *Synthese* 172.2, p. 283-299. DOI : 10 . 1007 / s11229-009-9507-y.
- Toon, Adam (mars 2010). « Novel approaches to models ». In : *Metascience* 19.2, p. 285-288. DOI : 10 . 1007 / s11016-010-9388-1.
- (1^{er} jan. 2012). *Models as Make-Believe*. Palgrave Macmillan. 148 p.
- (2016). « Fictionalism and the Folk ». In : *The Monist* 99.3, p. 280-295. DOI : 10 . 1093 / monist / onw005.
- Vaihinger, Hans (1923). *La Philosophie du « comme si »*. Trad. par Christophe Bouriau (2008). Éditions Kimé.

- Van Fraassen, Bas C. (1^{er} oct. 1980). *The Scientific Image*. Oxford University Press. 248 p. ISBN : 0198244274.
- Verreault-Julien, Philippe (fév. 2019). « How could models possibly provide how-possibly explanations? » In : *Studies in History and Philosophy of Science Part A* 73, p. 22-33. DOI : 10 . 1016 / j . shpsa . 2018 . 06 . 008.
- Walton, Kendall (1990). *Mimesis as make-believe : on the foundations of the representational arts*. Cambridge, Mass : Harvard University Press. ISBN : 9780674576032.
- Wheeler, Billy (2018). *Idealization and the Laws of Nature*. Springer, p. 128.
- Winsberg, Eric (déc. 2006a). « Handshaking Your Way to the Top : Simulation at the Nanoscale ». In : *Philosophy of Science* 73.5, p. 582-594. DOI : 10 . 1086 / 518328.
- (sept. 2006b). « Models of Success Versus the Success of Models : Reliability without Truth ». In : *Synthese* 152.1, p. 1-19. DOI : 10 . 1007 / s11229 - 004 - 5404 - 6.
- (2009). « A Function for Fictions : Expanding the Scope of Science ». In : *Fictions in science : philosophical essays on modeling and idealization*. Sous la dir. de Mauricio Suárez. New York : Routledge.
- Worrall, John (1989). « Structural Realism : The Best of Both Worlds? » In : *dialectica* 43.1-2, p. 99-124. DOI : 10 . 1111 / j . 1746 - 8361 . 1989 . tb00933 . x.
- Wray, K. Brad (19 mars 2019). *Resisting Scientific Realism*. Cambridge University Press. 238 p. ISBN : 1108415210.
- Yablo, Stephen (2001). « Go Figure : A Path through Fictionalism ». In : *Midwest Studies in Philosophy* 25.1, p. 72-102. DOI : 10 . 1111 / 1475 - 4975 . 00040.

Yablo, Stephen (déc. 2019). « Models and Reality ». In : *The Scientific Imagination*. Oxford University Press, p. 128-153.