

Les modèles comme fictions

Pascal Ludwig, Anouk Barberousse

► **To cite this version:**

Pascal Ludwig, Anouk Barberousse. Les modèles comme fictions. Philosophie, Les éditions de Minuits, 2000, 68, pp.16 -43. <ijn_00000501>

HAL Id: ijn_00000501

https://jeannicod.ccsd.cnrs.fr/ijn_00000501

Submitted on 7 May 2004

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Les modèles comme fictions¹

Anouk Barberousse

Pascal Ludwig

Que dit-on [...] d'un savant sans imagination ? Qu'il a appris tout ce qui, ayant été enseigné, pouvait être appris, mais qu'il ne trouvera pas les lois non encore devinées. L'imagination est la reine du vrai, et le *possible* est une des provinces du vrai.

(Baudelaire, *Salon de 1859*)

Qu'est-ce qui différencie un *modèle* scientifique d'une *théorie* scientifique ? Dans nombre de cas, cette distinction semble aller de soi ; cependant, elle n'est pas toujours parfaitement claire. On en voudra pour preuve le long article de 1913 où Bohr présente ce qu'il appelle indifféremment son "modèle" ou sa "théorie" de l'atome d'hydrogène, destiné à former le point de départ de ce qui allait devenir, 12 ans plus tard, la mécanique quantique. Qu'est-ce qui fait que l'on n'hésite jamais à qualifier la mécanique quantique de *théorie*, alors que "l'atome de Bohr" reçoit parfois le titre de *modèle* ? C'est à ces questions que le présent article veut répondre de façon générale.

Nous défendons la thèse selon laquelle les modèles scientifiques doivent être interprétés comme des *fictions*. Cette thèse a été également soutenue par N. Cartwright dans *How the Laws of Physics Lie*² ; cependant, le sens que nous lui donnons ici est notablement différent. Notre point de départ est l'idée selon laquelle les modèles scientifiques, de même que les théories, sont des *représentations*. Parmi toutes les représentations que l'on peut former, certaines possèdent la propriété de représenter directement les circonstances dans lesquelles elles ont été produites, c'est-à-dire d'être porteuses de propositions que l'on peut évaluer directement comme vraies ou fausses dans ces circonstances. Les reportages, les chroniques économiques, les descriptions contenues dans les guides touristiques sérieux sont des exemples de telles représentations. Cependant, toutes les représentations ne possèdent pas cette propriété. Par exemple, un tableau de Giorgione, un film de Chaplin ou un roman de Jane Austen, qui sont autant de représentations, n'ont pas pour but de représenter *directement* les différentes circonstances dans

¹ Les idées défendues dans cet article ont été présentées devant divers audiences. Nous remercions Jérôme Dokic et Daniel Averbuch pour de précieuses remarques, ainsi que les participants du séminaire "Modèles, modélisations en sciences physiques et en sciences sociales".

² Cf. Cartwright, 1983.

lesquelles ils ont été produits□ Chaplin ne cherche pas dans ses films à communiquer des propositions susceptibles d'être vraies ou fausses des États-Unis de l'entre deux guerres. Nous fondant sur ces exemples, nous appellerons "fictions" les représentations dont la fonction n'est pas de représenter directement les circonstances dans lesquelles elles sont produites. Nous soutenons que les modèles appartiennent à cette seconde catégorie de représentations, alors que les théories appartiennent à la première□ nous soutenons que les modèles, contrairement aux théories, sont des fictions.

1 L'imagination dans les jeux de faire-semblant et dans l'activité modélisatrice

1.1 Quelques remarques sur l'imagination et la fiction

Il n'est pas question, dans le cadre de cet article, d'entrer dans les méandres des théories contemporaines de la fiction. Les détails de ces théories ne sont d'ailleurs guère pertinents pour notre propos. Aussi nous contenterons-nous d'une conception certainement grossière, mais selon nous néanmoins adéquate, de la fiction : nous nommerons "fiction", au sens large, toute représentation dont la fonction est de susciter, chez un interprète, l'imagination de certains contenus.³ On peut considérer un roman, ou une pièce de théâtre, comme des représentations dont le but est de susciter, chez les lecteurs ou les spectateurs, des projets imaginatifs ayant certains contenus intentionnels. A la suite de Walton, nous considérerons que telle est également la fonction d'autres artefacts culturels comme les films, les tableaux, les sculptures.

Quelle est la nature du contenu intentionnel des épisodes imaginatifs suscités par de tels artefacts□ La plupart des auteurs qui se sont intéressés à cette question considèrent que l'on peut identifier ces contenus à des propositions, exprimables par des *phrases* d'une langue naturelle. On doit en effet constater qu'à nombre d'états d'imagination correspondent des attributions d'attitudes vis-à-vis d'un contenu exprimé par une phrase. Ainsi, pour décrire le projet suscité par tel ou tel passage de *La Chartreuse de Parme*, l'on pourra dire que le lecteur doit imaginer que *Fabrice a accompli telle ou telle action lors de la bataille de Waterloo*. De tels énoncés constituent cependant un sérieux défi pour l'approche propositionnaliste des contenus d'imagination. Considérons en effet la clause :

³De fait, cela correspond à une prise de parti en faveur des approches de la fiction qui partent du concept d'imagination. Ces approches ont été récemment défendues par de nombreux auteurs, parmi lesquels Walton (cf. Walton, 1990) et Currie sont les plus notables. Currie résume de façon fort claire la thèse centrale de ces approches :

"L'intention directrice d'un auteur de fiction est que le lecteur forme un certain type d'attitude vis-à-vis des propositions énoncées dans la narration. Il s'agit de l'attitude qu'on décrit, d'une façon assez vague, à l'aide des termes d'"implication imaginative", ou, de façon plus adéquate, de faire-semblant (*make-believe*). L'auteur a l'intention que nous fassions semblant que l'histoire, telle qu'elle est énoncée, soit vraie." (Cf. Currie, 1990, p.□8).

(1) Fabrice était présent lors de la bataille de Waterloo.

Supposons, à la suite de Russell et de Kripke, que la fonction essentielle d'un nom propre soit de représenter son porteur. Puisque le nom propre "Fabrice del Dongo" n'a pas de porteur en français, cette supposition conduit immédiatement à la conséquence selon laquelle le sujet de (1) ne représente rien, et donc que (1) n'exprime aucune proposition.

Une telle conséquence n'est évidemment pas acceptable si l'on veut identifier le contenu de projets imaginatifs à l'aide de propositions. Pour résoudre ce problème, la plupart des philosophes de la fiction adoptent une sémantique descriptiviste *ad hoc* pour les noms propres à l'intérieur des contextes de fiction. Telle ne sera pas notre stratégie. Nous proposons plutôt l'hypothèse suivante :

Le contenu intentionnel d'une phrase de fiction doit être identifié à une propriété de contextes, à savoir la propriété instanciée par tous les contextes dans lesquels un interprète de la fiction peut à juste titre imaginer se trouver conformément au sens de la phrase.

Cette hypothèse permet de résoudre de façon élégante les problèmes posés par les noms fictionnels. Un lecteur peut en effet imaginer se trouver dans un contexte dans lequel un nom, vide dans le contexte réel, possède un référent, instanciant les propriétés qui sont attribuées au porteur du nom dans le monde de la fiction. Elle possède d'autre part un avantage important pour l'entreprise qui est la nôtre : elle permet de comprendre comment certains énoncés fictionnels peuvent exprimer d'apparentes contradictions. Considérons un nouvel énoncé, (2) :

(2) Paul imagine que Napoléon et Talleyrand ne sont qu'une seule et même personne.

Cette phrase semble spécifier un projet imaginatif, qui pourrait être mis en oeuvre à l'intérieur d'une fiction. Pourtant, il n'existe aucun monde possible dans lequel Napoléon et Talleyrand ne seraient qu'une seule et même personne — si l'on accepte comme nous le faisons, encore une fois, la sémantique kripéenne des noms propres⁴. D'un point de vue logique, (2) exprime donc une antilogie. On a souvent remarqué, par ailleurs, que de nombreuses fictions ne pouvaient être vraies dans aucun monde possible. Ici encore, de nombreux philosophes ont proposé une solution *ad hoc* à ce problème, en acceptant la légitimité d'un concept de monde impossible.

On peut s'étonner du caractère démesuré de cette réaction. Mettons à part les fictions dont les parties sont explicitement contradictoires. Dans la plupart des cas intéressants, les fictions n'expriment pas tant des contradictions logiques explicites que des contenus qui sont en contradiction avec certaines théories qui d'une part semblent bien confirmées, et qui surtout doivent être acceptées pour donner un sens bien défini aux concepts mêmes qui sont mobilisés par la fiction. L'exemple le plus souvent discuté à cet égard est celui de *La machine à remonter le temps* de H. G. Wells. Il est clair que, si l'on tire certaines conséquences des actions décrites dans cet ouvrage, on arrive rapidement à des contradictions patentes — du moins si l'on

⁴ Rappelons que dans ce cadre, qui est largement accepté en philosophie du langage, la vérité d'un énoncé d'identité comme "a = b" implique sa vérité nécessaire. Réciproquement, deux individus dont on peut énoncer de façon vraie qu'ils diffèrent dans un monde, ce qui est le cas de Napoléon et de Talleyrand dans notre monde, doivent différer dans tous les mondes possibles.

présuppose que les concepts physiques et métaphysiques les plus usuels sont valables. Ainsi, le même personnage peut être dit avoir existé en 1935, et, au futur, ne pas pouvoir exister à cette date s'il tue son père après avoir remonté le temps. Notre approche permet au moins d'envisager une stratégie de solution à cette énigme, qui ne nécessite pas de faire appel à un concept aussi tératologique que celui de monde impossible. Rien ne dit que les mêmes théories physiques aient été développées dans les différents contextes spécifiés par la fiction, ni surtout que les symboles utilisés par ces théories possèdent les mêmes dénnotations que dans notre monde. Nous soutenons qu'une implication de *La machine à remonter le temps* est que la signification des concepts fondamentaux des théories vraies dans le monde du roman est telle que les contradictions apparentes qui naissent d'une application de nos théories aux situations décrites disparaissent. Il n'est pas difficile de montrer le caractère plausible de cette thèse — le simple fait que nous puissions interpréter le roman comme décrivant un monde cohérent plaide en sa faveur. Le processus cognitif par lequel nous arrivons à une telle interprétation nous semble en fait d'une simplicité désarmante : nous mettons entre parenthèses l'ensemble des concepts susceptibles de mener à des contradictions — ou au moins nous nous refusons à utiliser ces concepts dans celles de leurs occurrences qui pourraient mener à de telles incohérences. Autrement dit, le sens des concepts physiques ou métaphysiques susceptibles de mener à une contradiction est laissé indéterminé, la seule contrainte clairement impliquée par la fiction, à leur sujet, étant qu'ils doivent pouvoir exprimer un ensemble cohérent de propositions. Insistons sur les différences qui existent entre notre approche de la fiction et certaines théories des énoncés contrefactuels. Ce que nous soutenons, c'est que la *signification* de certains concepts dépend du contexte dans lequel ces concepts sont employés. En un sens, cette thèse est trivialement vraie, puisque la signification des mots exprimant ces concepts dépend d'un langage, et que les faits qui fixent les conventions d'un langage peuvent être considérés comme faisant partie de l'ensemble des faits contextuels. Lorsque nous imaginons les contenus exprimés par une fiction comme *La machine à remonter le temps*, notre projet imaginaire ne porte pas seulement sur des circonstances contrefactuelles, qui seraient spécifiées à l'aide des concepts à notre disposition, une langue étant fixée — il porte également, de façon indirecte, sur la signification des concepts mêmes que nous utilisons pour spécifier les circonstances imaginées. Nous imaginons que nous remontons le temps, mais cela n'a pas nécessairement les conséquences contradictoires dénoncées par les physiciens, car le concept de temps lui-même peut être imaginé comme étant libéré des contraintes théoriques imposées par nos théories physiques.

Une autre caractéristique distingue nettement notre approche des contenus fictionnels de l'approche courante : nous considérons que ces contenus sont des propriétés de contextes *dans lesquels l'interprète de la fiction imagine être*. Une caractéristique essentielle d'un *contexte* est d'être organisé, pour ainsi dire, autour du sujet possible d'un acte de langage — ou, pour généraliser, d'un acte mental identifiable linguistiquement. La caractéristique sur laquelle nous venons d'attirer l'attention découle directement de notre théorie. Cette conséquence nous paraît heureuse, car elle semble correspondre à la phénoménologie de l'imagination. De fait, nous

soutenons qu'un interprète participe toujours au monde de la fiction qu'il imagine, au sens où il imagine les faits vrais dans un tel monde à partir d'un point de vue⁵. Nous n'entendons pas soutenir par là que l'on doit imaginer être soi-même *personnellement, en chair et en os*, dans le monde de la fiction que l'on imagine⁶. Nous voulons plutôt dire que l'on imagine toujours les contenus fictionnels à partir d'un *point de vue subjectif*, qui se trouve d'ailleurs parfois être le point de vue d'un des personnages d'un récit. Chaque situation imaginée doit être envisagée à partir d'une perspective donnée⁷.

La dernière remarque que nous ferons sur la fiction concerne son statut épistémique. Il est courant de considérer qu'une fiction, par opposition à un discours vrai, exprime intrinsèquement des propositions fausses. Nous nous élevons, à la suite de nombreux auteurs cette fois, contre cette vision simplificatrice. Imaginer, c'est former des pensées qui ne concernent pas directement la réalité, au moins en ceci qu'elles portent sur des contextes dans lesquels un certain nombre des propositions tenues pour vraies dans nos meilleures théories sont mises entre parenthèses. Toute fiction, du coup, implique une réorganisation de nos concepts — ce qui explique que certains symboles, dénués de signification dans leurs usages sérieux, puissent véhiculer un contenu à l'intérieur d'un monde fictionnel. Les fictions, même si elles ne portent pas directement sur notre monde, véhiculent un contenu. Ce contenu a ceci de particulier qu'il est exprimé à l'aide de concepts en quelque sorte idiosyncratiques, formés au sein de la fiction elle-même. Nous soutenons que de tels contenus peuvent jouer des rôles importants dans l'entreprise de production de connaissances. C'est une banalité, hélas souvent occultée, de dire que le contenu d'une fiction peut nous informer de façon indirecte sur notre monde. Nous aurons l'occasion à la fin de cet article d'indiquer de quelle façon les modèles nous informent, de façon indirecte également, sur leurs domaines d'objets.

1.2 L'interprétation des modèles comme fiction

Comme nous l'avons signalé, la thèse que nous défendons a déjà été soutenue par Nancy Cartwright, dans *How the Laws of Physics Lie*. Ainsi, Cartwright écrit-elle qu'"un modèle est

⁵ Ce point a été souligné par Berkeley, dans un passage justement fameux du *Premier dialogue entre Hylas et Philonous*. Voir l'édition française publiée sous la direction de G. Bryckman, PUF, pp. 68-69. L'argument de Berkeley est discuté par Williams, 1973 et par Peacocke, 1985.

⁶ Une personne, disons Pierre, peut imaginer être P sans imaginer que Pierre est P. L'un des auteurs peut par exemple imaginer être Napoléon, sans imaginer que Pascal Ludwig est Napoléon — ainsi peut-il imaginer être Napoléon, et décorer Pascal Ludwig. La raison de cette propriété inattendue des rapports d'attitudes *de se* serait trop longue à expliquer. Nous renvoyons aux travaux de David Lewis sur cette question. Cf Lewis, 1979 et Lewis, 1986.

⁷ Le fait que les philosophes et les théoriciens de la fiction aient concentré leur attention sur l'art du roman a considérablement obscurci cette caractéristique de l'activité imaginative. Pourtant, elle s'applique aussi au roman, d'une façon souvent cachée. Interpréter correctement un roman, selon nous, c'est toujours imaginer être dans un contexte où, au minimum, nous sommes informés de façon véridique sur les faits narrés. L'identité de la personne qui nous informe, ou du canal par lequel nous avons été informés, est parfois spécifiée par le contenu du roman — mais le plus souvent, ces détails sont indéterminés.

une oeuvre de fiction"⁸. Il s'agit plus dans ce texte, cependant, d'un slogan que d'une thèse soigneusement argumentée. Cartwright ne dit rien des caractéristiques précises des modèles qui, par opposition à celles des théories, devraient nous amener à les concevoir comme des fictions. A vrai dire, elle ne dispose pas vraiment, de son propre aveu, d'une théorie satisfaisante du contenu des fictions. L'identification des modèles à des fictions a donc dans son oeuvre un caractère principalement programmatique.

Notre but est de développer un argument en faveur de la thèse énoncée par Cartwright. Mais avant d'accomplir cette tâche, nous voudrions montrer le caractère plausible, *prima facie*, de cette thèse. Nous le ferons en soulignant quelques ressemblances entre modèles scientifiques et fictions. Comme les fictions, les modèles scientifiques sont des représentations. Ils possèdent un contenu représentationnel, puisqu'ils décrivent un certain domaine d'objets — ou plusieurs domaines différents — comme instanciant certaines propriétés et certaines lois. En cela, les modèles ressemblent aux théories scientifiques. Mais à l'inverse des théories, les modèles ne peuvent généralement pas être décrits comme visant à produire des descriptions vraies du monde. Pour de nombreuses raisons, sur lesquelles nous reviendrons en détail par la suite, considérer un modèle comme exprimant des vérités littérales a peu de sens. De même que les fictions, les modèles expriment des contenus qui ne se situent pas d'emblée dans la dimension du vrai et du faux□un modèle ne peut pas être rejeté au simple motif qu'il serait littéralement faux.

Cette ressemblance entre modèles et fictions, dont on reconnaîtra l'importance, ne doit pas masquer certaines dissemblances. On remarquera d'abord qu'un modèle possède une fonction essentiellement épistémique, contrairement aux fictions, même si comme nous l'avons vu, le but des fictions ne peut certainement pas être réduit au pur *divertissement* de l'esprit. Inversement, le rôle joué par l'activité modélisatrice dans la recherche de la vérité, dès lors qu'on reconnaît que les modèles ne sont pas immédiatement vrais ou faux, soulève des interrogations. D'autre part, on soulignera que les théories, de même que les données expérimentales, imposent des *contraintes* à l'activité modélisatrice. Même si le propre d'un modèle est de pouvoir contenir certains éléments incompatibles avec les meilleures théories disponibles à un moment donné, le jeu laissé à l'imagination dans la construction de modèles scientifiques est restreint, du moins en comparaison avec la fiction — l'importance accordée par cette activité aux *données expérimentales*⁹ en est le signe le plus apparent. On aurait tort, néanmoins, de croire que l'imagination, dans l'activité de création fictionnelle, serait libérée de toute contrainte. Les contextes dans lesquels nous faisons comme si nous étions projetés, dans nos projets d'imagination, sont constitués pour l'essentiel d'éléments empruntés au contexte réel.

⁸ Cf. Cartwright, 1983, p.□53.

⁹ Dans la dernière partie de cet article, nous montrons cependant que cette contrainte elle-même, contrairement à ce qu'un empiriste un peu strict pourrait être amené à penser, n'a elle-même rien d'absolu.

2 Que représentent les modèles scientifiques☐

La comparaison entre fictions et modèles semble motivée. Peut-on pour autant considérer que la thèse qui identifie ces deux types de représentations est justifiée☐ Afin de répondre à cette question, nous allons développer un argument qui vise à montrer qu'il ne semble pas exister d'analyse alternative convaincante à la nôtre, du moins en ce qui concerne de nombreux modèles scientifiques.

Afin de pouvoir présenter cet argument, une rapide clarification du concept de modèle s'impose. Comme nous y avons fait allusion plus haut, nous considérerons qu'un modèle est une représentation qui porte sur un domaine d'entités. Le but d'un modèle est d'établir des généralisations à propos de ces entités, qui sont exprimées, dans le cas des modèles quantitatifs, par des équations. Sous ces aspects, les modèles ne se distinguent pas des théories. La différence entre un modèle et une théorie réside dans les relations entre les généralisations que le modèle exprime et celles qui sont acceptées dans la communauté scientifique au moment où le modèle est construit. Les généralisations exprimées par une théorie entretiennent en effet d'autres relations avec les généralisations couramment acceptées. De même que les différentes hypothèses d'une même théorie, les hypothèses de théories différentes conçues comme complémentaires plutôt que comme concurrentes doivent être mutuellement compatibles. En revanche, les généralisations ou les hypothèses valables à l'intérieur d'un même modèle peuvent être en contradiction avec les théories scientifiques disponibles au moment où le modèle est construit, y compris avec les théories les mieux acceptées. Cette caractéristique des modèles a été souvent soulignée. Ainsi, Michael Redhead souligne-t-il que certains modèles au moins, ceux dont le rôle est de simplifier les calculs, sont en contradiction logique avec la théorie au sein de laquelle ils trouvent place¹⁰. Mais Redhead ne semble pas conscient du défi que représente cette propriété des modèles pour le philosophe des sciences. En général, la nature des entités sur lesquelles porte le modèle n'est pas explicitement déterminée *par le modèle lui-même*, mais bien plutôt par l'arrière-plan théorique dans lequel il prend place. Supposons donc qu'un modèle M porte sur des entités de type S , et supposons que la nature des entités du type S soit explicitée par une théorie T . Supposons de plus que certaines des hypothèses que M exprime soient en contradiction avec certaines des généralisations de T . Si nous nous demandons maintenant dans quelles situations le modèle s'applique véridiquement — quelles situations il décrit☐—, nous arrivons à la conclusion paradoxale selon laquelle il n'en décrit aucune. C'est un fait bien connu des logiciens, en effet, qu'aucune situation ne saurait rendre vraies deux propositions contradictoires. En conséquence, tous les modèles qui possèdent la propriété de contredire par certaines de leurs hypothèses des généralisations théoriques s'appliquant à leur domaine d'entités s'avèrent informationnellement équivalents☐ils ne véhiculent *aucune* information.

¹⁰ Cf. Redhead, 1980.

Cette conclusion n'est évidemment pas défendable. Le partisan d'une conception plus commune des modèles pourrait soutenir que nous avons commis une erreur dans notre argument. Bien sûr, admettra-t-il, certaines hypothèses de certains modèles contredisent des généralisations théoriques censées valoir de leur domaine d'entités. Mais on suppose précisément, lorsque l'on se place *dans le modèle*, que ces généralisations ne s'appliquent pas. Il y a trois manières d'interpréter une telle proposition. La première consiste à penser qu'on pourrait faire l'hypothèse, à propos d'une entité de type S soumise à une loi L , qu'elle n'est plus soumise à la loi dans certaines circonstances. Cette option semble peu plausible. Lorsque, dans un modèle classique de la matière, on fait comme si les particules n'étaient pas soumises à certaines des contraintes imposées par la mécanique quantique, on n'entend évidemment pas soutenir que *dans les circonstances particulières décrites par le modèle*, les lois de la mécanique quantique ne s'appliquent plus¹¹. Personne ne pense que de telles circonstances existent, ni même qu'elles *puissent physiquement* exister.

La deuxième interprétation envisageable de la proposition de notre adversaire est de renoncer au concept de modèle. Après tout, il arrive couramment, dans la pratique scientifique, d'avancer des hypothèses théoriques incompatibles entre elles. Le nier reviendrait à endosser une vision irénique du cours de l'histoire des sciences, faisant peu de cas du rôle des controverses et des désaccords théoriques. Peut-être les modèles incompatibles avec leur cadre théorique introduisent-ils en fait, par la bande, de *nouvelles hypothèses théoriques*. Nous considérons que cette interprétation est cohérente, mais empiriquement inadéquate. Il n'est tout simplement pas vrai que les scientifiques considèrent tous les modèles comme des théories à l'état naissant. Bien des modèles n'ont pas ce caractère prospectif, et en particulier tous ceux qui sont construits faute de puissance de calcul. On n'imagine pas quelqu'un refusant une modélisation d'un phénomène donné pour le seul motif qu'elle contredirait une théorie mieux confirmée. De fait, les modèles et les théories ne sont pas mis, dans la communauté scientifique, sur un pied d'égalité. Théorisation et modélisation sont deux activités qui possèdent leurs fonctions propres et le but d'une analyse philosophique des modèles devrait être de mettre en lumière ces fonctions, plutôt que de nier leur différence.

La troisième interprétation n'est en réalité qu'une reformulation de notre thèse centrale. Même dans la situation décrite par notre adversaire, il est possible *d'imaginer* que les généralisations énoncées par la théorie ne contredisent pas les hypothèses du modèle. Bien entendu, une telle affirmation a peu d'intérêt si l'on n'explicite pas la sémantique de l'imagination. Dans le cadre de notre approche, imaginer que les généralisations énoncées par la théorie ne contredisent pas les hypothèses du modèle revient à se placer dans un contexte où certains concepts de la théorie voient leur interprétation varier, de sorte que les prédictions de la théorie, appliquées au domaine d'entités du modèle, deviennent compatibles avec les hypothèses centrales du modèle. Un tel projet imaginatif s'apparente en tous points au projet normalement

¹¹ Un modèle de ce type est analysé en détail ci-dessous.

suscité par une fiction. Dans la fiction, nous imaginons que certains termes qui, dans le contexte réel, ne possèdent pas de dénotation, réfèrent à des individus en chair et en os. Dans un modèle, nous imaginons que certaines des théories que nous connaissons, et dont les prédictions contredisent de fait certaines hypothèses centrales du modèle, conduisent à des prédictions différentes de celles qu'elles impliquent habituellement. De la même façon que nous outrepassons, dans la fiction, les limites du langage naturel, en utilisant des symboles qui sont vides dans leur emploi normal, nous outrepassons dans un modèle certaines contraintes imposées par le langage de telle ou telle théorie disponible¹². Remarquons que dans un modèle, la théorie dont certaines prédictions posent problème continue d'exister en tant que telle. Elle continue de déterminer certaines catégories, applicables au domaine d'entités du modèle. Mais nous imaginons qu'elle diffère, par certains de ses concepts centraux, de la théorie réelle qui porte le même nom. En général, la nature de cette différence est laissée complètement indéterminée. Le modèle impose seulement à la signification imaginée des concepts théoriques qu'elle n'entre pas en conflit avec ses hypothèses. Cette caractéristique explique l'impression d'indétermination que suscitent de nombreux modèles scientifiques. Qu'est-ce exactement qu'une particule dans un modèle classique de la matière? Nul ne s'en préoccupe, car l'extension exacte du concept n'a pas de pertinence dans le modèle. Ce qui importe est que l'extension du concept dans le modèle n'est pas son extension normale.

3 Illustrations

3.1 Le modèle des Ehrenfest

A titre d'illustration de notre thèse concernant la nature des modèles scientifiques, nous nous proposons d'examiner en détail deux exemples de modèles tirés de la physique. Le premier a été élaboré dans le cadre de la mécanique statistique afin de mettre en lumière certaines propriétés très générales du phénomène de *thermalisation*. Considérons deux corps *A* et *B* de températures différentes, placés en contact thermique et isolés de tous les autres corps. La théorie qui rend compte du phénomène de thermalisation — c'est-à-dire d'égalisation des températures — qui a alors lieu est la thermodynamique, la science des échanges de chaleur. La thermodynamique prédit que la chaleur s'écoulera, en quelque sorte, du corps le plus chaud vers le corps le plus froid. Ce flux de chaleur est décrit dans cette théorie comme étant *unidirectionnel*. La thermodynamique, et plus particulièrement son Deuxième Principe, prédit en outre que la différence entre les deux températures tend exponentiellement vers 0 au fur et à mesure que le temps s'écoule.

¹² C'est ce dont M. Black avait eu l'intuition en écrivant que "l'usage de modèles théoriques consiste en l'introduction d'un nouveau langage ou dialecte". Cf. Black, 1962, p. 229.

La mécanique statistique est la théorie du mouvement des molécules composant les corps qui a pour but d'exhiber les causes microscopiques des phénomènes thermiques et hydrodynamiques. En ce qui concerne nos corps A et B décrits ci-dessus, elle conduit à des prédictions différentes de celles de la thermodynamique. Considérons par exemple que les corps A et B sont des gaz. La mécanique statistique les représente comme composés de molécules se mouvant en tous sens et rebondissant les unes sur les autres et sur les parois du récipient de façon apparemment aléatoire. Or une telle conception est incompatible avec la conséquence du Deuxième Principe de la thermodynamique selon laquelle la thermalisation s'effectue selon un flux *rigoureusement unidirectionnel*. En effet, dans le cadre de la mécanique statistique, il y a fort à parier que les molécules de plus grande vitesse, responsables de la température élevée de A , si A est le corps le plus chaud, iront de A vers B , *mais également, de temps à autre, de B vers A .*

Dans le cas très simple présenté, nous sommes donc face à un problème fâcheux, celui de l'incompatibilité entre les prédictions de deux théories qui ont pour but de rendre compte des mêmes phénomènes, la thermodynamique et la mécanique statistique. Cette incompatibilité est aggravée par une importante conséquence des principes sur lesquels repose la mécanique statistique classique, à savoir le *théorème de récurrence de Poincaré*, selon lequel les molécules soumises aux lois de la mécanique newtonienne reviennent nécessairement au bout d'un temps très grand, mais fini, dans un état arbitrairement proche de leur état initial. Cette prédiction est en contradiction avec le Deuxième Principe de la thermodynamique, car elle implique que le flux de chaleur se produira, si l'on attend suffisamment longtemps, de B vers A après que les températures des deux corps se sont égalisées.

Le modèle des Ehrenfest a pour but de clarifier ce problème issu de la confrontation entre deux représentations théoriques antagoniques d'un même phénomène. Considérons une boîte constituée de deux compartiments séparés par une cloison. Le compartiment A contient initialement N balles numérotées. Imaginons le processus suivant : nous tirons au hasard un nombre entre 1 et N , et nous déplaçons la balle qui porte le numéro correspondant du compartiment A vers le compartiment B . Au bout d'un certain temps, il arrivera que la boule portant le numéro correspondant au numéro tiré se trouve dans le compartiment B . Dans ce cas, nous la replacerons dans le compartiment A . Au fur et à mesure que le jeu se déroule, les balles vont de A vers B , mais aussi, de temps en temps, de B vers A . Chaque tirage est indépendant des autres, et tous les numéros entre 1 et N sont équiprobables. Tant qu'il y a plus de balles dans le compartiment A que dans le compartiment B , la probabilité que la balle tirée aille de A vers B est plus grande que la probabilité qu'elle aille de B vers A . Dans ce cas, le flux global de balles est unidirectionnel. Le nombre de balles dans A à chaque instant obéit à ce que l'on appelle une "chaîne de Markov". Le nombre moyen de balles dans A décroît exponentiellement jusqu'à $N/2$. En outre, on peut prédire avec une probabilité égale à 1 que le système reviendra dans son état initial, dans lequel toutes les balles sont dans le compartiment A . On peut même calculer au bout de combien de temps cela se produira : le système reviendra dans son état initial

au bout de 2^N pas, ce qui représente un temps extrêmement long même si le nombre des balles n'est que de quelques centaines.

Il est évident que ce modèle représente une situation qui est incompatible avec nos meilleurs théories physiques, et en particulier avec la mécanique quantique, ou même avec la mécanique newtonienne appliquée au domaine microscopique, puisqu'il n'est pas envisageable de faire passer des molécules réelles d'un compartiment à l'autre en les manipulant une à une. L'intérêt de ce modèle est qu'il nous permet de nous placer en imagination dans une situation fictive où les molécules des gaz ont les mêmes propriétés que des balles macroscopiques. Cela revient à mettre en parenthèses tout ce que nous savons par ailleurs des molécules réelles, et à observer certains aspects du comportement de ces molécules fictives. Dans ce contexte imaginé, le terme "molécule" acquiert une signification nouvelle — nous manipulons dans le modèle un concept différent de celui de "molécule réelle".

Le modèle des Ehrenfest tire son utilité épistémique de ce que le comportement d'un grand ensemble de ces molécules fictives présente de nombreuses ressemblances avec celui d'un corps réel. Ainsi, on atteint dans le modèle un état d'équilibre au bout d'un temps assez bref — de même, on peut prédire un retour à l'état initial. La conclusion que l'on peut tirer de ce modèle est que puisque les états initiaux et finaux des deux systèmes — le système réel et le système fictif — offrent certaines ressemblances, c'est que les processus qui ont lieu entre ces deux états sont eux aussi comparables l'un à l'autre. Cela signifie que les molécules réelles sont soumises à un processus qui ressemble étroitement au processus de tirage au sort décrit dans le modèle. L'information que nous pouvons tirer de ce modèle est même plus précise — nous pouvons saisir grâce à lui la façon dont le conflit entre mécanique statistique et thermodynamique peut être résolu. En effet, l'évaluation numérique du temps de récurrence qui est rendue possible par le modèle permet d'évacuer le conflit entre le théorème de récurrence de Poincaré et le Deuxième Principe de la thermodynamique en raison du caractère non physique — trop grand pour correspondre à quelque phénomène physique concevable que ce soit — du temps de récurrence.

Le modèle des Ehrenfest représente ainsi une situation fictive, qui contredit les théories disponibles pertinentes pour le phénomène concerné, puisque les molécules ne peuvent être représentées à la fois comme des balles que l'on peut manipuler individuellement et comme des corps soumis aux lois de la mécanique newtonienne ou à celles de la mécanique quantique. Cependant, le fait qu'il ne représente pas une situation réelle constitue précisément la raison de son intérêt scientifique. En effet, il nous permet d'*isoler* en imagination une hypothèse particulière concernant le comportement thermique des corps macroscopiques, à savoir celle selon laquelle le *nombre* des molécules qui les constituent joue un rôle déterminant, pour l'étudier indépendamment d'autres hypothèses, comme celles qui concernent la *dynamique* des molécules. C'est précisément la dynamique des "molécules" du modèle qui leur confère leur caractère fictif — nous interprétons cette caractéristique du modèle comme la mise entre

parenthèses intentionnelle du concept de dynamique moléculaire, mise entre parenthèses qui est destinée à concentrer l'étude sur une autre composante importante des systèmes réels, à savoir le nombre des molécules en jeu. Ainsi le modèle des Ehrenfest permet-il de tester l'hypothèse selon laquelle le nombre des molécules joue un rôle déterminant□ hypothèse dont le versant négatif est que la dynamique moléculaire joue un rôle moindre dans le phénomène de thermalisation□, et ce en l'isolant des autres hypothèses de la théorie qui sert de cadre ici, la mécanique statistique. Ce modèle a donc une fonction épistémique de clarification et de test□il permet en quelque sorte d'aller à l'encontre de la caractéristique des théories qui est décrite par la thèse Duhem-Quine, à savoir que les hypothèses des théories sont solidaires les unes des autres et ne peuvent être testées séparément. Le rôle du modèle des Ehrenfest est précisément de battre en brèche cette impossibilité en décrivant une situation fictive où il est possible de tester l'hypothèse selon laquelle le nombre des molécules joue un rôle plus important que leur dynamique indépendamment d'autres hypothèses, qui portent sur la dynamique moléculaire ou la nature des molécules.

3.2 Le modèle de l'atome de Bohr

Même si le modèle des Ehrenfest représente une situation manifestement fictive□ les molécules qui composent les gaz ne sont pas des boules numérotées que l'on tire au hasard□, on peut se demander si les autres représentations scientifiques couramment appelées "modèles" peuvent également faire l'objet d'une analyse fictionnaliste. Cette question semble se poser de façon particulièrement aiguë pour l'un des modèles les plus célèbres de l'histoire de la physique, le "modèle de l'atome de Bohr", développé pour l'hydrogène en 1913□ en effet, cette représentation de l'atome d'hydrogène, la première à rendre compte de façon remarquablement précise de toutes les données disponibles à ce moment, est-elle vraiment une représentation d'un atome qui n'existe pas, d'un atome fictif, ou bien la première pierre d'un édifice théorique, ayant pour but avéré de représenter les atomes réels□ Le "modèle de l'atome de Bohr" ne constitue-t-il pas une étape cruciale vers la constitution de la mécanique quantique, vers une *théorie* de la structure atomique□ Bohr n'a-t-il pas l'ambition dès 1913 de construire une telle théorie□ Si c'est le cas, c'est bien qu'il a en vue de *représenter la réalité*, et non une situation fictive. Dans ce qui suit, nous nous proposons de discuter cet exemple afin de voir si l'analyse fictionnaliste des modèles peut vraiment lui être appliquée. Il s'agit pour nous de montrer que notre analyse est pertinente pour un cas qui semble à première vue très défavorable.

3.2.1 La construction du modèle de l'atome de Bohr

Le modèle de l'atome de Bohr est la première représentation scientifique de la constitution des atomes qui ait été capable de rendre compte d'un très vaste ensemble de données nouvelles à l'époque. On peut donc le considérer comme une remarquable avancée scientifique. Afin de l'analyser correctement, il est nécessaire de rappeler brièvement ce qui constituait l'arrière-plan commun à tous les physiciens de l'époque¹³.

Spectroscopie

En 1913, rares étaient les physiciens qui doutaient encore de l'existence des atomes□ les expériences de Jean Perrin, en 1906, les avaient presque tous convaincus de la justesse de l'idée selon laquelle la matière possédait une structure granulaire. Cette idée faisait l'objet depuis les années 1860 d'un développement théorique cohérent dans la mécanique statistique, dont les succès ont également contribué à l'acceptation de l'atomisme moderne. Cette théorie, la mécanique statistique, avait permis d'élaborer certaines hypothèses sur la structure interne des molécules□ cependant, les données expérimentales recueillies depuis la naissance de la *spectroscopie* constituaient un ensemble de connaissances plus riches et plus sûres que les hypothèses dérivées des calculs de la mécanique statistique.

Le modèle de l'atome de Bohr ne peut être analysé correctement que si l'on juge exactement de ce qui était à l'époque de sa création de l'ordre de l'hypothèse, plus ou moins bien acceptée, et de l'ordre des données expérimentales, certes toujours à interpréter, mais présentant une stabilité certaine relativement à la variété des hypothèses possibles. La spectroscopie est l'étude expérimentale du rayonnement électromagnétique des atomes. Les atomes, convenablement excités, émettent une énergie dont on recueille la trace sur des plaques photographiques. L'image produite est celle de raies discrètes et bien différenciées, d'épaisseur variable. La configuration de telles raies, appelée "spectre" constitue une sorte de "signature" des divers types d'atomes□le spectre de l'hydrogène, par exemple, est toujours le même quel que soit l'échantillon étudié.

Quantification

Outre les hypothèses issues de la mécanique statistique et les données de la spectroscopie, on disposait dans les années 1900-1910 d'un important ensemble théorico-expérimental élaboré par Planck autour du phénomène appelé "rayonnement du corps noir"¹⁴. Des études de Planck il ressortait que l'on ne pouvait rendre compte des propriétés de ce rayonnement (qui, contrairement à celui des atomes, est continu) qu'au prix d'une hypothèse profondément étrange dans ces années d'apogée de la physique classique, celle selon laquelle l'énergie totale émise par un "corps noir" à une fréquence donnée est *quantifiée*, c'est-à-dire

¹³ Les rappels historiques de cette section et de la suivante sont inspirés de Lurçat, 1990, cf. en particulier le ch.□VI.

¹⁴ Pour des études historiques de ces travaux, voir en particulier Darrigol, 1993 et Kuhn, 1978.

qu'elle ne peut prendre que des valeurs multiples d'un certain quantum¹⁵. Il est très important de noter que cette hypothèse allait à l'encontre de toutes les conséquences des théories les mieux admises à l'époque, à savoir la mécanique classique et l'électromagnétisme, dans lesquelles l'idée même de quantum d'énergie ne pouvait recevoir aucune justification. L'idée de quantum irréductible d'énergie entrainait en contradiction avec tous les principes et toutes les conséquences de ces théories, qui, soulignons-le, constituaient un fondement théorique jugé particulièrement sûr à cette époque. L'hypothèse de Planck a été à bon droit qualifiée de révolutionnaire. En raison de la grande précision avec laquelle elle permettait cependant de rendre compte des données expérimentales relatives au rayonnement du corps noir, elle a été rapidement acceptée par les physiciens. L'autre raison de cette acceptation est l'utilisation qu'en a faite Einstein en 1905 en montrant qu'un certain nombre de phénomènes électromagnétiques pouvaient être expliqués de façon simple si l'on supposait que le rayonnement électromagnétique lui-même, et non seulement son interaction avec la matière, comme dans le cas du corps noir, était quantifié. Les "paquets" de lumière ainsi identifiés furent appelés "photons". L'hypothèse selon laquelle le rayonnement électromagnétique est quantifié, même si elle permettait de rendre compte de façon particulièrement simple de nombreux phénomènes, entrainait tout autant en contradiction avec les théories admises que celle de la quantification de l'énergie d'interaction entre matière et rayonnement. Elle choquait de nombreux physiciens, et d'importants efforts ont été menés à cette époque pour rendre compatibles ces hypothèses de quantification avec l'électromagnétisme classique et la mécanique classique.

Électron, radioactivité, modèles de Thomson et de Rutherford

A la toute fin du XIXe siècle, deux découvertes expérimentales avaient permis de rendre plus précises les hypothèses qu'il était possible de former alors sur la structure de l'atome au vu des théories disponibles. Il s'agit de celle de l'*électron*, et de celle de la *radioactivité*. La première de ces découvertes est à l'origine de la série des modèles de l'atome dont celui de Bohr constitue l'aboutissement. Elle a également fait l'objet de théories diverses, dont la plus satisfaisante était au tournant du siècle celle de Lorenz, selon laquelle les électrons pouvaient être considérés comme des oscillateurs émettant par rayonnement de l'énergie électromagnétique¹⁶.

La découverte de l'électron est en fait la découverte expérimentale de charges électriques discrètes, dont les porteurs ont été nommés "électrons". On a immédiatement supposé que ces charges, dites négatives par convention, devaient compter parmi les composants des atomes, dont on savait par ailleurs qu'ils étaient électriquement neutres. Cela imposait qu'ils contiennent des charges positives. L'une des premières représentations de la structure des atomes, l'un des premiers modèles de l'atome, est due à J.J. Thomson, le découvreur de l'électron. En 1903, celui-

¹⁵ Ce quantum est le produit de la fréquence en question par une certaine constante, h , appelée depuis "constante de Planck".

¹⁶ La théorie de Lorenz était conçue comme une théorie générale et unifiée de tous les phénomènes physiques, électromagnétiques et mécaniques.

ci fit l'hypothèse selon laquelle un atome est une boule portant une charge positive uniforme, et truffée d'électrons — on parle à son sujet de "*plum pudding model*". Le développement de cette hypothèse en un modèle complet de l'atome passe par l'application à cette boule des lois de l'électromagnétisme. Le résultat de cette application est que les électrons peuvent osciller à l'intérieur de la boule — un comportement qui est en accord avec la théorie de Lorenz, ce qui constitue un argument important en faveur des hypothèses incluses dans le modèle de Thomson.

La seconde découverte évoquée ci-dessus, celle de la radioactivité, ruine cependant le modèle de Thomson. A partir de 1896, on étudie de façon systématique le rayonnement émis naturellement — c'est-à-dire sans excitation — par certains atomes comme celui d'uranium — et on utilise ce phénomène pour tenter d'étudier directement la structure des atomes. Les expériences qui aboutissent au rejet de du modèle de Thomson sont celles de Geiger et Marsden, qui bombardent d'un rayonnement α (c'est-à-dire, comme on l'a su plus tard, d'un jet d'ions d' He^{++}) une mince feuille d'or. La plupart des particules α traversent la feuille d'or en ligne droite, sans être déviées — mais dans un cas sur huit mille environ, les particules α sont déviées d'un angle supérieur à 90° . Le modèle de Thomson ne permet pas d'expliquer ces déviations, qui sont dues à ce que la quantité de mouvement des particules α est changée par l'action d'une force importante. Or la boule chargée positivement et truffée d'électrons par quoi Thomson représente un atome — un atome d'or en l'occurrence — est incapable d'exercer une telle force. Le champ électrique créé par cette boule chargée décroît en effet à l'extérieur de la boule en raison inverse du carré de la distance, et à l'intérieur, il décroît de la surface vers le centre, où il s'annule — cela signifie qu'une boule chargée crée un champ électrique d'autant plus fort qu'elle est plus petite. Or la boule chargée de Thomson est de "grande" taille, puisqu'elle a la taille de l'atome lui-même. Elle ne peut être à l'origine du champ responsable de la déviation des particules α . Le champ électrique qu'il est nécessaire de supposer pour expliquer cette déviation est beaucoup plus fort que celui créé par l'atome de Thomson, ce qui signifie que l'élément de l'atome qui porte la charge positive est beaucoup plus petit que l'atome lui-même.

Les résultats des expériences de Geiger et Marsden sont analysées par Rutherford en 1911. Puisque les électrons contenus dans les atomes d'or sont beaucoup trop légers pour agir sur la trajectoire des particules α — qui sont en effet plusieurs milliers de fois plus lourdes que les électrons, et arrivent sur la feuille d'or à une vitesse de l'ordre de 10^7 km par seconde —, on doit supposer selon les considérations qui précèdent que les particules α sont déviées par une charge positive de taille beaucoup plus petite que celle de l'atome. Cela conduit Rutherford à construire un modèle dit "saturnien" ou "planétaire" de l'atome d'or, selon lequel cet atome est constitué d'un *noyau* central très petit¹⁷ et de charge positive, entouré d'électrons disposés sur des anneaux.

Rutherford n'avait nullement l'ambition, en élaborant ce modèle saturnien de l'atome, de construire une *théorie* atomique en bonne et due forme, mais uniquement de proposer une

¹⁷ Le diamètre du noyau est , comme nous le savons aujourd'hui, environ 10^{-14} fois plus petit que celui de l'atome.

représentation de l'atome d'or qui tienne compte de façon *ad hoc* des résultats des expériences de Geiger et Marsden. Il est en effet impossible que la représentation de l'atome d'or inventée par Rutherford soit une représentation exacte, pour les raisons suivantes. Les électrons sont nécessairement en mouvement dans l'atome, sinon ils seraient attirés par le noyau chargé positivement, et tomberaient sur lui. L'atome serait dans ce cas instable. Or les données de la spectroscopie indiquent sans aucune doute possible que les propriétés électromagnétiques des atomes sont remarquablement stables. Les électrons ne peuvent donc être immobiles dans le modèle de Rutherford. Ils doivent être constamment en mouvement. Ce mouvement est nécessairement un mouvement circulaire, sinon les électrons s'éloigneraient indéfiniment de l'atome, qui perdrait sa stabilité. Or les corps chargés en mouvement accéléré¹⁸ rayonnent. C'est là une des premières conséquences de l'électromagnétisme. En rayonnant, ils perdent de l'énergie, et donc de la vitesse, jusqu'à tomber sur le noyau. Le modèle de Rutherford est donc définitivement instable. Il est plus satisfaisant que le modèle de Thomson en ce qu'il tient compte des résultats des expériences de Geiger et Marsden, mais est en contradiction vis-à-vis de l'électromagnétisme. L'atome qu'il décrit en viole les lois de la théorie électromagnétique. C'est donc un atome fictif au sens où il est incompatible avec l'une des théories les mieux acceptées du moment, qui interdit irrémédiablement à un tel atome d'être stable.

3.2.2 Analyse du modèle de l'atome de Bohr

Nous allons voir maintenant que le modèle de l'atome de Bohr, comme ceux de Thomson et de Rutherford, peut être considéré comme représentant un atome fictif. Ce modèle est en effet incompatible, au sens strict, avec la mécanique classique et l'électromagnétisme. Cependant, il manifeste une avancée considérable vis-à-vis du modèle de Rutherford. L'atome représenté est *stable*. En 1913, Bohr construit un modèle détaillé de l'atome d'hydrogène, qui est l'atome le plus simple qui soit puisqu'il ne comporte qu'un seul électron. Les données relatives à cet atome sont constituées par une série de raies spectrales¹⁹, dont les longueurs d'onde (ou les fréquences) obéissent à une formule empirique découverte en 1885 par Balmer.

$$f = \nu / 4\pi^2 / n^2,$$

où $n \in \mathbb{N}$, et ν est une "fréquence fondamentale"

En 1913, les raisons théoriques pour lesquelles la formule de Balmer représente les données de façon aussi précise reste un profond mystère.

Du point de vue des théories classiques, l'atome d'hydrogène est décrit de la façon suivante. Le mouvement circulaire de l'électron est périodique, et la fréquence du rayonnement qu'il émet est égale à celle de ce mouvement. Dans ce rayonnement peuvent aussi intervenir des

¹⁸ Un mouvement circulaire est un mouvement accéléré.

¹⁹ D'autres séries ont été découvertes par la suite.

harmoniques, c'est-à-dire des multiples entiers de cette fréquence — on notera qu'aucune harmonique n'apparaît dans la formule de Balmer, qui est cependant d'une grande précision empirique. Outre les théories classiques, Bohr a à sa disposition le travail de Planck sur le rayonnement du corps noir. Ce dernier a établi, comme nous l'avons vu, que lorsque l'on multiplie une fréquence par la constante h on obtient une énergie. Si l'on tient compte de cela, on peut envisager l'équation

$$h\nu = E_2 - E_1 = h\nu / n^2 \quad (1)$$

comme un bilan énergétique, c'est-à-dire comme la représentant la différence entre les énergies de deux états qui sont respectivement le début et la fin d'un processus.

Fort de tout cela, Bohr introduit trois hypothèses tout aussi révolutionnaires que celle de Planck, et qui lui permettent d'élaborer le modèle de l'atome d'hydrogène qui porte son nom.

1) *Dans l'atome d'hydrogène, l'énergie de l'électron est quantifiée.*

Elle ne peut prendre que certaines valeurs (négatives), qui sont les suivantes

$$E_0 = 0, E_1 = -h\nu / 4, E_2 = -h\nu / 9, \dots, E_n = -h\nu / n^2, \dots \quad (2)$$

Chacune de ces énergies est dite être celle d'un "état stationnaire" de l'atome. Ces états sont numérotés par le nombre entier n qui figure au carré au dénominateur. n est appelé un "nombre quantique", et peut prendre n'importe quelle valeur entière.

Le concept d'*état stationnaire* est un concept radicalement nouveau, qui n'appartient à aucune théorie classique, et qui n'a en outre aucun sens en leur sein. Comme le montre ce qui suit, il est mobilisé par des hypothèses qui sont en contradiction avec les théories classiques. Notre analyse philosophique de ce fait est que le concept d'état stationnaire, dans sa radicale nouveauté, est un concept "en attente de théorie" — Bohr décrit un atome fictif d'hydrogène à l'aide de concepts qui appartiennent à une théorie à venir. La différence entre ce modèle et celui de Rutherford tient précisément à son caractère *prospectif*. L'atome de Bohr est décrit par des hypothèses susceptibles d'être justifiées par une théorie future, non encore constituée, mais dont certains concepts sont utilisés à l'aveuglette, en quelque sorte, dans le modèle. L'atome de Rutherford en revanche n'est décrit par aucun concept nouveau de ce type — les éléments qui composent cette description sont introduits de façon *ad hoc* afin de prendre en compte les résultats des expériences de Geiger et Marsden.

2) *L'électron ne rayonne pas lorsqu'il est dans un état stationnaire.*

L'atome décrit par le modèle rayonne uniquement lorsqu'il "saute" d'un état stationnaire à un autre, dont l'énergie est plus faible. Dans ce cas, la fréquence du rayonnement émis est déterminée par une équation du même type que (1). Le concept de "saut" d'un état stationnaire à un autre est lui aussi un concept radicalement nouveau, qui ne peut prendre sens dans aucune théorie classique.

3) Le rapport entre l'énergie cinétique de l'électron et sa fréquence de rotation est un multiple entier $h/2$.

Cette hypothèse permet de calculer la liste (2) des énergies permises, et autorise un accord remarquable avec les données expérimentales, en particulier avec la valeur numérique mesurée de la constante $h\nu$.

Dans une lettre à Mc Laren du 1er septembre 1913²⁰, Bohr reconnaît la nécessité de l'introduction d'hypothèses nouvelles pour explorer la constitution des atomes, mais qualifie les siennes d'"horribles", montrant par là sa répugnance à aller ainsi à l'encontre des lois classiques. Comme nous l'avons vu, ni la première hypothèse, ni la deuxième ne sont susceptibles de trouver aucune justification dans l'électromagnétisme classique, théorie qui, soulignons-le, ne contient aucun élément qui pourrait suggérer qu'elle cessât d'être valide à l'intérieur des atomes. La troisième est tout aussi "horrible", car elle a pour conséquence que l'électron émet un rayonnement dont la fréquence est différente de sa fréquence de rotation. En outre, autre mystère complètement opaque en 1913, la fréquence du rayonnement émis est déterminée d'une part par l'énergie de l'électron avant l'émission, et d'autre part par son énergie *après* l'émission□ comment l'électron pourrait-il "savoir" à quelle fréquence il doit rayonner pour arriver dans le bon état stationnaire□ Comment rendre compte de ce qui ressemble à la manifestation d'une téléologie aux antipodes des principes les mieux ancrés de la physique□ Autant de questions auxquelles le modèle de Bohr est incapable de répondre en 1913. Son mérite est de permettre de les *poser*, et d'en faire autant de guides pour la constitution de la future théorie dans laquelle toutes les étrangetés du modèle seront justifiées.

Au vu de ce qui précède, nous pouvons analyser le modèle de l'atome de Bohr comme une fiction au sens où nombre de concepts et d'hypothèses utilisés dans cette description sont inouïs dans les théories classiques, et sont, comme nous pouvons l'affirmer rétrospectivement, les indices d'une théorie à venir dans laquelle ils trouveront leur justification. Ainsi peut-on dire que le modèle de l'atome de Bohr entretient deux types de rapports avec des *théories*. Avec la mécanique quantique à venir, il entretient des rapports de prospection, comme nous pouvons en juger rétrospectivement□ il contient des concepts et des hypothèses qui ne prennent un sens précis que dans le cadre de la mécanique quantique. Vis-à-vis des théories classiques, il entretient des rapports de rupture, mais aussi, et il est important de le noter, de fondement□ le modèle est en effet fondé sur ces théories au sens où elles y sont utilisées partout où elles peuvent l'être, par exemple pour évaluer l'énergie des états stationnaires.

En guise de conclusion sur cet exemple, on peut rappeler que ce modèle de l'atome, s'il permet d'obtenir des résultats remarquables pour l'hydrogène, n'est pas généralisable tel quel aux autres atomes, plus complexes. Comme Bohr lui-même y insistait, il s'agit d'un modèle

²⁰ Citée par Lurçat, 1990, p.192.

provisoire □ en attente de théorie, mais sur la voie d'une théorie apte à généraliser aux autres atomes ses éléments les plus robustes.

4 Vers une typologie des modèles scientifiques

Notre hypothèse fondamentale, d'après laquelle le contenu représentationnel d'un modèle scientifique doit être interprété exactement comme le contenu représentationnel d'une fiction, possède l'avantage d'expliquer comment les hypothèses d'un modèle peuvent à la fois contredire des hypothèses théoriques bien admises, et néanmoins nous dire quelque chose sur le monde. Reste à savoir comment le contenu en question contribue à l'entreprise de production du savoir scientifique. Certes, notre théorie rend compte de la nature du contenu qu'on peut attribuer à un modèle. Mais puisque ce contenu y est envisagé comme un contenu *d'imagination*, elle ne semble rien dire sur la fonction épistémique des modèles.

Or doit-on vraiment supposer que les modèles n'ont qu'*une seule* fonction épistémique □ L'erreur des tentatives antérieures pour caractériser la distinction entre théories et modèles nous paraît être, précisément, d'avoir supposé que c'était le cas, et que c'était *la* fonction des modèles qui les distinguait clairement des théories. Pour notre part, nous possédons déjà un critère de distinction entre théories et modèles □ les théories sont élaborées dans le but d'être crues, soutenues, assertées □ les modèles présentent en revanche un contenu qui doit être imaginé. Nous n'avons donc pas besoin de supposer l'existence d'une fonction épistémique unique qui devrait jouer ce rôle de démarcation. De fait, nous pensons que les modèles jouent des rôles divers dans l'entreprise scientifique. Pour comprendre ce qui unifie tous ces rôles, il faut partir de la nature des pensées qu'ils visent à susciter.

A la suite de Ramsey²¹, nous pouvons nous représenter un système de croyances comme une cartographie du monde, nous permettant de nous y orienter. Les théories scientifiques contribuent, de façon complexe, à une telle cartographie, pour autant qu'elles peuvent faire l'objet de croyance. La cartographie qui repose sur les théories scientifiques ressemble cependant plus à celle des géographes du XVI^e siècle qu'à celle de nos contemporains. Les territoires inexplorés y sont légion. Pire, on ignore comment mettre exactement en relations mutuelles les territoires qu'on estime bien connus. Pour le dire d'un mot, la représentation scientifique du monde apparaît comme fondamentalement incomplète²². Ainsi peut-on, dans certains domaines du savoir, disposer d'un grand nombre de données empiriques d'une part, et de principes théoriques bien confirmés, d'autre part, sans être pour autant capable de mettre d'aucune manière en relation les données et les principes théoriques. Puisque les scientifiques sont alors dans l'incapacité de

²¹ Voir E. P. Ramsey, 1929.

²² Nous ne discuterons pas la question de savoir si une telle incomplétude est accidentelle ou essentielle à l'entreprise scientifique.

savoir effectuer de telles mises en relation, c'est-à-dire de pouvoir formuler des hypothèses justifiables qui le permettrait, nous soutenons qu'ils les *imaginent*. De cette activité imaginative naissent deux sortes de modèles, selon que l'imagination porte sur les *principes théoriques* ou sur les *données expérimentales*. Un troisième type de modèles peut en outre être associé à la fonction de *test* d'une hypothèse (ou plusieurs) indépendamment de l'ensemble théorique à laquelle elle appartient.

4.1 Les modèles prospectifs

Une première stratégie possible, dans une situation d'incomplétude du type de celle que nous venons de décrire, revient à imaginer un contexte dans lequel d'une part un certain nombre d'hypothèses permettent de subsumer les données expérimentales sous des lois, *et* où d'autre part ces hypothèses peuvent être déduites de principes théoriques fondamentaux. Nous appelons "prospectif" un modèle constitué par un tel ensemble d'hypothèses, car il suppose de se projeter en imagination, pour ainsi dire, dans un avenir dans lequel les hypothèses pourront être justifiées par des raisons théoriques.

En quoi un tel ensemble d'hypothèses diffère-t-il d'un ensemble *théorique* d'hypothèses? Qu'est-ce qui fait de cet ensemble d'hypothèses un modèle plutôt qu'un morceau de théorie en devenir? Il nous semble que la formation d'hypothèses théoriques est contrainte par un idéal d'intégration aux théories existantes. Une hypothèse théorique ne peut être formée que si elle est ou bien compatible avec les principes théoriques déjà acceptés par une communauté, et intégrable dans le réseau formé par ces principes, ou bien si elle peut être mise en concurrence avec l'un de ces principes²³. Il n'en va pas de même pour ce que nous appelons les modèles prospectifs. Les hypothèses qui constituent ces modèles, bien que difficiles à intégrer à l'ensemble des principes théoriques acceptés au moment où ils sont construits, et impossibles en général à déduire de ces principes, ne peuvent pas non plus être vues comme étant directement en concurrence avec eux. Une telle concurrence n'est que prospective: si les hypothèses du modèle permettent de subsumer les données, on peut imaginer qu'un jour, on disposera des moyens de les intégrer à la théorie.

Le modèle de l'atome de Bohr est un exemple particulièrement frappant de modèle prospectif. En effet, comme cela a été indiqué, Bohr introduit dans sa description de l'atome d'hydrogène trois hypothèses qui ne peuvent être conçues comme étant directement en concurrence avec les théories acceptées, la mécanique classique et l'électromagnétisme, puisqu'elles n'ont à strictement parler pas de sens dans leur cadre. Ces hypothèses révolutionnaires ne peuvent être qualifiées de théoriques: elles ne s'intègrent pas en 1913 dans

²³ Il nous semble qu'on peut comparer, de ce point de vue, la formation de nouvelles hypothèses théoriques et la formation de croyances. On peut former une nouvelle croyance si elle est compatible avec l'ensemble des autres croyances, et intégrable à cet ensemble, ou bien, dans le cas d'une incompatibilité, si les principes de la dynamique des croyances arbitrent en sa faveur.

une structure unifiée et hiérarchisée, mais apparaissent comme des éléments isolés, capables de représenter les données expérimentales sans qu'il soit possible de justifier leur introduction autrement que par ce remarquable accord avec les données. Elles ne seront intégrées dans un tout théorique qu'en 1925-26, lorsqu'une véritable théorie du domaine microscopique sera élaborée.

L'intérêt épistémique de telles constructions est facile à concevoir. Elles permettent d'entamer l'exploration de nouveaux territoires, d'inventer de nouvelles connexions entre des parties déjà cartographiées du monde. La conclusion qu'on peut tirer de la construction d'un modèle prospectif, c'est que dans une situation — pour le moment purement imaginaire — dans laquelle les hypothèses du modèle seraient déductivement justifiées à partir de principes théoriques, les données problématiques seraient expliquées. Un tel modèle, s'il parvient à subsumer l'ensemble des données de façon élégante, incite donc à modifier la théorie dont on part — par exemple, la mécanique classique pour ce qui est du modèle de Bohr — dans une direction qui permettra de l'y intégrer.

4.2 Les modèles-ponts

On peut, pour établir une connexion manquante entre théorie et données, imaginer une situation dans laquelle la *théorie* se trouve être différente de celle qui est réellement acceptée relativement à un domaine d'entités. On peut aussi imaginer une situation dans laquelle les *données* diffèrent, selon certains aspects, des données réelles. La construction de modèles repose très souvent sur un tel point de départ. Cartwright a insisté, d'une façon très claire, sur l'importance de telles situations en physique mathématique

La théorie n'a qu'un stock de principes très limité permettant de passer des descriptions [d'un domaine d'entités] à des équations, et ces principes nécessitent une information d'un type très particulier, structurée d'une façon elle aussi très particulière. Les descriptions (...) portant sur ce qu'il y a [dans le domaine d'entités] sont choisies en raison de leur adéquation descriptive. Mais les descriptions (...) donnant lieu à des équations doivent être choisies, dans une large mesure, pour leurs propriétés mathématiques. Il s'agit là d'une caractéristique de la physique mathématique. Les descriptions qui décrivent de la façon la plus satisfaisante ne sont généralement pas celles auxquelles on peut faire correspondre des équations²⁴.

Disposant d'un stock de descriptions tirées d'observations, le physicien qui veut subsumer ces descriptions sous une théorie se doit de leur donner une forme mathématique. Mais comme Cartwright y insiste à juste titre, cela présuppose une certaine dose de fiction. Nous devons imaginer que nos données, par exemple, portent sur des points matériels, au lieu de porter sur des

²⁴ Cf. Cartwright, 1983, p. 131.

solides en trois dimensions. Dans ce cas, l'idéalisation qui conduit à la fiction est "contrôlée"²⁵ □ lorsque nous représentons un corps par un point matériel, nous savons malgré tout comment représenter l'action d'une force, par exemple, sur d'autres points que le centre de gravité de ce corps, auquel il est rapporté dans l'idéalisation du point matériel. En outre, nous savons qu'une telle représentation, qui peut être supposée plus fidèle que l'idéalisation du point matériel, constitue en fait une sophistication inutile, puisque dans la plupart des cas, elle n'améliore la qualité des prédictions que de façon négligeable, et à un coût important en ce qui concerne les calculs effectués. Ce sont de telles considérations qui permettent de *justifier* l'usage de l'idéalisation du point matériel en mécanique newtonienne □ remplacer une représentation plus fidèle du corps étudié par un point matériel ne coûte presque rien du point de vue de la fidélité de la représentation tout en facilitant grandement les calculs.

Les modèles globaux du changement climatique permettent également d'illustrer la notion de "description préparée" chère à Cartwright, mais en reposant sur des idéalizations non contrôlées. Ils ont pour but de représenter la façon dont le climat global de la planète évolue sur de longues périodes de temps. La théorie physique sur laquelle ils reposent contient un petit nombre d'équations simples, mais impossibles à résoudre même par les ordinateurs les plus puissants dont nous disposons aujourd'hui en raison du nombre gigantesque des variables nécessaires. Les observations qui fournissent les valeurs de ces variables sur les périodes de référence sont faites par des sondes très nombreuses et de natures variées. La simple compilation de cet immense corps de données disparates requiert des moyens informatiques exceptionnels. La caractéristique majeure des représentations globales du climat est qu'elles requièrent une modélisation dès l'étape de la collecte des données □ en effet, avant même de pouvoir effectuer une première mise en équations de ces données, il est nécessaire de construire des modèles permettant de les homogénéiser et de remplir par extrapolation les lacunes qu'elles présentent. Dans ce cas, les activités d'"observation" et de modélisation, qui sont séparées dans les analyses classiques de la démarche scientifique, ne font qu'une. C'est le modèle lui-même qui fournit les données, tout en assumant son rôle normal d'aide à la prédiction. Cela ne signifie évidemment pas que les données soient produites de part en part par l'imagination des constructeurs de ces modèles, mais que leur collecte nécessite ce qu'on a pu appeler un "art scientifique"²⁶ requérant de se placer en imagination dans une situation où les données ne seraient ni ambiguës, ni lacunaires, ni hétérogènes.

Nous appellerons "modèles-ponts" ces mises en équations des données, permettant de les subsumer au moins partiellement sous une théorie²⁷. Il nous semble clair que les scientifiques ne considèrent en général pas ces modèles-ponts comme des hypothèses justifiées — ni même à vrai dire comme des hypothèses justifiables. Il s'agit bel et bien de fictions, qui

²⁵ Cf. Sklar, 1993, p. □58.

²⁶ Nous empruntons cette expression à un exposé de Paul Edwards, prononcé en mai 1999 dans le cadre du séminaire "Modèles, modélisations en sciences physiques et en sciences sociales".

²⁷ Ils correspondent dans une large mesure à ce que Suppes et van Fraassen appellent des "modèles de données". Cf. par exemple Suppes, 1974, et van Fraassen, 1994, pp. □49-351.

imposent la considération de situations imaginaires dans lesquelles les données auraient un contenu immédiatement traduisible en mathématiques. Quelle est l'utilité épistémique de telles fictions ? Pour répondre à cette question, il faut comprendre que les situations imaginaires dans lesquelles il faut se placer pour mathématiser les données peuvent être plus ou moins éloignées du monde réel dans l'espace logique²⁸. Pour certains buts, la distance qui sépare ces situations de la situation réelle n'importe pas. Redhead a tenté de produire une théorie de cette notion de "distance" entre situations fictives et situations réelles, à partir d'une analyse de la notion d'*approximation*²⁹. Cependant, la "distance" qu'il introduit reste extrêmement abstraite et n'autorise qu'un jugement purement formel sur la ressemblance entre la situation fictive représentée par le modèle et la situation réelle. En effet, lorsque des ensembles de solutions d'équations sont "proches" au sens de Redhead, cela ne garantit pas que les situations qu'ils représentent sont semblables relativement aux aspects qui nous intéressent : la notion de distance introduite par Redhead n'est pas nécessairement la mesure d'une ressemblance pertinente entre la situation fictive représentée par le modèle et la situation réelle.

4.3 Les modèles-tests et les modèles opportunistes

Partir de la spécificité de l'acte mental d'imagination nécessaire à la construction et à l'interprétation des modèles scientifiques nous a permis de distinguer les deux grandes catégories de modèles qui précèdent. Nous ne voulons cependant pas soutenir que tous les modèles appartiennent à l'une de ces catégories. Ainsi, la variation imaginative à l'origine d'un modèle peut porter non pas sur une théorie dans son ensemble, mais uniquement sur un petit nombre des principes, et du même coup des concepts, qui constituent une théorie. Dans ces modèles, les principes en question sont mis entre parenthèses, parce qu'on pense qu'ils n'ont aucune pertinence dans l'explication d'un ensemble de données. La construction du modèle fonctionne alors à la façon d'un test : si les hypothèses qui le constituent, et dont la cohérence suppose la mise entre parenthèses d'un ensemble de principes ou d'hypothèses *P*, permet de subsumer correctement des données, on peut en conclure de façon fondée que les principes *P* n'ont effectivement pas de pertinence dans l'explication du phénomène considéré.

Le modèle des Ehrenfest présenté ci-dessus est un exemple important de modèle-test. Les principes qui sont mis entre parenthèses dans ce modèle sont comme nous l'avons vu ceux qui décrivent la *dynamique* des molécules. Lorsque l'on fait abstraction des caractéristiques du mouvement des molécules et qu'on les remplace en imagination par un processus de tirage au sort, on parvient, si l'on envisage un nombre suffisamment grand de "molécules", à un accord remarquable avec certaines propriétés des systèmes réels. La conclusion de ce modèle-test est

²⁸ Cf. Lewis, 1973, pour des éclaircissements concernant cette notion.

²⁹ Cf. Redhead, 1980, p. 50.

que les hypothèses théoriques portant sur la dynamique moléculaire n'ont pas de pertinence pour l'explication de *ces* propriétés³⁰. Il peut donc être dit avoir comme fonction de tester les diverses hypothèses que l'on peut faire sur les effets de la dynamique moléculaire dans les phénomènes de thermalisation.

Parfois, la nature ne se prête tout simplement pas, ou se prête difficilement, au test empirique d'hypothèses théoriques. Une possibilité ouverte est alors de *simuler* un tel test théorique. Considérons ainsi les hypothèses complexes que l'on peut formuler à propos du système constitué par un avion d'un type nouveau et son pilote. Il va sans dire que la question de savoir si de telles hypothèses sont ou non justifiées possède un caractère crucial. Pour d'évidentes raisons cependant, elles ne peuvent être soumises à un test expérimental. Elles sont soumises en revanche à des *simulations* de tests expérimentaux, à l'aide de simulateurs de vol. Simuler le comportement d'un nouveau type d'avion et d'un pilote, dans certaines circonstances, revient à imaginer que les données fournies par le simulateur, qui servent d'entrées perceptives au pilote, sont des données factuelles. Un vol simulé, pour nous, est un vol modélisé dans un type particulier de modèle — un modèle qui tire partie du comportement réel du pilote dans un jeu de faire-semblant pour tester certaines hypothèses sur ce que serait son comportement réel dans un vol réel. En général, notre conception de la modélisation permet de faire place à des modèles qui, comme le "modèle du bateau, le bulldozer miniature, le modèle architectural d'un campus, le modèle de bois ou d'argile d'une voiture³¹", ou comme le simulateur de vol, ne formulent pas explicitement de généralisations sur un domaine d'entités. En construisant de tels modèles miniatures, ou simplifiés, d'entités ou de phénomènes, on imagine que le changement d'échelle ou la simplification n'ont pas de pertinence, et que les données dont nous pouvons disposer sur les modèles réduits sont des données réelles, en provenance des entités réelles. De la sorte, on parvient tester, au moins en imagination, des hypothèses portant sur le comportement des entités en question. Nous nommons modèles opportunistes de telles simulations, car elles tirent partie de lois réelles dans des jeux de faire-semblant. Ces modèles opportunistes ressemblent à la fois aux modèles-tests, puisqu'ils permettent d'évaluer des hypothèses, mais partagent également certaines des caractéristiques des modèles-ponts, puisque l'activité imaginative porte sur des données expérimentales.

BIBLIOGRAPHIE

BLACK, M, 1962, *Models and Metaphors*, Cornell University Press.

³⁰ Les caractéristiques de la dynamique moléculaire peuvent cependant avoir un effet important sur d'autres propriétés des phénomènes de thermalisation.

³¹ Cf. Goodman, 1990, p. 207.

BOHR, N., 1913, "On the Constitution of Atoms and Molecules", *Phil. Mag.*, 26, 1-25, 476-502, 857-875.

CARTWRIGHT, N., 1983, *How the Laws of Physics Lie*, Oxford University Press.

CURRIE, G., 1990, *The Nature of Fiction*, Cambridge University Press.

DARRIGOL, O., 1993, *From c-numbers to q-numbers: The Classical Analogy in the History of Quantum Mechanics*, The University of California Press.

EHRENFEST, P. & T., 1907, "Über zwei bekannte Einwände gegen Boltzmann's H-Theorem", *Phys. Zeitschrift* 8, p. 311.

GOODMAN, N., 1968, *Langages de l'art*, Jacqueline Chambon, 1990.

KUHN, T., 1978, *Black Body Theory and the Quantum Discontinuity*, Oxford University Press.

LEWIS, D., 1973, *Counterfactuals*, Harvard University Press.

LEWIS, D., 1979, "Attitudes *De Dicto* and *De Se*", *The Philosophical Review*, 88□513-543.

LEWIS, D., 1986, *On the Plurality of Worlds*, Basil Blackwell.

LURÇAT, F., 1990, *Niels Bohr*, Paris, Criterion.

PEACOCKE, C., 1985, "Imagination, experience and possibility: a Berkeleian view defended", in *Essays on Berkeley: a Tercentennial Celebration*, Oxford Clarendon Press.

RAMSEY, F. P., 1929, "General propositions and causality", in *Philosophical Papers*, Cambridge University Press, 1990.

REDHEAD, M., 1980, 'Models in Physics', *British Journal for the Philosophy of Science*, 31 : 145-163.

SKLAR, L., 1993, 'Idealization and Explanation: A Case Study from Statistical Mechanics', *Minnesota Studies in Philosophy*, vol. XVII: Philosophy of Science, The University of Notre-Dame Press, 258-270.

SUPPES, P., 1974, "The Structure of Theories and the Analysis of Data", in F.Suppe, ed., *The Structure of Scientific Theories*, University of Illinois Press.

VAN FRAASSEN, B., 1994, *Lois et Symétrie*, Vrin.

WALTON, K., 1990, *Mimesis as Make-Believe. On the Foundations of the Representational Arts*, Harvard University Press, 1990.

WILLIAMS, 1973, "Imagination and the Self", in *Problems of the Self*, Cambridge University Press.