



HAL
open science

Une psychiatrie sans esprit ?

Hugo Bottemanne, A. Chevance, Lucie Joly

► **To cite this version:**

Hugo Bottemanne, A. Chevance, Lucie Joly. Une psychiatrie sans esprit ?. *L'Encéphale*, 2021, 47 (6), pp.605-612. 10.1016/j.encep.2021.05.006 . halshs-03565374

HAL Id: halshs-03565374

<https://shs.hal.science/halshs-03565374>

Submitted on 5 Jan 2024

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.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Une psychiatrie sans esprit ?

Psychiatry without mind?

Hugo Bottemanne^{1,2,3}, Astrid Chevance⁴, Lucie Joly⁵

¹ Paris Brain Institute - Institut du Cerveau (ICM), UMR 7225 / UMRS 1127, Sorbonne University / CNRS / INSERM, 75013 Paris, France

² Sorbonne University, Department of Philosophy, Science Norm Decision Research Unit, UMR 8011 / FRE 3593, Paris, France

³ Sorbonne University, Department of Psychiatry, Pitié-Salpêtrière Hospital, Assistance Publique-Hôpitaux de Paris (AP-HP), 75013 Paris, France

⁴ Centre of Research in Epidemiology and Statistics Sorbonne Paris Cité, Institute for Health and Medical Research, and French National Institute of Research for Agriculture, University of Paris, 75013 Paris, France

⁵ Sorbonne University, Department of Psychiatry, Saint Antoine Hospital, Assistance Publique-Hôpitaux de Paris (AP-HP), 75012 Paris, France

Corresponding author: Hugo Bottemanne

e-mail : hugo.bottemanne@gmail.com

Résumé

La philosophie de l'esprit (Philosophy of Mind) représente aujourd'hui l'un des champs de recherche les plus prolifiques en philosophie et connaît une hybridation progressive avec les sciences cognitives. Elle s'attache à étudier des questions fondamentales pour les neurosciences et la psychiatrie, comme la nature des états mentaux et des processus cognitifs le rapport entre le cerveau et l'esprit, ou encore la relation des états mentaux avec le monde. La philosophie de l'esprit a posé les fondements de l'espace conceptuel au sein duquel les sciences cognitives et la psychiatrie ont essaimé, avec une trame résolument naturaliste. Pourtant, la psychiatrie contemporaine se montre souvent réticente à se nourrir de la philosophie de l'esprit, si ce n'est par le biais des fragments qui affleurent des neurosciences. Dans cet article, nous décrivons brièvement l'évolution de la philosophie de l'esprit contemporaine, et nous détaillerons ses contributions pour la réflexion autour de trois thématiques centrales pour la psychiatrie : la naturalisation de l'esprit, la causalité mentale, et la subjectivité des états mentaux. Nous verrons comment la philosophie de l'esprit peut être utilisée pour appréhender la clinique et les théories en psychiatrie : du biologique au fonctionnel, de la neurophysiologie à la cognition, de la matière à l'esprit.

Mots-clés : Philosophie; Psychiatrie; Philosophie de l'esprit; Epistémologie; Croyance; Neurosciences; Bayésianisme

Abstract

The Philosophy of Mind is currently one of the most prolific fields of research in philosophy and has witnessed a progressive hybridization with cognitive science. It focuses on fundamental questions to neuroscience and psychiatry, such as the nature of mental states and cognitive processes, the relationships between brain and mind, and between mental states and the world. The Philosophy of Mind laid the foundations for the conceptual space within which cognitive science and psychiatry have sprung up, with a resolutely naturalistic framework. Yet contemporary psychiatry is often reluctant to feed on the Philosophy of Mind, other than through the fragments that emerge from neuroscience. In this article, we will briefly describe the evolution of contemporary Philosophy of Mind, and we will detail its contributions to thought around three central themes for psychiatry: naturalization of mind, mental causality, and subjectivity. We will see how Philosophy of Mind can be used to understand the clinic and theories in psychiatry: from biological to functional, from neurophysiology to cognition, from matter to mind.

Keywords: Philosophy; Psychiatry; Philosophy of Mind; Epistemology; Belief; Neuroscience; Bayesianism

Introduction

La philosophie de l'esprit (*Philosophy of Mind*) représente aujourd'hui l'un des champs de recherche les plus prolifiques en philosophie. Son essor au cours des cinquante dernières années a été marqué par son expansion académique, mais aussi par son hybridation progressive avec les sciences cognitives. La philosophie de l'esprit a aujourd'hui un objet d'étude qui lui est propre : il s'agit spécifiquement de la discipline qui étudie les *états mentaux*, et les *phénomènes mentaux* (1). Les phénomènes mentaux désignent généralement l'ensemble des processus cognitifs, et les états mentaux leurs produits, parmi lesquels on compte par exemple les croyances, les émotions, ou encore les perceptions. La philosophie de l'esprit s'attache dès lors à explorer la nature des états mentaux et des processus cognitifs, les interactions causales entre l'esprit et la matière, ou encore la relation des états mentaux avec le monde (1,2). Anticipant l'accumulation des données expérimentales, parfois

contradictoires, elle offre une trame pour la génération de théories, tant et si bien qu'une grande partie des théories contemporaines en sciences cognitives proviennent d'une hybridation de conceptions forgées par des philosophes de l'esprit et de données produites par des neuroscientifiques (3,4).

En France, la réflexion philosophique en psychiatrie s'est principalement constituée autour de la dimension phénoménologique, c'est-à-dire la méthode d'exploration des états mentaux et du réel tel qu'ils apparaissent *subjectivement* à la conscience. Cette phénoménologie initialement plutôt psychanalytique a participé à la réflexion autour des troubles mentaux, avec un positionnement large allant de l'éthique à la métaphysique (6,7). Mais cette méthodologie subjective s'est retrouvée en difficulté lors de la confrontation avec les sciences empiriques, et sous le feu des logiciens de la philosophie analytique (8). Cette réflexion a été associée à la perpétuation de conception dualiste corps-esprit en psychiatrie, et plus largement au sein de la communauté médicale dans son ensemble. Le dualisme consiste à considérer qu'il existe une différence de nature, de propriété, ou de causalité, entre les phénomènes mentaux (comme les perceptions, les croyances, les émotions) et les phénomènes physiques (comme l'activité neuronale). En particulier, la distinction abusive entre organicité et psychisme, ferment de la séparation entre les spécialités médicales dites somatiques et la psychiatrie, constitue encore une barrière conceptuelle forte aux échanges entre ces différentes disciplines.

Etonnamment, la psychiatrie contemporaine se montre souvent réticente à se nourrir de la philosophie de l'esprit, si ce n'est par le biais des fragments qui affleurent des neurosciences. Malgré son importance dans la génération actuelle des théories étio-pathogéniques en psychiatrie (9), la philosophie de l'esprit n'est pas proposée dans le cursus de formation des psychiatres, et souvent absente des débats contemporains francophones. Pourtant, la philosophie de l'esprit a posé les fondements de l'espace conceptuel où ont essaimé les sciences cognitives, et le succès empirique des sciences cognitives a favorisé l'évolution moniste et matérialiste de la philosophie de l'esprit (5). Cette hybridation peut probablement jouer un rôle non négligeable pour la psychiatrie, en lui fournissant un socle ontologique avec une trame résolument naturaliste.

Dans cet article, nous décrirons brièvement l'évolution de la philosophie de l'esprit contemporaine, et nous détaillerons ses contributions pour la réflexion autour de trois thématiques centrales pour la psychiatrie : la naturalisation de l'esprit, la causalité mentale, et la subjectivité des états mentaux. Nous verrons comment ces

thématiques peuvent constituer des défis pour la psychiatrie, dans ses dimensions tant théoriques que cliniques, et de quelle manière elle pourrait bénéficier de ces apports de la philosophie de l'esprit.

Pour un ? sci?nc? d? l'?sprit

L'éclosion de la philosophie de l'esprit contemporaine a été favorisée par un certain nombre de facteurs synergiques, indissociables de l'essor des sciences cognitives (10). Elle s'est ainsi nourrie des découvertes empiriques provenant de disciplines multiples comme la psychologie expérimentale, les neurosciences, l'anthropologie, la linguistique, l'intelligence artificielle ou encore les sciences computationnelles, tout en participant largement aux fondements conceptuels de ces disciplines (2). Pour le philosophe Daniel Andler, la philosophie de l'esprit « prépare le terrain pour les sciences cognitives, (...) principalement en la débarrassant de divers obstacles » (11). Ces obstacles multiples, comme la théorie du dualisme corps-esprit, ont dominé les réflexions philosophiques en Europe pendant de nombreux siècles (12). Ces obstacles ont été fissurés progressivement par les travaux de philosophes comme William James, Charles Sanders Peirce, ou Clarence Lewis, puis fracturés par l'apogée de la philosophie de l'esprit anglo-saxonne au milieu du vingtième siècle avec des auteurs comme Nelson Goodman, John Austin, ou encore David Armstrong (13). La publication de *The Concept of Mind* en 1949 par le philosophe Gilbert Ryle ouvre par exemple un espace conceptuel pour les sciences empiriques en réfutant vigoureusement le dualisme de type cartésien et son « dogme du fantôme dans la machine » (14). Ces fractures marquent une première brèche dans la relative indépendance de la réflexion philosophique vis-à-vis des progrès scientifiques contemporains, et posent les bases d'une hybridation des deux disciplines.

L'exploitation de l'héritage de la philosophie du langage et du cadre formel hérité de la logique propositionnelle marque aussi une nouvelle étape pour cette discipline (15). Cette évolution profite de la naturalisation de la logique opérée par le philosophe Willard Van Orman Quine (16), c'est-à-dire de l'inscription des lois de la logique dans le monde physique. Cette naturalisation traduit les phénomènes mentaux, sous leur forme logique, en phénomènes naturels, en utilisant les prémisses de la psychologie scientifique contemporaine. Si les phénomènes mentaux constituent une classe particulière de phénomènes naturels, et reposent sur des règles formelles, alors la philosophie de l'esprit peut se proposer de faire la

synthèse entre la logique et les sciences de la nature : une grande partie des travaux en philosophie de l'esprit s'attaque dès lors à la formulation de l'esprit en termes de structures logiques, et cette évolution fournit les bases pour l'étude du langage et de la pensée sous la forme d'une manipulation de symboles ou de *représentations* (2). La philosophie n'est plus seulement la réflexion métaphysique autour des concepts mentaux et du paradoxe dualiste, mais une exploration des mécanismes formels de la cognition (17).

Enfin, la troisième étape est marquée par l'essor des sciences computationnelles à la suite des propositions du logicien Alan Turing, et la focalisation sur les processus de traitement de l'information plutôt que sur leurs états représentationnels (18). Il s'agit dès lors pour les sciences de l'esprit de caractériser, d'expliquer, et de modéliser les états mentaux et leurs processus logiques : pour cela, elles doivent étudier les rôles fonctionnels qu'ils occupent dans les relations entre les entrées sensorielles, les états représentationnels, et les sorties comportementales d'un organisme (19). Ces nouvelles approches suggèrent que la pensée n'est pas un enchaînement de représentations statiques, mais un cheminement inférentiel qui se nourrit de l'expérience du réel, et dont les règles peuvent être définies de manière formelle sous la forme d'une computation. Ces règles formelles, aussi appelées *algorithmes*, définissent la manière dont les représentations sont générées et modifiées au cours du temps. Dans ce cadre, la théorie du bayésianisme subjectif, formalisée par le mathématicien Frank Ramsey en 1926 (20), s'est progressivement imposée comme la trame théorique majeure pour représenter et modéliser les processus logiques qui gouvernent la pensée, menant à l'hypothèse contemporaine du *cerveau bayésien* (21). Elle suppose notamment que les états représentationnels sont équivalents à des degrés de probabilité, et que ces états représentationnels peuvent être modifiés en utilisant des algorithmes bayésiens. Cette logique bayésienne a marqué de son empreinte l'évolution théorique des sciences cognitives, participant pleinement à une grande partie des théories contemporaines en psychiatrie (22).

Mais ces nouvelles approches conceptuelles à la frontière des sciences cognitives et de la philosophie de l'esprit génèrent de nombreuses questions tant pour les sciences de l'esprit que pour leurs applications cliniques en psychiatrie. En particulier, la question de la causalité du mental au physique, ou encore de la nature qualitative des états mentaux, compliquent la dimension expérimentale et théorique des sciences de l'esprit. Par ailleurs, il persiste un fossé entre la participation de la philosophie de l'esprit pour la génération

des théories en psychiatrie, et son utilisation par les psychiatres dans le cadre de leur pratique clinique. Dans la suite de cet article, nous détaillerons comment ces fondements conceptuels issus de la philosophie de l'esprit questionnent les théories et la pratique clinique en psychiatrie, et nous proposerons plusieurs manières d'y répondre.

L'esprit dans un monde physique

La question de la place de l'esprit, et des *états mentaux* en général, dans un monde régi par les lois de la physique est au cœur de la philosophie de l'esprit (23). Le matérialisme, au sens prototypique, est une position métaphysique affirmant que le monde est entièrement constitué de matière, et donc dépendant de la physique (24). L'un des arguments centraux de cette position est la *clôture causale* (ou complétude) du monde physique (25) : les causes physiques suffisent pleinement à expliquer les effets physiques, autrement dit, tout ce qui est causé est causé par le physique, il n'existe pas de cause non physique, ou encore tout ce qui n'est pas physique ne peut pas avoir de pouvoir causal (24). Cette position métaphysique est l'un des fondements de la pratique scientifique, en tant que méthode d'investigation des lois causales du monde physique sous toutes ces formes (biologie, biochimie, physique nucléaire, génétique, immunologie) : si la causalité est clôturée dans le physique, alors la méthodologie expérimentale peut reposer sur le physique (26). Cette position associe ainsi une dimension ontologique (i.e. chaque propriété du monde est physiquement constituée), et méthodologique (i.e. chaque propriété du monde peut être étudiée à l'aide des sciences physiques) (24).

En philosophie de l'esprit, le physicalisme représente l'extension des conceptions matérialistes aux relations complexes entre le corps et l'esprit (27). Si le monde est uniquement constitué de matière, et que la matière est la seule chose qui soit dotée d'un pouvoir causal, alors chaque chose (dont l'esprit) est physique, provient du physique, ou survient sur le physique, selon le niveau d'implication de la thèse. Le premier niveau d'implication, le plus fort sur le plan métaphysique, est la thèse de l'identité entre le corps et l'esprit. Formellement introduite par les philosophes Herbert Feigl, John Smart et Ullin Place sous la forme du *monisme matérialiste*, cette thèse suggère que les événements mentaux sont, à titre exclusif, des événements physiques, et rien de plus (28–30). Cette thèse offre la possibilité de résoudre un certain nombre de questionnements sur les rapports entre le corps et l'esprit, ou encore la causalité mentale au domaine physique,

car si les propriétés mentales sont identiques à des propriétés neurobiologiques, alors il n'y a pas de problème de relation entre le monde physique et mental, et toute causalité mentale n'est plus qu'une espèce de causalité physique (1). Profitant de l'accumulation des données expérimentales, de nombreux courants contemporains en philosophie de l'esprit ont essaimé autour de la position *matérialiste* (tout ce qui existe est physique) et *moniste* (l'esprit survient sur le physique) (24), participant à fonder une véritable trame théorique transversale allant de la psychologie cognitive aux neurosciences moléculaires (31,32). Cette trame a participé à l'émergence d'une psychiatrie *naturaliste*, c'est-à-dire fondée sur l'idée que les symptômes psychiatriques reposent sur des états physiques, ou biologiques, et sur laquelle veut se fonder une partie de la psychiatrie scientifique contemporaine. Cette naturalisation des troubles psychiatriques s'est appuyée sur le développement des descriptions neurobiologiques issues des neurosciences, tout en conservant les descriptions psychologiques traditionnelles, notamment les états mentaux comme la croyance, l'émotion, et l'humeur.

Ce rapport entre les concepts issus de la philosophie de l'esprit, les états mentaux traditionnels de la psychologie, et les produits empiriques des sciences cognitives est pourtant ambigu. Si les symptômes psychiatriques ne sont que l'expression de dysfonctionnements physiques, et que les descriptions neurobiologiques subsument les descriptions psychologiques, pourquoi conserver les états mentaux traditionnels en psychiatrie clinique ? Si le monde causal est clôturé dans le monde physique, que tout ce qui est mental survient sur le physique, et que la science est la méthode d'investigation principale du monde physique, alors pourquoi se préoccuper de philosophie ? C'est pourtant ici que la philosophie de l'esprit prend tout son sens. On peut dire qu'elle a en grande partie traité le problème du réductionnisme neurobiologique, et participé à libérer la psychiatrie des paradoxes du matérialisme fort. Elle a critiqué en leur cœur les prétentions initiales de la neuropsychologie et des hypothèses issues des premières techniques d'imagerie cérébrale. Contre la réduction initiale des neurosciences à une tâche de localisation cérébrale des facultés cognitives, elle a offert des arguments formels invitant à remettre de l'esprit au cœur des sciences cognitives, et indirectement au cœur de la psychiatrie.

La position physicaliste que nous avons évoquée a ainsi été l'objet d'un débat complexe mené par les philosophes de l'esprit, majoritairement focalisé autour de l'argument de la *réalisation multiple* et la thèse de

l'anomisme du mental. Le premier argument de *réalisation multiple*, avancé par le philosophe Hilary Putnam, suggère que différents systèmes physiques, quelle que soit la matière qui les composent, pourraient théoriquement donner naissance au même état mental, au sens où ce ne sont pas les propriétés physiques qui déterminent les propriétés mentales, mais les propriétés fonctionnelles, ou relationnelles (33). En particulier, l'évolution de la science computationnelle a montré que des processus de traitement de l'information isomorphes à des phénomènes mentaux comme le raisonnement sous incertitude ou la prise de décision pouvaient être implémentés dans des machines, avec des réseaux en silicium, c'est-à-dire que le substrat biologique n'était pas nécessaire à la réalisation de certains états fonctionnels, ou de certaines propriétés computationnelles (34,35). Certaines opérations logiques peuvent être implémentées sur des réseaux artificiels, sans lien avec un substrat biologique, en reproduisant le profil relationnel (ou computationnel) des réseaux de neurones. Cette position a bénéficié d'arguments formels ou de résultats expérimentaux (36,37), comme la réalisabilité multiple des prédictions descendantes dans le cadre du codage prédictif (38), la variation temporelle des bases neurales sous-tendant le traitement des émotions (39). En psychiatrie, ce débat est particulièrement déterminant pour les programmes de recherche qui explorent les biomarqueurs de la symptomatologie psychiatrique, c'est-à-dire les marqueurs biologiques (concentration d'une molécule dans le sang, activité d'une aire cérébrale) associés aux symptômes, avec l'objectif de proposer des thérapies ciblées sur ces dysfonctionnements. En particulier, les biomarqueurs fonctionnels acquis au cours de procédures d'imagerie cérébrale sont une cible privilégiée pour les thérapies contemporaines de neuromodulation (stimulation magnétique transcranienne, stimulation profonde), notamment dans le trouble obsessionnel compulsif ou les hallucinations résistantes au traitement pharmacologique. Les notions de réalisabilité multiple et de variation temporelle des aires fonctionnelles pourraient constituer une trame d'explication pour la résistance de certains patients aux thérapeutiques validées chez d'autres, malgré la similitude de leur symptomatologie psychiatrique. Ces hypothèses invitent à explorer plus ardemment les profils relationnels, et particulièrement *computationnels*, impliqués dans les troubles mentaux, notamment à l'aide de techniques d'imagerie bayésienne comme la Modélisation causale dynamique (DCM, Dynamic Causal Modeling, pour une revue, voir (40)).

Le second argument avancé contre le matérialisme fort, et proposé par le philosophe Donald Davidson, est *l'anomisme du mental*. Il suggère que les états mentaux *surviennent* sur les états physiques, mais sont

irréductibles aux propriétés physiques (41). En conséquence, le mental est *anomique* au physique, c'est-à-dire qu'il ne peut être formulé de lois strictement physiques qui subsume les propriétés du mental. Cette position exploite l'hypothèse formelle de la *survenance*, un concept issu de la philosophie de la biologie et de la physique permettant de déjouer les paradoxes associés à la notion d'identité entre propriétés de nature différente. On peut ainsi parler de propriétés survenantes lorsque l'on étudie des systèmes hiérarchiques avec de multiples niveaux de complexité pour lesquels on ne retrouve pas nécessairement de relation d'identité entre les différentes propriétés hiérarchiques, comme c'est le cas par exemple pour les relations causales entre plusieurs entités physiques à différentes échelles. On considère ainsi formellement qu'un état mental peut survenir sur un état physique lorsque deux états physiques semblables provoquent deux états mentaux identiques, mais que l'on ne trouve pas de relation d'identité entre ces états physiques et mentaux, ou entre les lois causales qui les caractérisent. Par exemple, le holisme des croyances, c'est-à-dire les contraintes logiques qui existent entre plusieurs croyances co-existantes, est l'une des facettes de l'anomisme du mental : bien qu'une croyance soit probablement survenante sur un état physique, son contenu est intrinsèquement dépendant de sa position au sein d'un réseau d'autres croyances, qui constituent l'ensemble du modèle du monde pour un sujet (42).

Les hypothèses les plus contemporaines en neurosciences théoriques reposent également sur ces théories anomistes et de réalisation multiple. Elles supposent majoritairement que des propriétés *représentationnelles* ou *computationnelles* puissent survenir sur des propriétés physiques, sans que ces propriétés fonctionnelles ne soient subsumés par les propriétés physiques sur lesquelles elles émergent. C'est par exemple le cas pour la théorie de l'espace de travail neuronal global du neuroscientifique Stanislas Dehaene (43), ou encore de l'inférence active de Karl Friston (44). Dans ces deux théories, les processus mentaux sont décrits de manière précise et formelle sur le plan logique, majoritairement en utilisant le cadre du bayésianisme, et associés à des profils fonctionnels sur le plan cérébral caractérisés par la notion de relation, ou de réseau. Cette hypothèse est majeure pour la psychiatrie, car elle suppose qu'un grand nombre des symptômes psychiatriques, par exemple les états mentaux qualifiés de pathologique (idée délirante, hallucination, cognition dépressive), *surviennent* sur des états physiques du cerveau mais ne sont pas réductibles à ces états physiques. Ces hypothèses issues de la psychiatrie computationnelle insistent encore une fois sur la prévalence des propriétés *relationnelles* et *mécanistiques* des phénomènes mentaux, parfois sans faire d'hypothèses précises à propos

des états physiques sur lesquels ils surviennent (9). La psychiatrie computationnelle a généré de riches descriptions de l'étiopathogénie des troubles psychiatriques comme la schizophrénie, la bipolarité, les addictions ou encore les troubles du spectre autistique, en se limitant aux mécanismes logiques impliqués dans ces pathologies (45,46). Ces descriptions ne permettent toutefois pas encore de véritables applications translationnelles, au sens d'une amélioration franche des thérapeutiques utilisées contre ces troubles. Nos traitements, particulièrement la pharmacothérapie et les thérapies de stimulation, ciblent les dysfonctionnements neurobiologiques impliqués dans les troubles. La connaissance précise des processus logiques qui gouvernent l'esprit ne permet malheureusement pas encore d'interventions thérapeutiques. Si l'esprit est survenant sur le physique, et gouverné par des processus logiques, le véritable défi repose encore sur la formalisation du lien entre ces processus logiques et la neurobiologie.

L' paradox de la causalité mentale

En sciences cognitives, et par extension en psychiatrie, la fonction fait office d'explication. Ces disciplines cherchent à expliquer fonctionnellement les comportements d'autrui en faisant appel à des états mentaux comme la croyance, la perception, le désir, l'émotion, l'intention, ou encore la sensation. Cette conception suggère que nos comportements sont causés par nos états mentaux, que nos comportements déterminent la nature de nos entrées sensorielles, et que nos entrées sensorielles déterminent en retour nos états mentaux (47). Ces états mentaux occupent alors un rôle causal dans la production des comportements, dépendant conjointement des caractéristiques de l'état mental considérée et de son contenu sémantique. Cette boucle causale de l'esprit au corps, puis du corps à l'esprit, permet de définir fonctionnellement les états mentaux par le rôle causal qu'ils occupent dans le système (48). Cette explication peut être réalisée à un niveau comportemental, au cours duquel l'observateur interprète un ensemble de signes, mais aussi au niveau neurologique en montrant la manière dont des processus neuraux réalisent ces fonctions causales (49).

Lorsque l'on étudie des relations causales entre plusieurs propriétés, il existe en réalité de nombreuses manières de les décrire (50). Les théories computationnelles issues de la philosophie de l'esprit, en particulier celles qui ont participé à la trame théorique contemporaine de la psychiatrie, utilisent majoritairement ce fonctionnalisme causal pour décrire les mécanismes impliqués dans les troubles (51). A partir de ce principe

unificateur de la causalité mentale, ou de la causalité du mental sur le physique, elles proposent de nombreux mécanismes putatifs sous-jacents aux comportements observés par le clinicien. Par exemple, un patient souffrant d'une addiction au *chemical-sex* (utilisation de produits dopants au cours de pratiques sexuelles) qui se rend dans une soirée libertine met en place une série de comportements en raison de croyances prospectives concernant le déroulement de la soirée, l'effet des produits administrés, l'estimation du plaisir ressenti. De la même façon, un patient suicidaire qui accroche une corde pour se pendre réalise des actions spécifiques qui modifient le monde physique en raison d'un ensemble d'états mentaux à propos des conséquences de son action sur le monde. Il peut croire que sa vie n'a plus aucun sens, que la corde sera suffisamment solide pour supporter son poids, et avoir le désir que la corde comprime sa trachée. Ces états mentaux, à ce niveau de description, occupent un rôle fonctionnel causal suffisant pour expliquer, ou décrire son comportement (52). De la corde nouée à la mort de notre patient, il y a ici une *causalité* du mental au physique, un lien entre la croyance comme état mental et les conséquences de l'action comme objet physique. Malgré son élégance, cette heuristique fonctionnelle soulève un certain nombre de paradoxes focalisés autour de la question de l'interprétation, et celle de la causalité. Imaginons par exemple une star de cinéma à un café de Saint-Tropez, par une chaude après-midi d'été. Nous la voyons mettre soudainement une large paire de lunettes de soleil noire. Nous pouvons interpréter son comportement, l'expliquer de manière fonctionnelle, en utilisant des états mentaux : elle suspecte qu'un journaliste est à proximité (croyance), et ne souhaite pas être photographié (désir), car elle pense qu'elle a les traits fatigués (croyance). Peut-être aussi qu'elle est éblouie par le soleil (perception), ce qui est pour elle désagréable (émotion), et qu'elle craint d'abîmer sa rétine (croyance). Peut-être encore qu'elle est triste (émotion), qu'elle essaye de se retenir de pleurer (action), parce qu'elle ne souhaite pas qu'on voit de larmes sur son visage (désir). Pour un même comportement, un même interprète, il existe une multitude d'états mentaux possibles, et probablement plusieurs voies neuronales potentiellement impliquées. Comment le psychiatre peut-il échapper à ce piège de l'interprétation causale, au vu des limites physiques de sa clinique ?

Le second paradoxe repose sur la question de la causalité du mental. Nous avons vu précédemment que la trame théorique en sciences cognitives, et par extension en psychiatrie, admettait que le monde physique était clos. Le monde est uniquement constitué de matière, l'esprit est physique, et la matière est la seule chose qui

soit dotée d'un pouvoir causal. C'est ici que le paradoxe de la causalité pose problème. Revenons à notre patient suicidaire qui réalise des actions pour se pendre en raison d'un ensemble d'états mentaux (croyances, émotions, désirs) à propos des conséquences de ces actions. Si le domaine physique est clos, et que l'ensemble des événements physiques qui se succèdent jusqu'à la mort de notre patient peut recevoir une explication causale limitée au domaine physique, alors quel besoin de faire appel à des états mentaux comme des croyances, des désirs, ou tout autre chose ? D'un autre côté, le cœur de la clinique psychiatrique est constitué par ces états mentaux, et que reste-il de notre discipline si nous les éliminons au profit d'états physiques ? Cette question centrale recoupe largement les débats autour des corrélats neuronaux des troubles psychiatriques, mais aussi de manière plus contemporaine le débat autour des biomarqueurs computationnels. A mesure que s'accumulent les données empiriques concernant le lien entre les propriétés physiques et les propriétés subjectives (intentionnelles et qualitatives), la causalité mentale semble dépossédée de sa force explicative.

En philosophie de l'esprit, cette question a été formulée au sein de la thèse *éliminativiste* affirmant que les états mentaux utilisés en psychologie du sens commun (Folk Psychology) seraient progressivement éliminés au fur et à mesure de l'évolution des neurosciences (53). Si toute chose est physique, et qu'il existe une méthodologie optimale pour expérimenter le monde physique, alors l'usage d'états mentaux comme les croyances ou les désirs disparaîtra en faveur des bases neurologiques qui les sous-tendent. Bien que l'on trouve déjà les prémisses de l'éliminativisme chez les philosophes Ludwig Wittgenstein ou Gilbert Ryle, ces hypothèses ont été formalisées initialement en psychologie dans le béhaviorisme logique : pour les behavioristes, l'esprit et les états mentaux qui le caractérisent peuvent toujours être réduit au comportement, en dernière instance (14,54). C'est d'ailleurs l'un des risques encourus par le monisme anomal de Donald Davidson évoqué précédemment. Si les phénomènes mentaux sont causés par des phénomènes physiques, et que la clôture causale du monde physique est maintenue, alors les phénomènes mentaux ne peuvent quant à eux rien causer du tout. Cette hypothèse, qualifiée d'*épiphéénoménalisme*, condamne les états mentaux à une inutilité causale. Dès lors, la causalité mentale est une illusion, et la focalisation de la pratique psychiatrique sur les états mentaux une erreur.

Un premier amendement proposé par plusieurs philosophes de l'esprit, déterminant pour la clinique en psychiatrie, a consisté à admettre que les attitudes propositionnelles comme la croyance et le désir étaient constituées à la fois de propriétés épistémiques et physiques. En effectuant cette distinction, ils supposaient que leurs propriétés épistémiques étaient dérivées du fait qu'elles se comportent comme des propositions avec un contenu et un sens sémantiquement évaluable, mais dénués de tout pouvoir causal dans le monde physique. A l'inverse, cette thèse assumait que leurs propriétés physiques garantissent une force causale, tout en étant dénuées de toute propriété sémantique. Cette distinction entre des propriétés physiques avec un pouvoir causal, et des propriétés épistémiques avec un contenu sémantique, ne constitue pas un renouveau du dualisme cartésien des substances, bien au contraire. Elle est solidement ancrée dans le monisme matérialiste que nous avons détaillé, en attribuant tout le pouvoir causal aux propriétés physiques, et en supposant par extension que les propriétés sémantiques en sont dénuées. Ainsi, cette position propose une voie intermédiaire entre le dualisme et le réductionnisme matérialiste, fondée sur la clôture de la causalité. Mais que penser alors de la prétention des théories computationnelles utilisées en psychiatrie, en particulier celles issues du bayésianisme subjectif, de faire des propriétés sémantiques et relationnelles le cœur de l'explication causale ?

Un second amendement voisin proposé par plusieurs philosophes de l'esprit au milieu du XX^e siècle a consisté à considérer que l'objectif des sciences de l'esprit reposait sur la notion d'*interprétation*, au sens de l'ordre des raisons et non des causes, tandis que les sciences biologiques relevaient de l'explication causale des faits naturels. Cet amendement suppose que les états mentaux permettent en première instance d'interpréter les comportements, c'est-à-dire de déduire indirectement la structure sémantique de l'esprit. Cette thèse de l'interprétation a connu un certain succès auprès des logiciens et des philosophes du langage, en grande partie en raison du renouveau d'intérêt qu'elle provoquait pour les états mentaux traditionnels. En effet, tout en s'opposant à la prétention de la psychologie de détenir la clé de l'explication causale des comportements humains, elle sapait le terrain des éliminativistes qui souhaitaient supprimer toute référence aux états mentaux. Elle réaffirmait notamment le rôle de la croyance comme état mental princeps, défendant sa place dans la trame théorique et empirique des sciences cognitives. Dans cette trame, on peut d'ailleurs imaginer un rôle intermédiaire pour le psychiatre, à l'interface de la dimension causale, dévolue aux neurosciences cognitives, et de la dimension sémantique, l'objet de la clinique. Cette interface pourrait être au cœur de la spécificité de

sa discipline, lui permettant de jongler entre ces deux dimensions, notamment lorsqu'il combine des outils pharmacologiques et psychothérapeutiques.

Cette position interprétative est toutefois elle-même sujette à de nombreuses limites. Un grand nombre d'états mentaux peuvent exister sans jamais produire directement de comportements, sans jamais produire de substrat nécessaire pour l'interprétation causale. Par exemple, nous pouvons avoir la croyance qu'il existe des épaves pleines de trésors dans l'océan Indien, ou une civilisation cachée dans la forêt d'Amazonie, sans qu'à aucun moment nos croyances n'engendrent le moindre comportement, ou *disposition à agir* dans le monde. Ce n'est pas parce que j'ai ces croyances que je vais me lancer dans une exploration des fonds marins, ou quitter mon foyer pour explorer l'Amazonie. Se pose aussi le problème de l'accessibilité des états mentaux pour le psychiatre. Les états propositionnels comme la croyance ou le désir sont souvent fournis verbalement par le patient au cours de l'entretien, ou interprétés subjectivement par le médecin au cours de l'observation clinique. En interprétant le contenu allusif du discours du patient, ou en inférant des hypothèses à partir de son attitude réticente, le psychiatre peut par exemple diagnostiquer un état délirant chez son patient. De la même manière, l'évocation par le patient de croyances avec un contenu de dépréciation associée à une volonté de repli au domicile peut amener son psychiatre à diagnostiquer un épisode dépressif. Pourtant, lorsque le patient expose les états mentaux qui ont mené à un comportement, il n'a accès qu'au reflet de son expérience consciente, et non aux mécanismes qui en sont la cause. En contrepoint, la meilleure caractérisation en termes neurobiologiques des symptômes psychiatriques n'offre qu'un aperçu limité de leur *rôle fonctionnel* pour les patients. Par exemple, la formalisation de la douleur morale ressentie par le sujet sous l'angle des processus neuronaux impliqués s'éloigne de l'expérience subjective de ce symptôme, au risque d'en perdre le sens. La philosophie de l'esprit, malgré sa défense des propriétés fonctionnelles et causales des états mentaux, ne propose pas encore de solutions véritablement satisfaisantes pour ce paradoxe complexe de la causalité mentale.

L'épreuve de la subjectivité

Pour les éliminativistes, l'évolution des neurosciences cognitives entrainera progressivement une élimination des états mentaux traditionnels utilisés en sciences cognitives. Pourtant, toute une classe de propriétés

mentales, primordiale pour la psychiatrie, résiste à l'éliminativisme. De nombreuses caractéristiques intrinsèques des états mentaux, comme leur intentionnalité (i.e. le fait de renvoyer à quelque chose) ou leur qualité subjective (i.e. le *quale*, la sensation que cela fait) sont en effet difficilement retranscrits en termes d'*états physiques* dans le cerveau (24). Je peux expliquer de manière physicaliste les réseaux cérébraux qui s'activent lorsque je vois une pomme rouge, mais je suis plus ennuyé lorsqu'il s'agit de décrire de la même manière *ce que cela me fait* de voir une pomme rouge. Pourtant l'intentionnalité de l'expérience phénoménale, mais aussi son contenu qualitatif, constituent deux facettes majeures de la subjectivation du monde par un individu. Malgré leur implication singulière dans de nombreux troubles psychiatriques, les neurosciences cognitives peinent encore à explorer, et formaliser en termes physiques, ces propriétés (27).

Dérivé du latin *intendere* (i.e. se diriger vers quelque chose), l'intentionnalité désigne le fait pour un état mental d'être dirigé vers un objet, de pouvoir renvoyer à quelque chose. De nombreux états mentaux traditionnels comme les croyances, les désirs, les jugements, ou les sentiments, sont caractérisés par leur relation à un contenu sémantique, et considérés comme des états intentionnels. On les qualifie fréquemment d'*attitude propositionnelle*, au sens où ils sont constitués d'une relation avec un contenu propositionnel, existant dans l'esprit du sujet ou dans le monde auquel il appartient. Le contenu de l'attitude propositionnelle est son objet, ce vers quoi elle est dirigée, et constitue la marque de sa nature intentionnelle. Je peux ainsi percevoir une étoile filante, croire que les caïmans sont végétariens, désirer vivre éternellement, ou avoir l'intention d'écrire une symphonie : mon attitude à propos de ces propositions définit ma relation subjective avec son contenu. Ce contenu peut aussi être évaluable, c'est-à-dire être inscrit dans une axiomatique formelle : ma croyance peut être vraie ou fausse, ma perception correcte ou inexacte, mon désir exaucé ou inassouvi, mon intention exécutée ou abandonnée.

Par ailleurs, les états mentaux conscients, qu'ils soient considérés comme des états intentionnels ou non, s'accompagnent quasiment systématiquement d'une expérience qualitative (55). Nos perceptions, sentiments, ou jugements sont associés à un effet subjectif, désigné comme le *quale*, ou *l'effet que cela fait* (56). Cela me fait un certain effet de sentir l'odeur de la fleur d'oranger, de me piquer avec une aiguille, de penser à l'être aimé, ou encore d'entendre le hurlement d'un loup dans la forêt. Le *quale* est parfois assimilé à l'émotion, mais il implique une dimension phénoménologique qui va bien au-delà des dimensions qualitatives de valence

et quantitative d'intensité. La phénoménologie de ces expériences a souvent des caractéristiques communes entre différents individus, tout en étant propre à un sujet particulier (56). Elles ne sont d'ailleurs pas indépendantes entre elles, mais liées par des relations logiques dont on peine à expliquer la nature : l'effet que cela me fait d'entendre le hurlement d'un loup est différent de celui d'entendre le brame d'un cerf ou le geignement d'un nourrisson, mais le quale associé au hurlement du loup me paraît plus proche du brame du cerf que du geignement du nourrisson. De la même manière, l'effet que cela me fait de voir le rouge vif est différent de celui de voir du noir ou de l'orange, mais l'expérience du rouge me paraît plus proche de l'orange que du noir.

La majorité des états mentaux sont caractérisés par ces propriétés intentionnelles et qualitatives. Ils ont une dimension objective au sens où un individu peut rapporter verbalement quel type d'état mental il possède, et qu'un autre peut comprendre qualitativement la nature de son expérience. Mais cette objectivité se limite à la similitude, ou l'isomorphisme, de la manière dont les deux individus traitent l'information. Nous pouvons comprendre quel type d'expérience subjective ressent un être humain qui entend une symphonie d'Antonín Dvořák, mais ne pouvons pas vraiment comprendre l'effet que cela fait d'être une chauve-souris qui perçoit son environnement par écholocalisation (56). Ces états mentaux sont ainsi subjectifs au sens où l'objectivisation de l'expérience n'est envisageable que lorsque le sujet qui observe est suffisamment semblable à celui qui est observé pour être en mesure d'adopter son point de vue. Ces effets ont été illustrés par l'expérience de pensée du *spectre inversé* : elle se propose d'imaginer que chez un sujet la couleur *rouge* apparaisse phénoménologiquement *verte*, mais qu'il continue à utiliser le mot *rouge* lorsqu'il désigne des objets qu'il perçoit comme vert, si bien que son inversion du spectre des couleurs passe totalement inaperçue aussi bien pour lui que pour ses pairs. Cet individu suivra les mêmes règles de circulation que tout le monde lorsqu'il sera au volant de sa voiture, désignera de la même façon les cerises qu'il vient de cueillir dans son jardin, c'est-à-dire que son action dans le monde sera *fonctionnellement* identique, bien qu'il éprouve des qualia différents du reste de la population. Dans notre clinique psychiatrique, ce caractère subjectif pourrait d'ailleurs limiter notre capacité à comprendre, au sens large, la nature des états mentaux de nos patients. Prenons la douleur morale retrouvée au cours d'un épisode mélancolique : lorsque nous écoutons le patient détailler ses états mentaux, nous ressentons qualitativement sa douleur, et nous pouvons l'objectiver en l'assimilant à des maux que nous avons déjà ressentis. Sa douleur fait écho en nous à d'autres états mentaux

précédents, marqués par leurs propriétés subjectives. Mais à moins d'avoir soi-même expérimenté un épisode mélancolique, pouvons-nous vraiment comprendre *l'effet que cela fait* d'être mélancolique, ou la douleur des patients a-t-elle pour nous la même incommensurabilité que l'écholocation de la chauve-souris ? Cet argument pourrait être répliqué pour de nombreux états mentaux, de l'apophénie retrouvée dans les prémices d'un épisode psychotique à la déréalisation de l'attaque de panique. Les patients en rémission d'un épisode dépressif d'intensité sévère décrivent d'ailleurs souvent le caractère étranger, inhabituel, des états mentaux qu'ils éprouvaient, bien loin de la notion classique de tristesse ou de douleur morale.

Cette dimension intentionnelle et qualitative des états mentaux a parfois aussi été utilisée comme un argument d'irréductibilité de l'esprit au physique (57). La difficulté des neurosciences à retranscrire en termes d'états physiques dans le cerveau l'intentionnalité et le qualia participerait à leur indépendance fonctionnelle, ou plus précisément à l'exigence du maintien de ces états mentaux pour l'explication fonctionnelle des comportements et de la cognition (12). Cette incommensurabilité est importante pour la psychiatrie, car la plupart des troubles mentaux impliquent des perturbations du caractère intentionnel (ce vers quoi est dirigée la pensée) et qualitatif (ce que cela fait d'avoir cet état mental). Dans le trouble anxieux par exemple, l'anxiété d'anticipation est caractérisée par des croyances avec un contenu intentionnel perturbé (à propos de la vraisemblance d'événements aversifs) et des *qualia* spécifiques (ce que cela fait de penser à ces événements aversifs). A l'inverse, les angoisses des patient psychotiques sont souvent marquées par leur absence de contenu intentionnel conscient (comme si elles étaient sans objet), et un effet qualitatif différent de l'anxiété du trouble anxieux. De la même manière, la bouffée anxieuse au réveil de certains patients mélancoliques est parfois caractérisée par le même objet intentionnel que l'anxiété d'anticipation, mais possède une dimension qualitative qui lui est spécifique. L'irréductibilité physique actuelle de ces états qualitatifs limite la possibilité pour la psychiatrie de s'appuyer sur les sciences de la nature pour les décrire. Cet argument est particulièrement important dans le cadre plus large du Research Domain Criteria (RDoC), un programme de recherche lancé en 2009 aux États-Unis par le National Institute of Mental Health (NIMH) ayant pour ambition de transformer la compréhension des troubles mentaux en partant non plus de la clinique psychiatrique mais de la physiopathologie génétique, moléculaire et cérébrale impliquée dans ces troubles (58). Le RDoC suppose que la compréhension d'une hiérarchie de mécanismes physiques (du gène à l'expression fonctionnelle des réseaux de neurones) permettra de comprendre comment les troubles mentaux surviennent, et d'améliorer les

propositions thérapeutiques (59). Ce programme ambitieux, qui a déjà souffert de nombreuses critiques, pourrait être limité par la difficulté pour les sciences empiriques d'appréhender ces propriétés spécifiques des états mentaux, survenant sur la matière mais irréductible à la matière (60,61).

Toutefois, la philosophie de l'esprit, en particulier dans sa dimension computationnelle, pourrait apporter des solutions contre le paradoxe de la subjectivité (62). En particulier, la défense de l'incommensurabilité de l'intentionnalité et du qualia se limite majoritairement à la dimension consciente des états mentaux. Les définitions contemporaines de la conscience en distinguent plusieurs types (conscience primaire, conscience phénoménale, conscience introspective), selon que l'on désigne les phénomènes perceptuels (sensoriels, proprioceptifs, nociceptifs), les qualités subjectives de l'expérience (les qualia, les sensations, les sentiments), ou encore les états mentaux de second degré comme la métacognition. Toutes ces définitions s'accordent toutefois sur un certain nombre de propriétés de la conscience, parmi lesquels l'intentionnalité (et à fortiori le qualia) sont en première place. La conscience est toujours conscience de quelque chose, en particulier pour la conscience primaire des phénomènes perceptuels. Mais les états mentaux complexes explorés en psychiatrie impliquent une dimension qui échappe largement à l'attribution consciente, et pourrait aussi être indépendant de l'exigence intentionnelle. L'humeur, dans sa dimension clinique, est typiquement un état dénué de caractère intentionnel (elle n'est pas dirigée vers un objet) mais caractérisée par des propriétés qualitatives (en lien avec la valence de l'humeur) (63).

Les modèles computationnels s'appuyant sur le bayésianisme, majoritairement hérité d'une longue tradition en philosophie de l'esprit, ont permis de formaliser ces états mentaux sous la forme de réseaux de croyances et de processus logiques, implémentés et traités par le cerveau. Ils associent les propriétés sémantiques, qualitatives, et intentionnelles de l'expérience avec sa dimension causale et physique, en proposant des mécanismes putatifs combinant processus logiques et processus biologiques. Plusieurs modèles computationnels contemporains des troubles de l'humeur la modélisent par exemple sous la forme d'un ensemble de croyances profondes, inconscientes, estimant la probabilité de récompense dans l'environnement (64), ou encore la probabilité de l'efficacité des actions mises en place par l'organisme pour acquérir ces récompenses (65). Ces modèles formalisent les propriétés intentionnelles (ce vers quoi sont dirigées les croyances) et qualitatives (ce que cela fait d'éprouver un type d'humeur) sous une forme hybride, à la fois

sémantique et biologique. Plus spécifiquement, ils admettent que l'organisation hiérarchique (des croyances entre elle, ou des réseaux) et la nature des processus logiques (notamment l'inférence bayésienne) sont implémentées dans la structure neuronale (au travers de la hiérarchie corticale), et que cette organisation détermine la nature sémantique et causale des états mentaux. Ces propriétés émergentes peuvent se passer de toute attribution consciente, et sont tout à fait susceptibles d'être étudiées de manière mécanistique sous l'angle des neurosciences.

Conclusion

Dans cet article, nous avons détaillé la manière dont la philosophie de l'esprit peut être utilisée pour appréhender la clinique et les théories en psychiatrie. Nous avons vu que dans un double effet de mise en abyme, les sciences cognitives et la philosophie de l'esprit évoluent en synergie. Les données empiriques sont générées à partir d'une trame théorique tissée par des conceptions philosophiques, et de nouvelles questions philosophiques éclosent sans cesse des résultats empiriques. De la même manière, la philosophie de l'esprit occupe une place majeure dans la genèse des théories en psychiatrie, d'autant plus à l'ère de la psychiatrie computationnelle. Partageant les mêmes objets d'intérêt que les sciences cognitives, mais proposant des méthodes d'investigation différentes, elle peut fournir à la psychiatrie l'ossature conceptuelle pour relier différents niveaux d'explication : du biologique au fonctionnel, de la neurophysiologie à la cognition, de la matière à l'esprit.

Les auteurs déclarent n'avoir pas de liens d'intérêts en rapport avec ce travail.

Bibliographie

1. Kim J. *Philosophy of Mind*. Westview Press; 1996.
2. Putnam H. *Philosophy and Our Mental Life*. In: *Mind, Language, and Reality*. Cambridge University Press; 1975.
3. Brook A. Introduction: philosophy in and philosophy of cognitive science. *Top Cogn Sci*. 2009 Apr;1(2):216–30.

4. Stepp N, Chemero A, Turvey MT. Philosophy for the rest of cognitive science. *Top Cogn Sci.* 2011 Apr;3(2):425–37.
5. Putnam H. *The Threefold Cord: Mind, Body, and World.* Columbia University Press; 1999.
6. Richardson WJ. Phenomenology and Psychoanalysis. *J Phenomenol Psychol.* 1980;11(2):1–20.
7. Thompson MG. *Psychoanalysis and Phenomenology: Toward a Human Synthesis*, by Miguel Iturrate. *J Phenomenol Psychol.* 1995;26:114–6.
8. Gomez L. *The Freud Wars: An Introduction to the Philosophy of Psychoanalysis.* Routledge; 2005.
9. Montague PR, Dolan RJ, Friston KJ, Dayan P. Computational psychiatry. *Trends Cogn Sci.* 2012 Jan;16(1):72–80.
10. Andler D. Cognitive Science. In: Kritzman L, editor. *The Columbia History of Twentieth Century French Thought.* 2006.
11. Andler D. De la philosophie des sciences à la philosophie de l'esprit. 2007 Jan 1;Mélika Ouelbani, dir. *La philosophie analytique dans tous ses états, Tunis : Université de Tunis, 2007, pp. 59–87.*
12. Bechtel W. How can philosophy be a true cognitive science discipline? *Top Cogn Sci.* 2010 Jul;2(3):357–66.
13. Bechtel W. Constructing a philosophy of science of cognitive science. *Top Cogn Sci.* 2009 Jul;1(3):548–69.
14. Ryle G. *The Concept of Mind.* Hutchinson & Co; 1949.
15. Searle JR. The future of philosophy. *Philos Trans R Soc Lond B Biol Sci.* 1999 Dec 29;354(1392):2069–80.
16. Quine WV. *Epistemology Naturalized.* In: *Ontological Relativity and Other Essays.* New York: Columbia University Press; 1969.
17. Davidson D. *Psychology as Philosophy.* In: Brown S, editor. *Philosophy of Psychology.* Harper & Row; 1974.
18. Turing AM. Computing Machinery and Intelligence. *Mind.* 1950;59(October):433–60.
19. Putnam H. *Mind and Body.* In: *Reason, Truth and History.* Cambridge University Press; 1981.
20. Weintraub ER. *Statistics on the Table: The History of Statistical Concepts and Methods.* By Stephen M. Stigler. Cambridge, MA: Harvard University Press, 1999. Pp. ix, 488. 19.95, paper. *J Econ Hist.* 2003 Jun;63(2):616–7.
21. Clark A. *Surfing Uncertainty: Prediction, Action, and the Embodied Mind.* Oxford University Press; 2015. 425 p.
22. Thagard P. Why cognitive science needs philosophy and vice versa. *Top Cogn Sci.* 2009 Apr;1(2):237–54.
23. Popper KR, Eccles JC. *The Self and Its Brain: An Argument for Interactionism.* Springer; 1977.
24. Kim J. *Mind in a Physical World.* Cambridge: MIT Press; 1998;
25. Gibb S. Non reductive physicalism and the problem of strong closure. *Am Philos Q.* 2012;49(1):29–41.

26. Hempel CG. *Philosophy of Natural Science*. Englewood Cliffs, N.J., Prentice-Hall; 1966.
27. Kim J. *Physicalism, or Something Near Enough*. Princeton University Press; 2005.
28. Feigl H. *Concepts, Theories, And The Mind-Body Problem*. University of Minnesota Press; 1958.
29. Smart J. Sensations and Brain Processes. *Philos Rev*. 1959;68(April):141–56.
30. Place UT. Is Consciousness a Brain Process. *Br J Psychol*. 1956;47(1):44–50.
31. Melnyk A. Formulating Physicalism: Two Suggestions. *Synthese*. 1995;105(3):381–407.
32. Dowell JL. Formulating the Thesis of Physicalism. *Philos Stud*. 2006;131(1):1–23.
33. Putnam H. *The Many Faces of Realism*. Open Court; 1987.
34. Putnam H. *Minds and Machines*. In: Hook S, editor. *Dimensions of Minds*. New York, USA: New York University Press; 1960. p. 138–64.
35. Putnam H. *Brains in a Vat*. In: Bernecker S, Dretske FI, editors. *Knowledge: Readings in Contemporary Epistemology*. Oxford University Press; 1999. p. 1–21.
36. Harbecke J. A realist approach to biological events. *Hist Philos Life Sci*. 2005;27(3–4):491–507.
37. Strappini F, Martelli M, Cozzo C, di Pace E. Empirical Evidence for Intraspecific Multiple Realization? *Front Psychol*. 2020;11:1676.
38. Ellis G. Top-down effects in the brain. *Phys Life Rev*. 2019;31:11–27.
39. Résibois M, Verduyn P, Delaveau P, Rotgé J-Y, Kuppens P, Van Mechelen I, et al. The neural basis of emotions varies over time: different regions go with onset- and offset-bound processes underlying emotion intensity. *Soc Cogn Affect Neurosci*. 2017 01;12(8):1261–71.
40. Stephan KE, Kasper L, Harrison LM, Daunizeau J, den Ouden HEM, Breakspear M, et al. Nonlinear dynamic causal models for fMRI. *NeuroImage*. 2008 Aug 15;42(2):649–62.
41. Davidson D. *Mental Events*. In: Foster L, Swanson JW, editors. Clarendon Press; 1970. p. 207–24.
42. Davidson D. *Belief and the Basis of Meaning*. *Synthese*. 1974;27(July-August):309–23.
43. Mashour GA, Roelfsema P, Changeux J-P, Dehaene S. Conscious Processing and the Global Neuronal Workspace Hypothesis. *Neuron*. 2020 04;105(5):776–98.
44. Friston KJ. The free-energy principle: a unified brain theory? *Nat Rev Neurosci*. 2010;11(2):127–38.
45. Hohwy J, Rosenberg R. Cognitive Neuropsychiatry: Conceptual, Methodological and Philosophical Perspectives. *World J Biol Psychiatry*. 2005;6(3):192–7.
46. Mujica-Parodi LR, Strey HH. Making Sense of Computational Psychiatry. *Int J Neuropsychopharmacol*. 2020 May 27;23(5):339–47.
47. Yablo S. *Mental Causation*. *Philos Rev*. 1992;101(2):245–80.
48. Kim J. Does the Problem of Mental Causation Generalize? *Proc Aristot Soc*. 1997;97(3):281–97.
49. Woodward J. Causation in Biology: Stability, Specificity, and the Choice of Levels of Explanation. *Biol Philos*. 2010;25(3):287–318.

50. Hohwy J. The Experience of Mental Causation. *Behav Philos.* 2004;32(2):377–400.
51. Hohwy J. Functional Integration and the Mind. *Synthese.* 2007;159(3):315–28.
52. Gerrans P, Hohwy J. *The Routledge Companion to Philosophy of Mental Disorder.* Routledge; 2011.
53. Hohwy J, Kallestrup J. *Being Reduced: New Essays on Reduction, Explanation, and Causation.* Oxford University Press; 2008.
54. Wittgenstein L. *Philosophical Investigations.* Wiley-Blackwell; 1953.
55. Jackson F. Epiphenomenal Qualia. *Philos Q.* 1982;32(April):127–36.
56. Nagel T. What is It Like to Be a Bat? *Philos Rev.* 1974;83(October):435–50.
57. Levine J. Materialism and Qualia: The Explanatory Gap. *Pac Philos Q.* 1983;64(October):354–61.
58. Insel TR. The NIMH Research Domain Criteria (RDoC) Project: precision medicine for psychiatry. *Am J Psychiatry.* 2014 Apr;171(4):395–7.
59. Lilienfeld SO. The Research Domain Criteria (RDoC): an analysis of methodological and conceptual challenges. *Behav Res Ther.* 2014 Nov;62:129–39.
60. Looi JCL, Liberg B. Between Scylla and Charybdis: DSM-5/ICD-11 and RDoC in neuropsychiatry? *Aust N Z J Psychiatry.* 2015 Jan;49(1):82–3.
61. Lilienfeld SO, Treadway MT. Clashing Diagnostic Approaches: DSM-ICD Versus RDoC. *Annu Rev Clin Psychol.* 2016;12:435–63.
62. Ibáñez A, Cosmelli D. Moving beyond computational cognitivism: understanding intentionality, intersubjectivity and ecology of mind. *Integr Psychol Behav Sci.* 2008 Jun;42(2):129–36.
63. Beedie CJ, Terry PC, Lane AM. Distinguishing Mood From Emotion. *Cogn Emot.* 2005;19:847–78.
64. Chang S-S, Chou T. A Dynamical Bifurcation Model of Bipolar Disorder Based on Learned Expectation and Asymmetry in Mood Sensitivity. *Comput Psychiatry Camb Mass.* 2018 Dec;2:205–22.
65. Clark JE, Watson S, Friston KJ. What is mood? A computational perspective. *Psychol Med.* 2018 Oct;48(14):2277–84.