

Une dissolution naturaliste du problème du fossé explicatif

(draft)

François-Igor PRIS

frigpr@gmail.com

Le problème du fossé explicatif, ou le problème difficile, dans la philosophie de l'esprit est une incarnation moderne du vieux combat entre le matérialisme et le dualisme. Le dualisme de substances avait été rejeté unanimement, mais les réfutations du dualisme de propriétés sont beaucoup moins convaincantes. (White 2006) En effet le problème du FE équivaut au problème du choix entre le physicalisme et le dualisme de propriétés. (Chalmers 1996, 2001, 2006; Block 2005) Tout d'abord le FE est le fossé épistémique entre les propriétés phénoménales et les propriétés physiques non-phénoménales sensées représenter ou décrire celles-ci. L'intuition nous dit que les propriétés phénoménales ne sont pas concevables, ou ne sont pas dérivables *a priori*, à partir des propriétés physiques non-phénoménales. Selon Joseph Levine, qui a introduit le problème du FE, le fossé épistémique n'implique pas nécessairement le fossé métaphysique (ontologique) entre deux types de propriétés. (1983) Pourtant, supposer l'existence d'une telle implication est assez naturel, car le concept de connaissance comprend la vérité de nos croyances et donc l'existence de faits connaissables. *Prima facie* on peut supposer l'existence des faits physiques et des faits phénoménaux qui ne sont pas de faits physiques. C'est ce que déduit par exemple Franck Jackson de son argument de la connaissance: outre la connaissance physique, il existe la connaissance phénoménale - la connaissance directe de l'expérience phénoménale - qui n'est pas celle physique et ne peut pas être dérivée *a priori* à partir de toute connaissance physique de l'expérience; par conséquent, la connaissance phénoménale est celle de faits non-physiques. (1986) Selon Stalnaker (2006) il n'y a pas de connaissance (ou d'information) phénoménale objective concernant des faits non-physiques. La connaissance complète inclut, selon lui, ce qu'il appelle l'"information subjective", ou l'"information essentiellement contextuelle". John Perry a traité l'argument de la connaissance et d'autres problèmes de la philosophie de l'esprit à l'aide de sa notion de la "proposition réflexive" qui entend une relation causale entre le sujet et le fait. (2001, 2004) Je partage l'avis de Ned Block (2005) que Perry ne prend pas en considération la phénoménalité proprement dite, et je pense que la même chose peut être dite à propos de l'idée de Stalnaker qui s'inspire du traitement que donne Perry aux puzzles posés par les croyances essentiellement indexicales et auto-situées. De mon côté, j'essaie de comprendre la phénoménalité en termes de "jeux de langage" phénoménaux, en m'inspirant de la deuxième philosophie

de Wittgenstein. A la différence de Stalnaker, pour moi, l'information phénoménale existe, mais dans un sens elle est "subjective" (par analogie avec la notion de l'information subjective chez Stalnaker) et implicitement physique. C'est cette information/connaissance spécifique qui permet de remplir le FE. La notion de l'information totale purement physicaliste n'a pas de sens. Pour employer le mot de Brandom (1994), le "physicalisme désenchanté" est faux. Le physicalisme/naturalisme acceptable doit avoir une dimension normative. A mon avis, Wittgenstein fournit un physicalisme/naturalisme semblable. J'introduis notamment les jeux de langage réflexifs dont les propositions réflexives de Perry sont un cas particulier. La relation causale entre le sujet et le fait est remplacée par la relation réflexive/introspective, à la fois naturelle et normative (spontanée), de la ressemblance familiale entre le jeu de langage instinctif et le jeu de langage réflexif.

Le dualisme conceptuel propose de remplir le FE ontologique à l'aide de deux sortes de concepts: physiques et phénoménaux (initialement introduits par Loar 1990). Néanmoins le problème du FE se déplace au niveau conceptuel. Selon Levine, le physicaliste doit expliquer l'existence du fossé conceptuel du point de vue physicaliste (la "contrainte matérialiste"). Autrement dit, il doit trouver une explication physicaliste de la nature spécifique des concepts phénoménaux. C'est le problème du second FE, posé par Levine (2003-2006). Levine arrive à la conclusion que le problème du second FE ne peut pas être résolu sans faire appel à quelque chose comme la notion de l'*accointance* de Russell. L'analyse de Chalmers du même problème aboutit au dilemme suivant: soit le concept phénoménal peut être expliqué du point de vue physicaliste, et dans ce cas il ne peut pas représenter notre situation épistémique, soit il peut représenter notre situation épistémique, mais dans ce cas il ne peut pas être expliqué du point de vue physicaliste. (2006) C'est une conclusion explicitement anti-physicaliste. Je suis d'accord avec Papineau (2006) qui affirme que les explications de Chalmers sont trop abstraites et exigeantes. La propriété *usage/mention* (voir plus bas) des CP permet de concilier leur nature physique avec l'existence du fossé épistémique. Sur cette question la position de Papineau s'accorde avec la mienne, à la différence près que j'ai recours à la notion wittgensteinienne de l'usage de la règle. La relation naturelle/spontanée entre les usages (jeux de langage) réflexifs et usages instinctifs joue le rôle de l'*accointance* évoqué par Levine.

D'autre part, je pense que ce que Block appelle « le mouvement méta-phénoménal » (2005) suggère que le dualisme conceptuel entraîne le dualisme

tout court.¹ Donc, encore une fois, le problème consiste à comprendre les concepts phénoménaux dans le cadre d'une théorie physicaliste. Pour cela Block (2005) et Papineau (2006) introduisent un trait fondamental des concepts phénoménaux: les concepts phénoménaux à la fois *utilisent* et *mentionnent* l'expérience phénoménale. L'expérience initiale est présente implicitement (est mentionnée) dans l'usage du CP. Ce qui est explicite c'est le mode de présentation de l'expérience. Selon Block, les CP sont des concepts anormaux d'une espèce naturelle : ils sont étroits et non-transparents à la fois. Ils sont implicitement (ontologiquement) physiques. Le problème du FE au niveau métaphysique se dissout. Pour Block le FE est purement épistémique. Je parle plutôt d'une illusion épistémique. Dans un sens le FE se dissout au niveau métaphysique ainsi qu'au niveau épistémique.

La science est capable de comprendre l'expérience phénoménale comme l'expérience physique; elle est capable de remplir le FE. Elle découvre et établit des identités comme $Q \text{ est } P$, où Q est une propriété phénoménale, P est la propriété physique correspondante à laquelle Q est identique. L'identité $Q \text{ est } P$ nous est donnée à travers des différents modes de présentation Q' et P' associés avec les termes Q et P respectivement. Le problème ontologique consiste à comprendre la copule *est*. Du point de vue épistémique la corrélation entre deux modes de présentation est une loi nomologique; du point de vue ontologique c'est en plus ce qu'on appelle l'identité nécessaire a posteriori (introduite par Kripke (1980)). Dans ce sens, la nature de l'identité psycho-physique n'est pas fondamentalement différente de celle de l'identité physico-physique, comme celle entre la température et le mouvement des molécules, $T=M$. La fameuse objection de Kripke (1972, 1980) consistant à dire que l'expérience phénoménale nous est donnée directement et donc qu'elle ne pourrait pas (n'aurait pas pu) être autre chose que ce qu'elle est, néglige la nature non-transparente des CP.

Les identités théoriques, comme $Q=P$ et $T=M$, entrent dans les théories scientifiques développées, considérées comme des pratiques normatives, en tant que règles implicites ou explicites d'usage, préalablement établies. Leur découverte et leur établissement sont deux autres pratiques normatives. La même règle $Q=P$ gouverne des pratiques de la découverte (de la naissance), de la justification de l'identité $Q \text{ est } P$, des pratiques de son usage en tant qu'identité établie (acceptée) et des pratiques théoriques développées. En fait la formulation propositionnelle de la règle peut être identifiée avec son usage établi (On sait par exemple que l'eau est l' H_2O . C'est un usage établi de la règle "L'eau est l' H_2O ",

¹ Block lui-même ne fait pas cette affirmation.

qui peut être identifié avec la règle elle-même. La même règle gouverne la pratique de la découverte du fait que l'eau est l'H₂O.), qui peut être vu comme une *proposition-gond* de Wittgenstein. J'ai plutôt de la sympathie pour la notion naturaliste de la pratique normative, développée par Rouse dans (2002), dans le cadre de son retournement naturaliste du pragmatisme normatif de Brandom (1994, 2000). On peut affirmer que le même type d'un "naturalisme normatif" contient la deuxième philosophie de Wittgenstein (W). (1953) Le jeu de langage spontané/naturel de W gouvernant par des règles généralement implicites est un exemplaire de la pratique normative naturaliste. On peut définir le "naturalisme normatif" ainsi : la nouvelle application de la règle doit à la fois être "naturelle/spontanée" et *post factum* naturellement/spontanément justifiable. La naturalité/spontanéité wittgensteinienne est une généralisation et une synthèse de la spontanéité purement intellectuelle et de la sensibilité purement matérielle kantienne. (Ch. 1)

Nous considérons donc l'identité $Q \text{ est } P$ comme une pratique normative ou un jeu de langage naturel/spontané, ayant notamment les règles $Q=P$ (proposition), Q (concept), P (concept), implicites ou explicites. Ces règles sont associées à l'identité dans un acte de réflexion/introspection ou dans le processus d'une investigation scientifique. Nous introduisons la distinction relative entre les jeux de langage instinctifs et les jeux de langage réflexifs. Si l'on néglige la copule *est*-réflexive (i.e. si l'on néglige l'acte de la réflexion en tant qu'acte) - c'est facile, car, de même que la copule *est*-instinctive, la copule *est*-réflexive, n'est pas transparente - l'identité réflexive $Q' \text{ est } P'$, où Q' et P' sont des modes de présentation des propriétés Q et P (i.e., les propriétés Q et P , faites *explicites* (comme dans le pragmatisme normatif de Brandom, naturalisé par Rouse) partiellement ou complètement), se transforme en corrélation entre les modes Q' et P' , dont j'ai déjà parlé ci-dessus. (Chapitre 2, 3)

En termes wittgensteiniens je comprends le FE comme un fossé entre la règle P et son application - le jeu de langage $Q \text{ est } P$. C'est la naturalité/spontanéité wittgensteinienne qui remplit ce fossé. Au niveau épistémique, c'est le passage continu (selon la relation de RF réflexive/introspective, ou selon la copule *est*-réflexive) entre les modes Q' et P' ayant le référent commun qui remplit le FE. La naturalité/spontanéité réflexive du mode de présentation joue le rôle de l'*a priori* naturalisé. Du point de vue wittgensteinien, le mode de présentation est donc à la fois *a priori* accessible (au cours d'une analyse conceptuelle ou phénoménologique), a une fonction cognitive (car il n'est pas transparent) et fournit un accès *a priori* au référent (en tant qu'application de la règle correspondante), tandis que du point de vue standard (Burge 1977), ces trois fonctions du mode de présentation ne vont pas toujours ensemble. On peut interpréter la distinction de Block (2005) entre les modes métaphysiques et modes conceptuels comme celle entre les modes (jeux de langage) plutôt instinctifs et les modes (jeux de langage) plutôt réflexifs et j'interprète sa notion

du CP et ses arguments contre le dualisme de propriétés en termes wittgensteiniens. Nous comprenons la propriété *mention/usage* des CP du point de vue du slogan fondamental de Wittgenstein "le sens est l'usage". Conformément à la philosophie de l'esprit il prend la forme "la conscience est l'usage", ou "le sens phénoménologique est l'usage de l'expérience physique correspondante". L'expérience phénoménale est donc à la fois utilisée et mentionnée en tant que règle physique de l'usage. Elle est donc implicitement (ontologiquement) physique. En général le physicalisme donc est "implicite", ou a une dimension normative. La théorie physique toute seule, qu'elle que soit riche, ne peut pas remplir le FE. Elle joue le rôle de la règle. La clause wittgensteinienne de la naturalité/spontanéité entre la théorie et l'expérience est nécessaire. Outre cela, la naturalité/spontanéité, comme un éther, est distribuée au sein de la théorie elle-même; elle est nécessaire pour comprendre celle-ci. (Ch. 2, 3)

Il y a une relation de ressemblance familiale (RF) entre les jeux de langage réels, mémoriels et imaginatifs, qu'ils soient instinctifs ou réflexifs. De façon assez conventionnelle, le concept phénoménal, vu comme une "copie" de l'expérience phénoménale, pourrait se rapporter par exemple au niveau réflexif ou mémoriel. Mais de façon plus stricte, nous considérons le concept comme une règle wittgensteinienne. De ce point de vue, le concept phénoménal est la capacité de produire ou d'identifier l'expérience phénoménale (instinctive, réflexive, imaginative, etc.) - une sorte de disposition spécifique acquise dans l'expérience quotidienne. Nous pensons que cette vision est assez proche de celle originale de Loar. (1990) Nous entendons l'affirmation de Papineau (2002, 2006) qu'il existe une similarité entre les CP de types différents et entre les CP et les concepts perceptifs au sens de la RF de Wittgenstein, que nous interprétons comme un jeu de langage du second ordre, dont la naturalité/spontanéité est plus faible que celle de jeux de langage du premier ordre, qu'elle relie. (Ch. 5)

Nous considérons la nécessité a posteriori de Kripke comme un instrument théorique approximatif pour décrire l'identité comme un jeu de langage naturel/spontané. Il est concevable que l'eau aurait pu ne pas être l'H₂O. Selon Kripke, c'est une illusion purement épistémique due à notre état d'ignorance: Si nous ne savions pas que la substance aqueuse (i.e. ayant les propriétés superficielles de l'eau: liquide, transparente, incolore, etc.) est l'H₂O, nous pourrions concevoir qu'elle ne l'est pas. On peut être d'accord avec Yablo (2002) sur le fait que l'explication de Kripke de l'illusion de la possibilité n'est pas correcte. En effet, nous avons la même illusion même si nous savons que l'eau est l'H₂O. De mon point de vue, l'illusion de la possibilité est due à la possibilité d'un usage de la règle (du concept) "eau", qui n'est pas celui "L'eau est l'H₂O". Cette possibilité ressemble par exemple à la "possibilité réelle" de Rouse (2002) ou la "possibilité naturelle" de Chalmers (1996.) L'eau aurait pu (et pourrait) ne pas être l'H₂O, car la règle "eau" aurait pu (et pourrait) être utilisée autrement. A

la différence de l'explication de Kripke, notre explication est aussi applicable aux identités psycho-physiques, $Q=P$. La douleur (Q) par exemple aurait pu (et pourrait) être quelque chose d'autre que l'excitation des C-fibres (P). (Ch. 4)

Actuellement, le schème le plus populaire de l'identification psychophysique $M=P$ est le suivant (Kim 1998, 2005, Levine 1999, Jackson 2000) : (1) L'état mental M = l'occupant du rôle fonctionnel F (une analyse conceptuelle *a priori*); (2) L'occupant du rôle F = l'état du cerveau P (une étude empirique); (3) $M=P$ (par transitivité). Pour Block et Stalnaker (1999), l'analyse (1) n'est pas *a priori*. Pour nous, le stade (1) est celui de la description (ou de l'explicitation) naturelle/spontanée de l'usage du concept mental. Le jeu de langage réflexif naturellement/spontanément survient sur le JL instinctif. Il est *a priori* au sens wittgensteinien. Par conséquent, pour nous, le stade (2) ne peut pas être purement empirique; il doit contenir une dimension normative: P est une application naturelle/spontanée empirique de la règle F . Autrement dit, entre le rôle fonctionnel et sa réalisation physique (entre la propriété du second ordre et celle du premier ordre) il existe une relation causale/normative à la fois naturelle et spontanée. Il existe donc une relation naturelle/spontanée entre M et P . Au niveau réflexif, cette relation se présente comme une relation nécessaire, i.e. comme la règle $M=P$. Grâce à la prise en compte de la spontanéité wittgensteinienne, les deux approches principales à la réduction physicaliste - le réductionnisme fonctionnel et la théorie de l'identité - se trouvent être compatibles.

Dans son article (2007, "Functional Reductionism") Block donne une critique du fonctionnalisme réductionniste. Ontologiquement, celui-ci est une position physicaliste, mais métaphysiquement c'est une position fonctionnaliste, car c'est le rôle fonctionnel commun et non pas une propriété commune de ses réalisations physiques qui explique l'identité ou les similarités phénoménales à un tel ou tel niveau d'abstraction. Le fonctionnalisme réductionniste n'est pas complet; ce n'est pas une réduction physicaliste. Selon Block, si le physicalisme échoue ontologiquement, il échoue; mais il est vrai aussi que le physicalisme échoue, s'il échoue métaphysiquement. Les réalisations physiques différentes d'une même propriété phénoménale doivent partager une propriété commune. (2007) En gros, nous partageons cette conclusion avec Block, conclusion à laquelle nous sommes parvenus indépendamment. La contradiction avec la position wittgensteinienne est apparente. De notre point de vue, il doit exister une relation de RF entre les réalisations du même rôle fonctionnel jouant le rôle de la règle wittgensteinienne. La « propriété commune » phénoménale – la relation de la RF - elle-même est implicitement physique. La critique de Block de la position "mixte" de Lewis (1980), combinant le fonctionnalisme et le physicalisme, a des affinités avec notre critique de la même position du point de vue wittgensteinien. (Voir § 2.9, ch. 3.)

Prenons le concept d'eau. Selon l'analyse conceptuelle (*a priori*) de Jackson (2000), l'eau est l'unique substance aqueuse de notre environnement. Block et Stalnaker (1999) modifient cette analyse: l'eau est l'actuelle et l'unique substance aqueuse autour d'ici. Une analyse encore plus précise dit: l'eau est l'actuelle et l'unique substance aqueuse de notre accointance. Cette dernière formulation est la plus correcte. En fait, c'est une description de l'usage principal du concept d'eau. La relation d'accointance joue le même rôle que la relation naturelle/spontanée entre le concept et son application. Si l'eau était définie simplement par son rôle macroscopique fonctionnel (Cet usage n'est pas principal, mais il est possible, et il ne doit pas être confondu avec l'usage principal, qui tient compte de la relation d'accointance.), la découverte de n'importe quelle autre substance, XYZ, remplissant le même rôle devrait être considérée comme découverte d'une autre forme de l'eau. Cela ne semble pas être toujours le plus naturel. (Ch. 6)

Donc le concept d'eau pourrait avoir des usages naturels/spontanés différents de celui l'H₂O. Si par exemple dans une partie lointaine du monde l'XYZ était l'actuelle et l'unique substance aqueuse de notre accointance, elle serait considérée comme l'eau (sous la condition qu'entre les deux usages il existe une relation de RF). Nous interprétons les usages actuels ou possibles de la même règle, du genre "L'eau est l'XYZ", comme *mondes contre-actuels*. Les *mondes contre-factuels* (ceux de Kripke) sont utilisés pour représenter l'identité naturelle/spontanée comme identité nécessaire a posteriori. (Ch. 7)

Chez Kripke lui-même la notion du monde possible (contre-factuel) est intuitive. Par exemple, l'eau, i.e. l'H₂O, pourrait ne pas être une substance aqueuse (dans un monde possible (contre-factuel) elle n'est pas une substance aqueuse). En fait, l'imagination de Kripke n'a pas de contraintes. Nous considérons les mondes contre-factuels comme usages de la règle "H₂O". La contrainte imposée sur notre imagination est celle de la ressemblance familiale entre les usages de la même règle. (Ch. 7)

Il semble que l'interprétation des mondes possibles en termes d'usages ou jeux de langage wittgensteiniens s'accorde avec l'interprétation actualiste du réalisme modal de Lewis (que Lewis lui-même ne partage pas), selon laquelle il y a (*actually are*) beaucoup de variantes en conformité avec lesquels le monde pourrait ou aurait pu être (*might be* ou *might have been*), mais il n'existe qu'un seul monde, qui est l'un de ces variantes. Les mondes possibles sont des propriétés du monde réel. *Pourrait être* est un type de propriétés. (Stalnaker 2005) On peut donc interpréter l'*usage possible* « X est ... » de la règle X comme l'*usage réel* « X pourrait être (aurait pu être) ... » de la même règle. Par exemple l'usage réel "L'eau aurait pu être l'XYZ" de la règle « eau » peut être vu comme l'usage possible "L'eau est l'XYZ" de la même règle.

Chez Lewis les mondes possibles sont des mondes "centrés", ce que signifie

qu'ils tiennent compte du point de vue de l'observateur, de sa position dans l'espace et le temps. Récemment, Stalnaker (2006) a critiqué la notion de Lewis. Selon lui, celle-ci ne permet pas de décrire correctement l'information subjective (voir ci-dessus) qui n'est pas séparable du contexte de l'énoncé. Il semble que la modification de la notion du monde possible proposée par Stalnaker va dans le même sens que notre proposition de voir les mondes possible comme usages wittgensteiniens.

Les mondes contre-actuels et les mondes contre-factuels ensemble, forment ce qu'on appelle le formalisme de la sémantique bidimensionnelle. (Jackson 1998, Chalmers 2006) L'identité nécessaire a posteriori de Kripke se rapporte à la fibre fixe et change avec variation du monde contre-actuel. Block et Stalnaker (1999) ont critiqué les tentatives de résoudre le problème du FE, fondées sur le formalisme bidimensionnel. De leur point de vue, ce formalisme ne permet que de décrire le problème du FE ou une solution déjà proposée. Nous pensons que la prise en compte de la dimension normative, i.e., du fait que les mondes possibles sont des usages (jeux de langage) entre lesquels il y a une relation de RF, permet d'affaiblir l'affirmation critique de Block et Stalnaker. (Ch. 7)

Nous réinterprétons quelques approches à la sémantique bidimensionnelle du point de vue du naturalisme wittgensteinien. Par exemple, nous pensons que l'approche de Chalmers (2005) doit être retournée de façon naturaliste. Chalmers définit la dimension contre-actuelle de la sémantique bidimensionnelle de l'expression - l'"*intension première*" - comme un ensemble des "possibilités épistémiques" dans lesquels l'expression est vraie. Le monde épistémique est un monde possible dont les caractéristiques sont compatibles avec tout ce que nous savons *a priori* à propos de l'expression ou de son *token* dans notre monde. L'expression *a priori* sera vraie dans tous les mondes épistémiques. Chalmers identifie l'*a priori* avec la nécessité de l'intension première. Stalnaker (2004) conteste l'identification de l'*a priori* avec la nécessité.

Chalmers avance deux thèses à propos de mondes épistémiques. (Chalmers 2005) (1) La *thèse de la scrutabilité* (*scrutability*) de la référence et de la vérité dans le monde actuel affirme l'existence d'une description assez complète ("canonique") permettant de dériver dans un langage neutre de notre monde de façon *a priori* (rationnelle) la référence (les valeurs de vérité) de notre expression. (Plus exactement, Chalmers dit que la description canonique "nécessite épistémiquement" l'expression. Selon la définition de Chalmers, ceci est vrai si et seulement si le conditionnel "D implique S", où D est la description canonique, S est l'expression, est *a priori*.) (2) La thèse de la curabilité généralisée affirme que la même chose est vraie dans tous les mondes épistémiques dans lesquels l'expression est vraie. Par exemple, selon (1), en principe on peut dériver de façon *a priori* que l'eau est l'H₂O. Selon la thèse (2), dans le monde épistémique dans lequel l'eau est l'XYZ on peut dériver de façon

a priori (et dans un langage neutre) que l'eau (en tant que l'eau dans ce monde) est l'XYZ. (Ch. 7)

Nous interprétons le monde épistémique comme l'usage naturel/spontané de la règle wittgensteinienne associée avec l'expression. La règle est justement ce qui est connu *a priori*. Son usage est justement ce qui est vrai et compatible avec l'*a priori*. Les deux thèses de Chalmers renvoient au problème du FE formulé en termes de jeux de langage. Il y a un fossé entre une description dite synoptique (conceptuelle) du jeu de langage et ce jeu. Le fossé est rempli "naturellement/spontanément". Le remplissage est justifié rationnellement. Le passage spontané de la description au jeu de langage correspond à la dérivation *a priori* ou à la nécessitation épistémique de Chalmers. La différence entre notre vision wittgensteinienne et celle de Chalmers est que Chalmers néglige le stade naturel/spontané qui précède le stade rationnel.

L'intension première correspond aux usages naturels/spontanés de la règle, liés entre eux par la relation d'une ressemblance familiale. Les mondes épistémiques dans lesquels l'intension première est fautive correspondent aux usages non-spontanés de la règle ou aux usages n'ayant pas de ressemblance avec la famille d'usages déjà prise en considération. A la différence donc de Chalmers nous ajoutons la relation de RF. Les deux thèses de la curabilité de Chalmers reflètent ce que nous venons de dire au niveau rationnel. Mais le niveau rationnel est secondaire. La critique de Chalmers (et Jackson) par Stalnaker (et Block) peut être justement comprise comme ayant à l'origine l'insatisfaction avec le rôle premier accordé au rationalisme.

La même opposition entre le rationalisme et l'empirisme se manifeste dans la controverse entre Chalmers & Jackson (2001) et Block & Stalnaker (1999) à propos de l'analyse conceptuelle et la dérivation *a priori*. Le principe de passage *a priori* de Jackson consiste à affirmer que chaque vraie proposition sur notre monde peut être dérivée de façon *a priori* à partir d'une certaine proposition physique microscopique. (1998, 2000) Par exemple pour eux (et aussi pour Levine (2001)) une vraie proposition sur l'eau peut être dérivée *a priori* à partir d'une vraie proposition sur l'H₂O. Selon Kripke, la relation entre l'une et l'autre est nécessaire, mais *a posteriori*. Selon Chalmers & Jackson, il existe une vérité contingente empirique - la « stop clause » - permettant de transformer la relation nécessaire en celle *a priori*. De notre point de vue la dérivation *a priori* au sens de Chalmers et Jackson est un énoncé naturel/spontané de W. L'introduction de la « stop clause » est une manière de passer de la « nécessité » à la « naturalité/spontanéité ». Selon Chalmers et Jackson (2001), la dérivation *a priori* des faits phénoménaux à partir de faits physiques n'est pas concevable, et donc le FE est ontologique. De notre point de vue, le FE entre le phénoménal et le physique peut être rempli par la naturalité/spontanéité wittgensteinienne, exactement comme le FE entre le physique macroscopique et le physique

microscopique. Chalmers et Jackson se situent du côté rationaliste. Au contraire, Block et Stalnaker (1999) parlent de l'identité empirique. Selon eux, la dérivation *a priori* des termes macroscopiques à partir des termes microscopiques n'est pas possible. En revanche, ils remplissent le FE phénoménal de façon empirique. Quant à nous, nous nous situons au milieu - du côté du naturalisme normatif. (Ch. 6)

Les questions de l'identité, de sa nature et son rôle dans l'explication, de la relation entre la corrélation et l'identité sont transversales. Il y a des divergences importantes entre les positions des philosophes (Smart (1959), Feigl (1958), Block et Stalnaker, Kim (2005), Levine, Jackson et Chalmers, Hill (1991), McLaughlin (2001), Hohwy (2003), etc.) sur ces questions. Nous essayons de les comprendre. Par exemple, la position de Block est la suivante: Les identités n'ont pas d'explications; c'est la corrélation qui peut être expliquée à partir de (ou évitée à l'aide de) l'identité. Les identités sont empiriques, fournies par la science. Mais elles font partie de (sont utiles pour) nos meilleures explications (scientifiques), et c'est la raison pour laquelle elles sont acceptées. (1999, 2007) Hill (1991) et McLaughlin (2001), eux aussi disent que les identités peuvent jouer le rôle explicatif, mais pour des raisons différentes. McLaughlin affirme que les identités métaphysiquement nécessaires peuvent expliquer le caractère nomologique des corrélations psycho-physiques. Pour Kim (2005) l'identité est une règle de ré-description de la meilleure explication, elle n'est pas impliquée dans l'activité explicative. En particulier, Kim nie l'affirmation que les identités expliquent les corrélations. (Kim 2005) Dans le contexte auquel Kim pose la question je suis d'accord avec lui: la "règle de ré-description" est la règle wittgensteinienne, rendue explicite. Récemment Block (2007) a distingué entre les explications causales (dans des contextes transparents) et les explications opaques, et a affirmé que sa position se rapporte au contexte opaque (C'est les explications opaques que les identités autorisent.), celle de Kim - au contexte transparent. Selon Block, pour lui, mais pas pour Kim, les explications opaques (et donc les identités) sont scientifiques; et c'est le noyau du désaccord. Pour comprendre ce problème, je tiens compte de la distinction entre l'identité naturelle/spontanée (ontologique) et l'identité formelle (épistémologique). Pour Smart (1959) et (pour Kim (2005)) les identités et les corrélations s'excluent mutuellement. Si l'on a l'identité il n'y a pas de sens de parler de la corrélation. Pour nous l'identité au sens fort se rapporte au niveau instinctif, la corrélation (ou l'identité au sens faible) - au niveau réflexif. Donc, dans un sens, l'identité et la corrélation s'excluent mutuellement et, dans un autre sens, au sens métaphysique, l'identité explique la corrélation. (Ch. 8)

Quant à l'explication de l'identité, nous distinguons quatre sens suivant lesquels on peut en parler. L'identité naturelle/spontanée est primaire au sens ontologique et dans ce sens elle n'a aucune explication. L'identité formelle (établie) est une règle épistémologique, et dans ce sens elle n'a aucune explication non plus. D'un

autre coté, l'identité ontologique a une structure conceptuelle implicite, qui peut être rendue explicite au moins partiellement. La structure conceptuelle explicite (*makes explicit*) l'identité ontologique et dans ce sens l'explique. En même temps elle remplit le FE entre l'identité formelle (ou la règle) et l'identité ontologique correspondante et le FE entre deux parties de l'identité formelle. Dans ce dernier sens l'identité formelle elle aussi peut être expliquée (peut être dérivée *a priori*). Donc il y a deux sens suivant lesquels l'identité peut être expliquée et il y a deux sens suivant lesquels l'identité ne peut pas être expliquée. Nous essayons d'éviter une confusion entre le niveau épistémique et le niveau ontologique, et tenir compte de la structure conceptuelle implicite de l'identité ontologique. (Ch. 8)

Nous sommes en partie d'accord avec Hohwy (2003) et McLaughlin (2003) sur le fait que l'usage que fait Block du principe de l'inférence à la meilleure explication (PIME) dans le traitement du *problème plus difficile* (Block 2002) n'est pas correct. Block ne prend pas en compte le fait que le PIME pourrait en partie constituer le phénomène qu'il vise à expliquer. De notre point de vue, Block considère des phénomènes qui sont déjà entièrement constitués et il cherche leurs meilleures explications. Selon lui, les identités théoriques, comme par exemple l'identité "L'eau est l'H₂O", font partie des meilleures explications des phénomènes et c'est pour cela qu'on les accepte. A notre avis, dans un sens, la situation est exactement inverse. Les identités sont déjà implicites dans des phénomènes constitués; ce sont des règles implicites. Les meilleures explications sont des *explicitations* de ces règles. Block justement ignore le stade d'explicitation. Pour lui les identités sont déjà explicites, et les meilleures explications sont extérieures par rapport aux phénomènes. Notre position est que l'identification « L'eau = l'H₂O » contribue à la meilleure explication parce que l'eau est l'H₂O. Dire que l'eau est l'H₂O parce que l'identification « L'eau = l'H₂O » contribue à la meilleure explication c'est confondre le niveau ontologique et celui descriptif. C'est mettre la charrue avant les bœufs. (On trouve le même genre d'erreur chez Braddon-Mitchell et Jackson (1996/2001).) (Ch. 8)

Nous supposons que l'application de la règle peut être plus ou moins réflexive. (On peut également parler de la règle instinctive et de la règle plus ou moins réflexive.) L'identité "L'eau est l'H₂O" peut être l'une de trois choses suivantes: un énoncé réflexif avec la règle explicite "eau" (si par exemple l'identité est une pratique de la découverte que l'eau est l'H₂O), un énoncé réflexif avec la règle explicite "H₂O" (si par exemple l'identité est une dérivation *a priori* du fait que l'H₂O est l'eau), une application réflexive ou instinctive de la règle "L'eau = l'H₂O". L'identité instinctive "L'eau est l'H₂O" ne contient aucune règle explicite (ni "eau", ni "H₂O", ni "L'eau = l'H₂O"). Un autre cas limite est celui de l'identité purement réflexive: l'identité coïncide avec sa règle formelle. Dans les cas intermédiaires, l'identité partiellement réflexive "L'eau est l'H₂O" avec la règle "L'eau = l'H₂O" peut être vue comme une corrélation (plus ou moins forte)

naturelle/spontanée entre les termes "eau" et "H₂O". Cette corrélation (l'identité réflexive) "survient" sur l'identité instinctive correspondante. C'est en ce sens que les identités expliquent les corrélations. (Chapitres 2, 3, 8)

Nous interprétons la relation naturelle/spontanée entre les énoncés réflexifs/introspectifs et les énoncés instinctifs, que nous appelons la *survenance naturelle/spontanée*, comme une sorte de naturalisation de la notion modale de la survenance de Kim. (1998) Nous comparons notre notion avec celle de la survenance humienne de Lewis et celle de la survenance naturelle (nomologique) de Chalmers (1996). A la différence de la notion purement physicaliste de Lewis et de la notion dualiste de Chalmers, notre notion de la survenance est à la fois naturelle et normative (spontanée). Kim lui même a constaté la faiblesse explicative et métaphysique de la notion de la survenance (modale). (1993) La survenance modale entre les propriétés mentales M et les propriétés physiques sous-jacentes P est une simple corrélation formelle, elle n'implique aucune relation de dépendance ni de détermination entre M et P. Notre notion de la survenance naturelle/spontanée remédie à ces défauts. Elle est notamment plus appropriée pour traiter le problème de la causalité mentale. (§ 10 du ch. 1, § 2.9.3 du ch. 3, § 5 du ch. 5.)

Décembre 2006, Paris