

Le naturalisme à l'épreuve du dilemme de Hempel

Résumé

Le dilemme de Hempel constitue une objection classique au physicalisme et, plus généralement, au naturalisme scientifique : défini par référence à la physique actuelle, le naturalisme est probablement faux ; défini par référence à une physique future complète, il devient trivial ou vide. Cet article examine plusieurs stratégies contemporaines visant à répondre à ce dilemme. Après avoir présenté la solution minimale défendue par Papineau et la version plus ambitieuse proposée par Melnyk, il soutient qu'une réponse naturaliste plus substantielle devient disponible si l'on prend au sérieux la structure effective de la physique contemporaine. En s'appuyant sur la notion de Core Theory et sur le rôle central des théories effectives en théorie quantique des champs, l'article défend l'idée que les lois physiques pertinentes pour la chimie, la biologie et les sciences spéciales sont robustes et largement découplées de la physique de plus haute énergie. Cette effectivité permet de formuler un naturalisme empiriquement informé, non trivial et faillibiliste, tout en reconnaissant que l'absence de consensus ontologique sur l'interprétation de la mécanique quantique demeure un défi interne majeur pour toute vision naturaliste du monde.

Le dilemme de Hempel

Ce dilemme est formulé par le philosophe des sciences Carl Gustav Hempel à la fin des années 1960, dans le cadre de discussions au sein du positivisme logique sur le physicalisme, le réductionnisme et l'unité des sciences. Il est originellement adressé au physicalisme, mais il peut s'appliquer de la même manière au naturalisme, qui constitue un cadre plus général.

Le physicalisme est la thèse selon laquelle tous les processus de l'univers (qu'ils soient chimiques, biologiques, mentaux ou sociaux) sont réalisés par des processus physiques. Hempel n'était pas hostile à cette vision du monde et la partageait dans une large mesure. Son dilemme a été conçu pour forcer les physicalistes à préciser dans quelles conditions une philosophie physicaliste pouvait être formulée. La prémisse du dilemme est que, la thèse physicaliste faisant référence à des processus physiques, elle doit être indexée sur les théories de la physique afin de préciser ce qu'elle énonce.

Première branche du dilemme :

Si, pour formuler la thèse du physicalisme, on doit se référer à ce que les théories physiques actuelles disent du monde, alors cette thèse est vraisemblablement fausse, car nous ne disposons pas des théories ultimes de la physique et il est probable que les théories actuellement acceptées soient incomplètes, imprécises ou fausses.

Deuxième branche du dilemme :

Si, pour formuler la thèse du physicalisme, on doit se référer à un état futur et achevé de la physique, alors cette thèse est peut-être vraie, mais elle est vide, car on ne sait pas quelles seront les théories ultimes de la physique.

Le dilemme laisse donc deux choix au physicaliste : soit endosser une thèse qui est fausse, soit endosser une thèse qui est vide.

La solution de Papineau

Le philosophe David Papineau cherche à échapper au dilemme en défendant une approche négative du physicalisme. Pour lui, le physicalisme n'a pas besoin d'être indexé sur des théories physiques pour avoir un contenu : il suffit de préciser ce qu'il exclut. Il exclut notamment toute entité ou causalité non physique, c'est-à-dire des entités comme des esprits désincarnés, une force vitale, une âme cartésienne, des forces psychiques ou des interventions surnaturelles. Le cœur de la position physicaliste de Papineau réside dans le principe de clôture causale du physique : tout effet physique possède une cause physique suffisante. Par exemple, lorsque je tape sur les touches de mon clavier, même si des processus mentaux interviennent, les chaînes causales tout au long de la séquence sont réalisées de part en part par des processus physiques qui se déroulent notamment dans mon système nerveux et dans mon cerveau. À aucun moment une causalité non physique n'est nécessaire pour expliquer le phénomène. Pour Papineau, les sciences spéciales (chimie, biologie, psychologie) n'ont jamais eu besoin d'introduire des causes qui n'étaient pas physiques. On peut donc émettre, par induction à partir de l'histoire des sciences, la thèse que cet état de fait se perpétuera à l'avenir. La définition du physicalisme proposée par Papineau peut ainsi correspondre à l'énoncé selon lequel il existe un rapport de détermination causale entre le physique et tous les autres phénomènes de l'univers. Cette thèse n'est pas vide, car on peut imaginer un état du monde dans lequel cette relation causale ne serait pas réalisée. Elle serait réfutée, par exemple, si l'on démontrait que l'esprit a besoin d'une âme pour fonctionner ou si l'on prouvait l'existence de Dieu, de la vie après la mort ou de phénomènes psychiques.

La solution de Melnyk

Je considère la solution de Papineau comme valable et comme dessinant un physicalisme à minima. Il s'agit cependant d'une approche purement négative et défensive, qui ne propose pas une vision du monde très structurée ni fortement engagée sur le plan ontologique. Elle peut néanmoins servir de solution de repli si d'autres définitions plus ambitieuses du physicalisme échouent. Andrew Melnyk propose, quant à lui, un physicalisme plus ambitieux. Pour cela, il choisit d'affronter directement la première branche du dilemme et accepte l'idée que le physicalisme soit indexé sur nos meilleures théories physiques actuelles. Il nous invite alors à donner à la thèse du physicalisme le même statut qu'à une théorie scientifique, c'est-à-dire celui de notre meilleure façon, à un moment donné, de rendre compte d'une réalité. Comme toute théorie scientifique, le physicalisme est une hypothèse faillible : il pourrait être faux, mais ce n'est pas une raison de ne pas le prendre au sérieux. La solution de Melnyk évite clairement la deuxième branche du dilemme, car elle est fondée sur les théories actuelles de la physique. Pour lui, le contenu du physicalisme est très riche — plus riche que celui de Papineau — puisqu'il fait référence à tout le contenu des théories physiques actuellement admises. Mais alors, comment répond-il à la première branche du dilemme ? C'est ici que les choses se gâtent. Melnyk admet que le physicalisme, tel qu'il est présenté, est probablement faux. Pour lui, ce n'est cependant pas un problème : puisqu'il a le même statut qu'une théorie scientifique, le fait

qu'il soit probablement faux ne constitue pas une raison de ne pas l'accepter comme notre meilleure hypothèse de travail. Il suffit, indique-t-il, qu'il soit plus probable que les hypothèses concurrentes. Pour Melnyk, accepter une hypothèse ne signifie pas penser qu'elle a plus de 50 % de chances d'être vraie. Cela veut plutôt dire la considérer comme plus probable que les hypothèses disponibles au même moment.

Je suis personnellement d'accord avec le raisonnement proposé par Melnyk jusqu'à un certain point. Le physicalisme peut être considéré comme faisant référence aux théories physiques actuelles et doit également être envisagé de la même manière qu'une théorie scientifique, c'est-à-dire comme notre meilleure façon, à un moment donné, de rendre compte d'une réalité, et comme quelque chose de révisable et de falsifiable. Cependant, je trouve un peu étrange l'idée qu'accepter une théorie ne signifie pas considérer qu'elle a plus de 50 % de chances d'être vraie. Je comprends pourquoi Melnyk peut le penser. Il considère que les théories physiques actuelles ne sont probablement pas vraies au sens strict, mais que nous avons en même temps les meilleures raisons du monde de les accepter. L'idée est qu'un dépassement des théories physiques actuelles permettrait d'affiner la position physicaliste sans la renverser radicalement, exactement comme la relativité a remplacé la physique newtonienne sans la reléguer à la poubelle de l'histoire des sciences. Je pense que Melnyk est largement sur la bonne voie pour proposer une version robuste et riche du physicalisme, mais qu'il fragilise inutilement son propos en admettant que le physicalisme tel qu'il est formulé aujourd'hui est probablement faux. Cette concession repose, selon moi, sur une façon inadéquate de considérer le problème de la vérité dans les sciences (voir mon article sur le mythe du mythe de la vérité). Je vais à présent formuler ma propre solution au dilemme de Hempel, qui repose en grande partie sur celle de Melnyk.

Généralisation de la discussion au naturalisme

Le dilemme de Hempel a été historiquement adressé au physicalisme, et c'est dans le cadre des discussions sur le physicalisme que Papineau et Melnyk y ont répondu. Cependant, ce dilemme s'applique à toute position ontologique indexée sur des théories scientifiques. Le naturalisme que je défends sur ce blog est donc pleinement concerné par ce dilemme. En réalité, le physicalisme n'est qu'un cas particulier du naturalisme auquel le dilemme de Hempel s'applique. J'adhère moi-même au physicalisme et j'aurais très bien pu en rester à une discussion sur celui-ci. Mais, pour rendre la discussion aussi générale que possible, il suffit de remplacer « physicalisme » par « naturalisme » et « théories physiques » par « théories scientifiques ». Le dilemme de Hempel généralisé s'énonce alors ainsi :

Première branche du dilemme :

Si, pour formuler la thèse du naturalisme, on doit se référer à ce que les théories scientifiques actuelles disent du monde, alors cette thèse est vraisemblablement fausse, car nous ne disposons pas des théories scientifiques ultimes et il est probable que les théories actuellement acceptées soient incomplètes, imprécises ou fausses.

Deuxième branche du dilemme :

Si, pour formuler la thèse du naturalisme, on doit se référer à un état futur et achevé des sciences, alors cette thèse est peut-être vraie, mais elle est vide, car on ne sait pas quelles seront les théories scientifiques ultimes.

The Core Theory

Je vais désormais proposer ma solution complète à ce problème. Mon raisonnement commence de la même manière que celui de Melnyk. Le naturalisme est une thèse qui fait référence aux théories scientifiques actuelles ; ma solution n'est donc pas concernée par la deuxième branche du dilemme. Le naturalisme est également une théorie révisable et faillible, comme n'importe quelle théorie scientifique. Cependant, contrairement à Melnyk, je ne considère pas que le naturalisme soit probablement faux. La raison principale est que je ne considère pas que les théories scientifiques actuellement les mieux établies soient probablement fausses. Comme je l'ai expliqué plus longuement dans mon article sur la vérité dans les sciences, les théories scientifiques les mieux établies contiennent des énoncés que l'on peut rendre strictement vrais en précisant les conditions sous lesquelles ils s'appliquent, ainsi que des modèles qui ne sont, à strictement parler, ni vrais ni faux, mais qui captent plus ou moins précisément des structures pertinentes de la réalité. Le point décisif pour mon propos est le suivant : nous avons aujourd'hui d'excellentes raisons de penser que les théories scientifiques qui décrivent les processus les plus fondamentaux de la physique de la vie quotidienne — si l'on exclut des domaines exotiques comme la dynamique galactique et extragalactique, l'expansion de l'Univers, les premiers instants après le Big Bang et ce qui se passe au centre d'un trou noir — sont définitivement stabilisées sur le plan épistémologique. « Définitivement stabilisées » ne signifie évidemment pas à l'abri de toute révision théorique, mais robustes au sens des théories effectives : leur domaine de validité et leurs structures causales pertinentes ne dépendent pas des détails de la physique plus fondamentale. Pour le dire plus simplement, nous pensons connaître les règles du jeu pertinentes qui sont à l'origine de tous les phénomènes de la vie quotidienne. Cette affirmation peut sembler surprenante et prétentieuse. Elle a pourtant été remarquablement défendue par le physicien théoricien Sean Carroll. Avant de la présenter rigoureusement, il faut expliciter ce que cette thèse ne dit pas :

- Elle ne dit pas que nous disposons d'une théorie du tout.
- Elle ne dit pas que nos théories physiques nous permettent de tout comprendre dans le monde de la vie quotidienne. Nous avons encore besoin de la chimie, de la biologie, de la psychologie, de la sociologie, etc., pour comprendre le monde dans lequel nous vivons. Le fait que les processus physiques fondamentaux qui sous-tendent tous ces domaines soient connus ne signifie pas qu'ils nous apportent un éclairage pertinent sur ces domaines. La thèse est ici celle d'un réductionnisme faible ou ontologique, et non d'un réductionnisme fort ou épistémologique. Autrement dit, cette thèse repose sur l'idée que, si les sciences spéciales comme la chimie, la biologie ou la psychologie sont conceptuellement indépendantes de la physique fondamentale, les processus qu'elles décrivent sont néanmoins réalisés par des processus physiques.

Alors, que dit cette thèse ? Le point suivant est technique, mais il est central pour comprendre pourquoi la thèse défendue ici n'est pas une simple pétition de principe philosophique. Elle part du constat que nous disposons d'un ensemble de théories, que Sean Carroll et quelques autres physiciens appellent The Core Theory (TCT), composé du modèle standard des particules et de la théorie de la relativité générale (ou d'une version quantique de cette dernière théorie à basse énergie). Le fait que TCT soit une théorie effective est absolument central pour l'argument : on sait qu'il ne s'agit pas d'une théorie fondamentale — même si c'est la théorie établie la plus fondamentale dont nous disposons actuellement —, mais sa structure générale ne dépend pas des détails de la physique encore plus fondamentale sur laquelle elle repose. Les théories effectives s'appliquent avec une précision impressionnante dans leur domaine de validité, tout en ignorant par construction les processus les plus fins à partir desquels elles émergent. Le concept de théorie effective est un développement de la fin du XXe siècle qui s'est révélé central en physique des particules et en physique de la matière condensée. Il repose sur le cadre théorique de la théorie quantique des champs (TQC), dans lequel le modèle standard de la physique des particules est formulé. Pour résumer, l'idée tient au fait que les calculs rigoureux en TQC sont impossibles à effectuer et que l'on mobilise, pour pouvoir utiliser la théorie, une procédure appelée renormalisation qui, en quelque sorte, gomme et « moyennise » tout ce qui se passe à partir d'une certaine échelle. Plus précisément, on introduit une « coupure » dans les calculs, qui correspond à l'échelle spatiale à partir de laquelle on choisit d'ignorer ce qui se passe. Bien sûr, il se passe quelque chose à ces échelles ; plus encore, si l'on comprenait la théorie de manière naïve, la contribution de ce qui se passe en dessous de ces échelles correspondrait à des nombres infinis. Deux raisons légitiment physiquement ce que beaucoup ont considéré, au début de l'établissement de cette procédure, comme un tour de passe-passe : 1) nous ignorons en réalité la physique de ce qui se passe en dessous de ces échelles, et il serait donc illégitime d'extrapoler nos théories jusqu'à elles ; 2) les paramètres associés aux champs quantiques — masses et charges des particules — dépendent de ces échelles, et les contributions ignorées dans les calculs à cause de la coupure sont en fait absorbées dans la définition de ces paramètres. Les résultats obtenus sont insensibles à l'échelle choisie, pourvu que celle-ci corresponde à un domaine d'énergie beaucoup plus élevé que celui du phénomène étudié. En physique, il existe une correspondance entre l'échelle spatiale et l'énergie mise en jeu : une particule d'une énergie donnée est associée à une longueur d'onde d'autant plus petite que son énergie est élevée. En pratique, cela signifie que les échelles sont découplées. La physique à relativement basse énergie qui correspond à TCT ne dépend pas des détails des théories à très haute énergie ou des théories de la gravitation quantique, dont elle n'est pourtant qu'une approximation aux basses énergies. Or, tous les phénomènes que l'on peut observer sur Terre, ou qui interagissent sensiblement avec ce que l'on peut y observer, relèvent d'un domaine de basse énergie entièrement compris dans le domaine de validité de TCT. En particulier, les processus chimiques, biologiques, psychologiques, sociologiques, etc., ne dépendent donc pas des détails de la physique située au-delà de TCT. Plus important encore, la découverte de cette nouvelle physique ne renversera pas TCT. D'une part, il existe d'excellentes raisons de penser que, si quelque chose d'essentiel manquait à TCT dans son domaine de validité, on l'aurait découvert dans les accélérateurs de particules disponibles ; cette affirmation repose sur un argument technique de théorie quantique des champs dont on peut ignorer ici les détails. D'autre part, il existe un théorème informel en physique — c'est-à-dire un théorème qui n'est pas établi rigoureusement, mais dont on a d'excellentes raisons de penser qu'il est vrai — selon lequel toute théorie qui respecte à la fois les principes de la mécanique quantique et

ceux de la relativité prend la forme d'une théorie quantique des champs aux basses énergies. En bref, TCT n'est pas le dernier mot de la physique fondamentale, mais le découplage des échelles indique que cette théorie et le cadre théorique sur lequel elle repose — la théorie quantique des champs — sont faits pour durer. Ils excluent l'existence de phénomènes comme l'élan vital, l'âme et les pouvoirs parapsychiques. Ils peuvent se révéler faux, mais c'est le lot de toute théorie scientifique. En attendant, il s'agit des théories les mieux confirmées de toute l'histoire des sciences, et aucun phénomène établi sur Terre ne les a jamais fait vaciller.

Ce qui est vraiment important pour l'argument développé ici, c'est que l'effectivité de TCT — à une condition que j'évoquerai dans la conclusion — est suffisante pour définir la thèse du physicalisme et, par extension, celle du naturalisme, de manière précise, sans risquer de les faire dépendre d'une théorie qui sera probablement balayée par l'histoire des sciences. Nous disposons donc d'une thèse extrêmement riche — qui échappe très facilement à la deuxième branche du dilemme — sans avoir a priori de raison de penser qu'elle est probablement fausse.

Conclusion

L'objectif de cet article n'est pas de montrer que le physicalisme ou le naturalisme est vrai. Il est de préciser une version à la fois robuste et riche du naturalisme, qui ne soit pas trivialement fausse tout en proposant un engagement ontologique positif à l'égard du monde. Selon cette approche, le monde que nous observons est entièrement réalisé par des processus physiques adéquatement décrits par TCT. Cette théorie ne suffit cependant pas à décrire toute la richesse structurelle que nous pouvons observer à d'autres échelles. Les phénomènes chimiques, biologiques et sociaux sont des phénomènes émergents, adéquatement décrits par les théories des sciences spéciales qui les concernent. Mais il s'agit d'une émergence au sens faible : TCT ne permet pas de déduire la chimie, la biologie ou la sociologie, mais elle contraint drastiquement les phénomènes que l'on peut y observer. Une analogie peut aider : la connaissance des règles du jeu d'échecs ne permet pas de déduire les meilleures stratégies possibles pour gagner une partie. Elle contraint cependant les stratégies utilisables et interdit celles qui violent ces règles. De la même manière, la connaissance de TCT ne permet pas de connaître notre monde à notre échelle, mais elle contraint fortement les formes qu'il peut prendre. Elle exclut notamment toute non-conservation de l'énergie, tout transfert d'information ou d'énergie dépassant la vitesse de la lumière, et tout nouveau type de force ou d'énergie qui aurait des effets causaux dans le domaine de la vie quotidienne. En particulier, si l'on excepte la gravitation, responsable de l'accélération vers le centre de la Terre que nous subissons, ainsi que les processus nucléaires, elle nous dit que tous les phénomènes de la vie quotidienne — chimie, vie, propriétés de la matière, processus mentaux — sont fondamentalement d'origine électromagnétique. Dans tous les cas, elle laisse entrevoir un monde qui n'est pas gouverné par des processus fondamentaux téléologiques, c'est-à-dire orientés vers des fins, mais par des processus impersonnels respectant des symétries abstraites et des lois de conservation. Elle ne laisse pas de place à une force vitale, à des âmes, à des pouvoirs parapsychiques, à la clairvoyance, à la télépathie ou à la télékinésie.

Pour s'imposer, cette vision du monde doit toutefois relever quelques défis. Elle doit bien évidemment résister à toute proclamation de l'établissement de phénomènes qui violeraient ses contraintes de base, ainsi qu'à toute tentative de démonstration d'un phénomène

strictement irréductible à TCT — les meilleurs candidats étant l'émergence de la vie et celle de la conscience phénoménale. Mais elle fait également face à un défi moins connu, reconnu par de nombreux physicalistes, et interne au physicalisme lui-même. TCT est fondamentalement une théorie quantique, et il est extrêmement difficile d'en extraire une ontologie. Il est aujourd'hui largement admis, en philosophie de la physique, que l'interprétation « des manuels » de la mécanique quantique n'est fondamentalement qu'un ensemble de recettes portant sur les probabilités de résultats expérimentaux. Son statut même de théorie physique — c'est-à-dire de théorie portant sur le monde en soi — est remis en question par des personnes positionnées de part et d'autre du débat entre réalisme et antiréalisme. Or, la position naturaliste que je défends ici implique le réalisme scientifique, et donc l'idée que les théories scientifiques nous donnent une description du monde réel. Le fait que la théorie physique qui occupe la position la plus centrale de tout cet édifice ne possède pas d'interprétation ontologique minimale consensuelle devrait constituer un motif de préoccupation majeur pour une position comme la mienne. C'est la raison pour laquelle je n'hésiterai pas, dans ce blog, à aborder les problèmes d'interprétation de la mécanique quantique.

Bibliographie

Carl Gustav Hempel

- **Hempel, C. G.** (1969). *Reduction: Ontological and Linguistic Facets*. In **S. Morgenbesser, P. Suppes & M. White (eds.)**, *Essays in Honor of Ernest Nagel*. New York: St. Martin's Press.

Texte fondateur où apparaît le dilemme portant sur la définition du « physique », devenu classique sous le nom de *dilemme de Hempel*.

David Papineau

- **Papineau, D.** (1993). *Philosophical Naturalism*. Oxford: Blackwell.

Présentation systématique du naturalisme et du physicalisme comme thèses empiriques et faillibilistes.

- **Montero, B. & Papineau, D.** (2005). *A Defence of the Via Negativa Argument for Physicalism*. *Analysis*, 65(3), 233–237.

Réponse explicite aux difficultés définitionnelles du physicalisme, directement pertinente pour le dilemme de Hempel.

Andrew Melnyk

- **Melnyk, A.** (2003). *A Physicalist Manifesto: Thoroughly Modern Materialism*. Cambridge: Cambridge University Press.

Défense ambitieuse d'un physicalisme a posteriori, explicitement formulé comme une hypothèse scientifique faillibiliste.

Sean Carroll

- **Carroll, S.** (2019). *Something Deeply Hidden: Quantum Worlds and the Emergence of Spacetime*. New York: Dutton.

Présentation accessible mais rigoureuse de l'interprétation d'Everett et de la structure ontologique sous-jacente à la physique contemporaine.

- **Carroll, S.** (2016). *The Big Picture: On the Origins of Life, Meaning, and the Universe Itself*. New York: Dutton.

Défense explicite d'un naturalisme fondé sur la *Core Theory* et les théories effectives.