

Zeitschrift: Revue de Théologie et de Philosophie
Herausgeber: Revue de Théologie et de Philosophie
Band: 42 (1992)
Heft: 3: Philosophie de l'action

Artikel: Externalisme et causalité mentale
Autor: Jacob, Pierre
DOI: <https://doi.org/10.5169/seals-381488>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. [Mehr erfahren](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. [En savoir plus](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. [Find out more](#)

Download PDF: 10.01.2026

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

EXTERNALISME ET CAUSALITÉ MENTALE*

PIERRE JACOB

Résumé

D'après le sens commun, les croyances (et les attitudes propositionnelles en général) sont des causes de l'action intentionnelle et des raisons de ce que nous faisons. Le contenu de ce que nous croyons est tenu pour une propriété causalement efficace de ce que nous faisons intentionnellement. Dans la conception externaliste adoptée par la majorité des philosophes contemporains, le contenu d'une croyance d'un individu dépend de l'environnement de l'individu. La thèse de l'efficacité causale des propriétés intentionnelles des attitudes propositionnelles s'expose à deux menaces épiphénoménalistes: si les propriétés intentionnelles sont tenues pour des propriétés fonctionnelles du cerveau d'un individu, leur efficacité causale risque d'être menacée de préemption par l'efficacité causale des propriétés physiques du cerveau de l'individu. Si on admet l'externalisme, alors les propriétés intentionnelles ne sont pas des propriétés fonctionnelles du cerveau d'un individu, et elles risquent d'être privées d'efficacité causale si on suppose que les propriétés causalement efficaces sont des propriétés locales de l'individu.

Deux thèses du sens commun sur la causalité mentale

Le sens commun souscrit, je crois, à deux thèses d'inégale force sur la causalité mentale. Selon la thèse la plus faible, les attitudes propositionnelles sont des causes de l'action intentionnelle et de la formation de nouvelles attitudes propositionnelles. Je me tiens à cloche-pied sur mon pied gauche parce que j'ai l'intention de me tenir à cloche-pied sur mon

* Je remercie Daniel Schulthess, Richard Glauser et Roberto Casati de m'avoir invité au colloque de l'Université de Neuchâtel sur la philosophie de l'action des 15 et 16 novembre 1991 qu'ils ont organisé. Je remercie les participants au colloque de leurs critiques — en particulier Kevin Mulligan qui avait préparé des commentaires écrits et Filip Buekens qui m'a envoyé ultérieurement des critiques tout à fait pertinentes. Enfin je remercie un relecteur anonyme de la *Revue de Théologie et de Philosophie*.

pied gauche. J'ai l'intention de boire un verre d'eau Perrier. Je crois qu'il y a une bouteille d'eau Perrier dans le frigidaire. J'ouvre le frigidaire pour découvrir qu'il contient non une bouteille d'eau Perrier mais une bouteille de limonade. Je renonce à ma précédente croyance et modifie mon intention de boire un verre d'eau Perrier en intention de boire un verre de limonade.

Selon le physicalisme des particuliers¹ (*token physicalism*), chaque état mental intentionnel d'un individu est un état cérébral de l'individu. Si le physicalisme des particuliers est correct — ainsi que je le supposerai —, alors la thèse causale faible (réhabilitée par Davidson 1963) est vraie: si chaque attitude propositionnelle particulière est un état cérébral particulier et si un état cérébral particulier peut être une cause, alors une attitude propositionnelle particulière aussi peut être une cause.

Selon la thèse causale forte, les attitudes propositionnelles sont des causes de l'action intentionnelle et de la formation de nouvelles attitudes propositionnelles *en vertu de leur contenu*. Mon intention de me tenir à cloche-pied sur mon pied gauche, non mon intention de lever ma main droite, explique causalement pourquoi (toutes choses étant égales) je me tiens intentionnellement à cloche-pied sur mon pied gauche. Mon intention de boire un verre d'eau Perrier, combinée à ma croyance qu'il y a une bouteille d'eau Perrier dans le frigidaire, non à ma croyance que Edith Cresson est l'actuel Premier ministre du gouvernement français, contribue à expliquer causalement pourquoi j'ouvre le frigidaire. Selon la thèse forte, le contenu des attitudes propositionnelles est une propriété causalement efficace.

Admettons que les types d'attitudes propositionnelles soient des propriétés mentales. Si le physicalisme des particuliers est correct, alors les propriétés mentales ne sont ni exprimables ni définissables au moyen de prédictats exprimant (ou dénotant) des propriétés physiques de base. Ma croyance particulière que Londres est jolie n'est autre que l'activation de l'un de mes états cérébraux. Mon état cérébral particulier possède lui-même plusieurs propriétés: outre la propriété mentale d'être une incarnation particulière de *la croyance (type) que Londres est jolie*, il a des propriétés physiques (électriques, chimiques, hormonales, etc.). Supposant — comme tous les tenants du physicalisme des particuliers — que les propriétés physiques du cerveau d'un individu sont causalement efficaces, la question

¹ Le physicalisme des particuliers s'oppose au physicalisme des propriétés (ou des types) selon lequel les propriétés intentionnelles ou mentales sont au minimum des propriétés physiques et au maximum des propriétés physiques nomologiques. La «réalisabilité multiple» du mental par le physique a été soulignée initialement par Putnam et est devenue une thèse fondamentale de la littérature *fonctionnaliste* en philosophie de la psychologie: un programme informatique donné peut être «implémenté» par des ordinateurs ayant des propriétés physiques distinctes.

se pose naturellement de savoir si les propriétés mentales peuvent, comme le veut la thèse causale forte du sens commun, être causalement efficaces aussi. Sinon, et quoique les états mentaux particuliers soient des causes, les propriétés mentales se révéleront épiphénoménales. Seront-elles pour autant inutiles dans des explications causales ?

Avec Davidson, j'admettrai que la relation causale entre événements ou états est une relation extensionnelle non épistémique. Mais le fait que la relation causale soit extensionnelle laisse indéterminée la question de savoir quelles *propriétés* d'une cause sont *causalement efficaces* dans la production de son effet (cf. Horgan 1989, Kim 1984 et Sosa 1984). De plus, je distinguerai la relation causale de la notion *d'explication causale* qui est sujette à des contraintes épistémiques et pragmatiques complexes. Une propriété peut jouer un rôle dans une explication causale sans posséder d'efficacité causale.

Deux menaces épiphénoménalistes

Je veux distinguer deux menaces épiphénoménalistes qui pèsent sur la thèse causale forte du sens commun — une menace modérée et une menace forte — selon que les propriétés mentales sont supposées covarier ou non (*supervene on*) avec les propriétés physiques du cerveau d'un individu. Cette covariation dépend à son tour des principes d'individuation du contenu des attitudes propositionnelles.

Supposons que les propriétés mentales d'un individu covariant avec les propriétés physiques de son cerveau. Alors la menace modérée consiste en ce que l'efficacité causale des propriétés mentales risque d'être (selon l'expression de Lepore & Loewer 1987) *filtrée* ou *préemptée* par l'efficacité causale des propriétés physiques du cerveau de l'individu. Si le mental est susceptible de «réalisation multiple» par le physique, alors deux individus — disons, un francophone et un anglophone — dont les cerveaux ont des propriétés physiques distinctes se verront assigner une seule et même propriété mentale - par exemple, la croyance que Londres est jolie. Le fait que le mental soit susceptible de réalisation physique multiple fait peser la menace modérée que les propriétés physiques distinctes du cerveau de chaque locuteur, non leur propriété mentale commune, aient un rôle dans l'explication causale du fait que l'un énonce la phrase française «Londres est jolie» et l'autre énonce la phrase anglaise synonyme «London is pretty».

La menace la plus grave est suscitée par l'externalisme, c'est-à-dire la doctrine selon laquelle, non seulement les propriétés mentales sont inexprimables au moyen de prédictats dénotant des propriétés physiques de base, mais elles ne covariant même pas avec les propriétés physiques du cerveau d'un individu. Comme le montrent les célèbres expériences de pensée de

Putnam (1975a) et de Burge (1979), selon nos attributions de sens commun, le contenu des attitudes propositionnelles ne covarie pas avec les propriétés physiques du cerveau d'un individu. Si l'externalisme est la conception correcte de l'individuation du contenu de nombre d'attitudes propositionnelles — comme je le supposerai —, alors les contenus respectifs d'une paire de croyances simultanément attribuables à une paire de jumeaux microphysiques peuvent différer l'un de l'autre. La tension entre l'externalisme et la causalité mentale repose sur l'idée plausible selon laquelle la causalité est *locale*. Comme l'a exprimé McGinn (1989: 133):

«... les aspects de la cause qui contribuent à faire advenir l'effet doivent se situer à l'endroit exact où l'interaction causale se produit. Il en va de la causalité cérébrale ou mentale comme des boules de billard. Leurs effets dépendent des propriétés locales de ces entités. Les pouvoirs causaux d'un état ou d'une propriété leur sont intrinsèques; ils ne peuvent dépendre essentiellement de relations avec des entités éloignées.»

Si l'externalisme est correct, alors les contenus d'une paire de croyances attribuables à une paire de jumeaux microphysiques peuvent être aussi différents l'un de l'autre que les paires suivantes: un autographe authentique de Winston Churchill et une imitation (Dennett 1983: 44); un tableau authentique de Picasso et une copie; un billet authentique de 500 francs et un faux billet; une photographie représentant Jules et une photographie représentant Gilles, son frère jumeau identique (Dretske 1990: 7).

Vraisemblablement, une imitation parfaite (d'un autographe de Churchill, d'un tableau de Picasso ou d'un billet de 500 francs) et l'original peuvent renvoyer des photons ou passer divers tests chimiques sans qu'un physicien ou un chimiste puisse détecter la moindre différence entre l'imitation et l'original. Mais quoiqu'ils puissent être physiquement et chimiquement indiscernables, l'original et la copie auront des valeurs financières (sinon esthétiques) différentes. Comme le dit Dretske (1990: 7), ce qui confère à l'original une valeur financière (et/ou esthétique) supérieure, ce sont «certains faits historiques et relationnels... qui ne covarient pas avec les propriétés physiques... de la toile ou du papier qui a cette valeur, propriétés (physiques) qui déterminent les pouvoirs causaux de l'objet». La différence historique entre l'original et la copie peut ne pas être physiquement détectable. L'éénigme suscitée par l'externalisme pour la causalité mentale tient à ce que les différences de contenu proviennent d'origines historiques distinctes et que ces différences historiques peuvent ne laisser aucune trace physique.

Geach (1969: 72, 99) a baptisé *changements de Cambridge* le changement advenu à Socrate lorsqu'il est devenu physiquement plus petit que Théétète à la suite de la croissance de Théétète et les changements «posthumes» qui adviennent à Socrate «chaque fois qu'il est l'objet de l'admiration d'un écolier». En adaptant l'expression forgée par Geach, on pour-

rait dire que l'externalisme fait peser sur la causalité mentale la menace que les propriétés mentales d'un individu se révèlent être des *propriétés de Cambridge* de l'individu.

Notez que si l'externalisme menace la causalité mentale, il est compatible avec le physicalisme: l'identité des propriétés mentales pertinentes dépend des principes externalistes d'individuation du contenu; quoique le contenu ne covarie pas avec les seules propriétés physiques du cerveau d'un individu, il peut covarier avec une paire de propriétés physiques dont un membre appartient au cerveau de l'individu et l'autre appartient à l'environnement physique de l'individu.

La menace de préemption

Supposons que j'aie supprimé ma douleur au pied droit en ayant avalé un comprimé d'aspirine. Le comprimé a soulagé ma douleur en vertu de sa propriété chimique — le fait qu'il est composé d'acide acétylsalicylique. Je pourrais aussi présenter une explication causale concurrente du même fait en vous informant que j'ai avalé un *antalgique*. Etre antalgique est une propriété *fonctionnelle* du comprimé d'aspirine — une propriété que le comprimé peut partager avec d'autres comprimés ayant une composition chimique différente. Cette propriété fonctionnelle, qui covarie avec la propriété chimique du comprimé, est définissable (à la Ramsey, Carnap, Lewis) comme une propriété chimique de second ordre qui consiste en ce que le comprimé possède l'une des propriétés chimiques de premier ordre (causalement efficaces) à l'intérieur d'une classe disjonctive de ces propriétés chimiques de premier ordre. Certes, comme le soutiennent Block (1990) et Jackson & Pettit (1989, 1990a), dans le processus grâce auquel ma douleur est soulagée, être antalgique n'a pas une efficacité causale directe comme la présence dans le comprimé de l'acide acétylsalicylique. Si la propriété chimique de premier ordre est causalement efficace, son efficacité causale ne *préempte-t-elle* pas l'efficacité causale de la propriété fonctionnelle? Kim (1989, 1990) baptise ce problème le problème de l'«exclusion explicative».

La menace de préemption (ou d'exclusion) explicative peut être écartée, selon moi, en conférant deux rôles distincts à des propriétés dans une explication causale ou en distinguant deux sortes d'explication causale. Considérons deux explications possibles d'un même *explanandum*: le verre s'est cassé parce que *Jules* l'a laissé tomber vs le verre s'est cassé parce que *quelqu'un* l'a laissé tomber. Le premier *explanans* est une proposition singulière, le second est une proposition générale. Si la première est vraie, alors la seconde est vraie en vertu de la première. Quoique les mécanismes causaux responsables du bris du verre soient identiques, les explications diffèrent par l'information supplémentaire. De manière analogue, une explication

causale de la suppression de ma douleur peut à l'évidence mentionner ou faire référence à la propriété chimique du comprimé d'aspirine que j'ai avalé. Jackson & Pettit (1989, 1990a) qualifient d'explication par le *processus* une telle explication. Une explication causale du même fait peut aussi consister à mentionner la propriété fonctionnelle qu'a le comprimé d'être un antalgique. Jackson & Pettit (1989, 1990a) qualifient d'*explication-programme* une telle explication que je nommerai simplement *fonctionnelle* et qui consiste à *quantifier* sur un ensemble de propriétés chimiques (de premier ordre) causalement efficaces. Si l'explication du premier type est vraie, alors l'explication du second type est vraie en vertu de l'explication du premier type.

Une explication causale qui mentionne une propriété chimique (de premier ordre) ne «préempte» pas une explication fonctionnelle parce qu'elles n'offrent pas la même information sur le processus chimique en cause. La première révèle le chemin réellement parcouru par le processus de suppression de la douleur. La seconde révèle une information plus générale sur une classe de chemins possibles et non réellement parcourus. Supposer que la première prive automatiquement la seconde de raison d'être, c'est admettre ce que Jackson & Pettit (1990b) ont nommé «la préférence explicative pour le détail», c'est-à-dire supposer que l'information la plus détaillée sur un processus causal est toujours préférable à une information plus générale. Avec Jackson & Pettit, et pour des raisons déjà mises en lumière par Putnam (1975b), je ne crois pas que la préférence explicative pour le détail soit justifiée. Soit le fameux dé de Putnam de 4,5 cm de côté. Il peut passer par un orifice carré de 5 cm de côté, mais non par un orifice circulaire de 5 cm de diamètre. Pourquoi? Putnam montre élégamment qu'une explication géométrique de ce fait qui ignore la composition physique du dé et des bords des orifices a conjointement une plus grande simplicité et une plus grande généralité que l'explication physique.

Propriétés de Cambridge et explication causale

Comme l'a fait observer Dretske (1988b: 80), toutes les propriétés d'une cause ne contribuent pas également à conférer à la cause son efficacité causale. Avaler un comprimé d'aspirine peut soulager une douleur. Mais aucune des propriétés suivantes du comprimé ne confère au comprimé son efficacité causale dans le processus chimique de suppression de la douleur: son prix, le nom de la rue dans laquelle j'ai acheté le flacon contenant le comprimé, la couleur des lettres sur l'étiquette du flacon.

Même la forme géométrique du comprimé — le fait qu'il ait la forme d'une section de cylindre — ne confère au comprimé aucune efficacité dans le processus de soulagement de la douleur. Comme l'a fait remarquer Dretske (1988a: 32), eût-il été sphérique ou cubique, il aurait soulagé ma douleur pareillement. Je ne veux pas suggérer que la forme géométrique du comprimé est dénuée de toute efficacité causale: avaler un comprimé pyramidal d'aspirine peut érafler la gorge en vertu de la forme du comprimé. Mais la forme est dénuée d'efficacité causale dans le processus chimique de suppression de la douleur. Même le fait d'avoir *une forme quelconque* est une propriété causalement inefficace du comprimé d'aspirine dans le processus de suppression de la douleur puisque le comprimé peut être réduit en poudre, privé de forme, et le fait d'avaler la poudre peut néanmoins soulager la douleur. Eu égard au processus de suppression de la douleur, la forme est une propriété de Cambridge du comprimé.

La menace externaliste sur la causalité mentale n'est pas que l'efficacité causale des propriétés mentales d'un individu soit préemptée par l'efficacité causale des propriétés physiques du cerveau de l'individu. Si l'externalisme est correct, alors le contenu n'est pas une propriété fonctionnelle du cerveau d'un individu. Il ne covarie pas plus avec les propriétés physiques du cerveau d'un individu que le prix (ou la forme) du comprimé d'aspirine (ou n'importe laquelle de ses propriétés de Cambridge) ne covarie avec sa propriété chimique. En vertu des principes d'individuation externaliste, le contenu risque d'être une propriété de Cambridge du cerveau d'un individu. Manifestement, la menace externaliste sur la causalité mentale ne peut pas être dissipée par la même manœuvre que la menace de préemption.

Notez que l'acuité des deux menaces d'épiphénoménalisme examinées ici — préemption et externalisme — dépend de la supposition rendue explicite notamment par Dretske (1988a), Kim (1984) et Sosa (1984) selon laquelle toutes les propriétés d'une cause ne concourent pas également à la production de son effet. Cette supposition n'est-elle pas incompatible avec la doctrine davidsonienne selon laquelle la relation causale est purement extensionnelle? Je ne le crois pas. Comme je l'ai indiqué plus haut, même si la relation causale est purement extensionnelle, on peut savoir qu'une relation de causalité tient entre deux événements sans savoir quelles propriétés de la cause sont causalement efficaces dans le processus causal. De surcroît, la causalité et l'efficacité causale des propriétés doivent à mes yeux être distinguées de la notion d'*explication causale* qui consiste à fournir une information pertinente sur un processus causal.

Le défi de l'externalisme et les explications non causales

Le sens commun a sans doute tort de tenir le contenu des attitudes propositionnelles pour des propriétés causalement efficaces. Aurait-il tort de tenir les explications dans lesquelles le contenu est pertinent pour des expli-

cations *causales*? Toutes les explications correctes ne sont pas des explications causales. Considérons les questions suivantes: pourquoi mon fils est-il un neveu? Pourquoi ma femme deviendra-t-elle veuve après ma mort? Pourquoi dois-je écrire deux fois «r» pour écrire «Pierre»? Pourquoi le nombre 4 est-il plus grand que le nombre 2?

Comme l'a fait valoir Kim (1973, 1974), la meilleure explication du fait que mon fils est un neveu consiste à présenter la définition du concept exprimé par le mot «neveu» et à montrer que mon fils la satisfait en vertu du fait que je suis son père et que j'ai des frères et sœurs. Et cela indépendamment du fait que mon fils est mon fils en vertu d'un processus biologique et qu'il a pu devenir un neveu *après* que j'ai acquis des frères et sœurs. De même, les réponses pertinentes aux trois autres questions seront des explications conceptuelles: elles révéleront des liens de dépendance conceptuels (et non causaux). Pas plus que la question de savoir pourquoi tel ou tel théorème d'arithmétique est vrai, ces questions ne sont des *explananda* pour des explications causales. Pourquoi tel mathématicien tient pour vrai tel théorème d'arithmétique est en revanche un *explanandum* pour une explication causale.

Il existe des questions ambiguës qui peuvent tantôt s'interpréter comme une requête pour une explication causale tantôt comme une requête pour une explication conceptuelle. Considérons la question: pourquoi est-ce que j'habite à Paris? Je peux répondre que j'habite à Paris parce que j'ai déménagé de Lille à Paris ou parce que j'ai trouvé un emploi à Paris. Dans le contexte d'une question portant sur mon droit de prendre part à un vote destiné à envoyer un député de Paris à l'Assemblée Nationale, je peux aussi répondre que j'habite à Paris parce que (ou en vertu du fait que) j'habite dans le troisième arrondissement de Paris. A la différence de la première explication, la seconde n'est pas une explication causale.

Pour percevoir la pertinence des explications non causales dans le défi de l'externalisme, considérons Jules et Gilles, deux jumeaux microphysiques identiques dont les cerveaux sont physiquement indiscernables. Chacune de leurs croyances simultanément attribuables peuvent avoir des contenus différents. Lorsque Jules forme la croyance perceptuelle que la tasse t_1 devant lui contient du café, Gilles entretient la croyance perceptuelle jumelle que la tasse t_2 située devant lui contient du café. Supposons que t_1 et t_2 soient deux tasses perceptuellement indiscernables contenant exactement la même quantité de café. Les conditions de vérité des pensées respectives de Jules et de Gilles diffèrent l'une de l'autre en vertu du fait que la première est vraie si et seulement si t_1 contient du café et la seconde est vraie si et seulement si t_2 contient du café. Si, à l'insu de Jules, nous permutions t_1 et t_2 de sorte que Jules attribue la propriété de contenir du café à t_2 qu'il prend pour t_1 , nous serons enclins à tenir sa croyance pour *fausse*, non pour vraie. Puisqu'elles ont des conditions de vérité différentes,

la croyance de Jules et celle de Gilles sont différentes. Puisqu'elles sont différentes, elles ne covarient pas avec les propriétés physiques du cerveau de Jules et de Gilles (qui sont par hypothèse identiques).

La croyance de Jules est une cause du fait qu'il boit du café de la tasse t_1 . La croyance de Gilles est une cause du fait qu'il boit du café de la tasse t_2 . Lorsque Jules et Gilles saisissent la tasse placée devant eux avec les doigts de leur main droite, la mettent en contact avec leurs lèvres, inclinent leurs têtes légèrement en arrière et boivent une gorgée de café, accomplissent-ils la même action? Tous les philosophes ne donnent pas la même réponse à cette question. Ils accomplissent les mêmes mouvements corporels. Mais accomplissent-ils la même action intentionnelle? Kim (1982) et Stich (1983: 160-70) sont de ceux qui pensent que, pour les besoins de la psychologie (scientifique), Jules et Gilles doivent être supposés faire la même chose. Evans (1982: 200-204), Hornsby (1986) et Peacocke (1981) sont de ceux qui font valoir qu'intentionnellement décrites les actions de Jules et Gilles diffèrent. S'ils ne font pas la même chose, la *différence* entre les contenus de leurs croyances respectives ne joue-t-elle pas un rôle dans l'explication causale de la différence entre leurs actions? Si c'est le cas, alors le contenu joue un rôle dans l'explication causale.

Qu'on s'accorde pour penser que Jules et Gilles font la même chose ou non, ce serait une erreur de croire que l'identité ou la non-identité de ce qu'ils font est un *explanandum* pour une explication causale. La croyance de chaque jumeau est une cause de ce qu'il fait. Les actions respectives de chaque jumeau sont deux processus physiques indépendants l'un de l'autre. Pour provoquer l'intuition que deux jumeaux microphysiquement identiques peuvent avoir des croyances ayant des contenus distincts, l'expérience de pensée typique de l'externaliste met en scène une *coïncidence* entre deux processus physiques causalement indépendants l'un de l'autre. La coïncidence est un événement conceptuellement (non physiquement) complexe: c'est la somme logique des deux processus séparés. Elle se compose des deux effets séparés de deux processus physiques indépendants l'un de l'autre. Mettre en lumière les similitudes et les différences entre les causes respectives et les effets respectifs des deux processus jumeaux n'est pas fournir une troisième explication causale en supplément des deux explications causales de ce que chaque jumeau fait séparément. C'est une décomposition de la coïncidence complexe en ses constituants conceptuels. La différence entre une explication causale et une explication (conceptuelle) non causale tient à ce que la première, non la seconde, fournit une information sur un processus ou un changement physique. Chaque explication causale de ce que fait un jumeau apporte une information sur un processus physique. La comparaison des deux processus physiques ne caractérise pas un troisième processus physique.

Jackson et Pettit (1989) distinguent deux sortes d'explication causale: l'explication-*processus* et l'explication-*programme*. Dans une explication-*processus*, il est fait référence à une propriété causalement efficace de la cause dans la production de l'effet. Mais dans une explication-*programme*, il est fait référence à «un trait qui *programme* causalement sans causer». Ils donnent plusieurs exemples d'explications-*programmes*: le prix de l'agneau au marché de Portland était le même qu'au marché de Ballarat parce que la relation entre l'offre et la demande était la même aux deux endroits. La longévité de la dynastie des Bourbons a été supérieure à la longévité des Windsors parce que les premiers ont été plus prolifiques que les seconds. Les arbres croissent plus vite à Melbourne qu'à Canberra parce qu'il gèle moins à Melbourne qu'à Canberra.

«La propriété supposée expliquer le résultat est une propriété commune aux différents membres d'un ensemble de situations possibles dont chacun des membres aurait produit le résultat et dont l'un a produit effectivement le résultat. Nous caractérisons cette propriété commune comme celle qui programme causalement le résultat et appelons les explications qui mentionnent une propriété qui programme le résultat sans le produire des explications-*programmes*» (Jackson et Pettit, 1989: 393).

Je suis donc en désaccord avec la notion d'explication-*programme* de Jackson et Pettit dans la mesure où leur concept ne permet pas de discriminer entre des explications causales et des explications conceptuelles non causales, comme l'explication du fait que si vous imprimez des forces égales à deux boules de billard de masse égale, vous conférez aux deux boules des accélérations égales. Avoir la même accélération n'est pas un *explanandum* pour une explication causale. Que deux boules aient la même masse et qu'on leur imprime la même force sont des propriétés relationnelles des boules pertinentes pour une réponse conceptuelle à une question non causale. A la différence de la masse particulière d'une boule et de la force particulière qui lui est imprimée, ces deux propriétés relationnelles ne font pas partie de l'*explanans* d'une explication causale. Chaque boule est assujettie à un processus physique indépendant de l'autre. Lorsqu'on compare ces deux processus, on ne fournit pas une explication causale en supplément des deux explications causales séparées qui décrivent deux processus physiques indépendants l'un de l'autre. Jackson & Pettit (1989) amalgament à tort, à mon avis, sous la notion de propriété programmatrice le rôle explicatif joué par une propriété fonctionnelle (comme la propriété d'un comprimé d'être un antalgique) et la relation d'avoir la même masse qui tient entre deux boules de billard.

Je me tourne maintenant vers trois exemples de propriétés historiques relationnelles d'un système qui ne covarient pas avec ses propriétés physiques. La question est de savoir si ces propriétés jouent un rôle dans des explications causales du comportement du système.

L'argument du remplacement de Stich

Stich (1978) défendait ce qu'il appelait le «principe d'autonomie» selon lequel, faute de covarier avec les propriétés physiques du cerveau d'un individu, les propriétés d'un individu ne seront pas pertinentes pour une explication psychologique causale de son comportement. En faveur de ce principe, Stich (1983) a offert ce qu'il appelle l'«argument du remplacement». Supposons que je sois kidnappé et remplacé par une réplique qui soit mon clone moléculaire. Quoique nous soyons physiquement et chimiquement indistinguables, je peux faire certaines choses que ma copie ne peut pas faire, comme divorcer d'avec ma femme ou vendre ma voiture.

Ce qui me confère à moi, et non à ma copie, le pouvoir de vendre ma voiture ou de divorcer d'avec ma femme est que, à la différence de ma copie, j'ai acheté ma voiture et j'ai épousé ma femme. En mentionnant respectivement le fait que la relation de mariage tient entre ma femme et moi, et non entre ma femme et ma copie, et le fait que je suis, à la différence de ma copie, propriétaire de ma voiture, que fais-je d'autre qu'expliquer causalement le fait que, à la différence de ma copie, je peux divorcer d'avec ma femme ou vendre ma voiture? Remarquez que le fait d'être uni par le mariage avec ma femme ne covarie pas avec mes propriétés physiques, *a fortiori* ne covarie-t-il pas avec les propriétés physiques de mon cerveau. Pas davantage que le fait d'avoir été peint par la main de Picasso ne covarie avec les propriétés physiques d'une toile. Si de telles propriétés relationnelles légales et historiques de ma personne peuvent jouer un rôle dans une explication causale, pourquoi le contenu d'une croyance ne le pourrait-il pas?

Qu'est-ce que le fait que je suis uni par le mariage avec ma femme permet exactement d'expliquer? Cela ne permet pas d'expliquer pourquoi j'ai divorcé d'avec ma femme puisque je n'ai pas divorcé d'avec elle. Cela explique pourquoi je *peux* divorcer. Si je n'ai pas divorcé, aucun processus physique qui puisse faire l'objet d'une explication causale ne s'est déroulé: le fait d'être marié me confère la possibilité juridique de divorcer; mais je peux conserver cette propriété juridique relationnelle toute ma vie sans qu'elle entre dans une explication causale si tant est que je ne divorce pas. De même, le fait que je suis propriétaire de ma voiture explique, non pourquoi je l'ai vendue, mais pourquoi je *pourrais* la vendre. Ni pourquoi je *peux* divorcer d'avec ma femme, ni pourquoi je *peux* vendre ma voiture ne sont des *explananda* pour une explication causale tant qu'aucun processus (divorce ou vente) ne s'est accompli.

Cependant, combinée avec mon intention de divorcer — si tant est que je forme cette intention — le fait d'avoir la propriété d'être uni à ma femme par le mariage ne pourrait-il pas contribuer à expliquer causalement pourquoi j'ai divorcé d'avec elle si jamais je le faisais? Considérons de

plus près le rôle joué par cette propriété juridique relationnelle dans l'explication causale du divorce d'un individu. Comparons deux hommes mariés ayant chacun l'intention de divorcer. L'un vit dans une communauté libérale dans laquelle le divorce est légal; l'autre vit dans une communauté religieuse dans laquelle le divorce est illégal. Le premier peut divorcer; le second ne le peut pas. Cela suggère que l'information (consignée dans son état civil) qu'un individu est marié peut sans doute entrer dans des explications causales; mais elle ne peut entrer dans des explications causales qu'à la condition d'être traitée et interprétée par une autorité ayant des croyances et un système de formation de croyances. Faute d'être traitée par le cerveau des autorités compétentes, l'information abstraite selon laquelle une personne est mariée reste causalement inerte.

L'appareil Gizmo de Dretske

Considérons l'exemple suivant de Dretske (1991). Supposons que, conformément à un manuel d'instruction, vous assemblez un appareil électromécanique — baptisé *Gizmo* par Dretske — censé avoir un comportement électrique déterminé. Vous branchez l'électricité et, au lieu du comportement attendu, l'appareil se met à fumer et à bourdonner. Vous voulez savoir pourquoi. Comme vous savez déjà que le fait que l'électricité soit branchée est un facteur pertinent dans l'explication causale du comportement de l'appareil, vous voulez en savoir davantage sur ce qui, dans le branchement interne de l'appareil, contribue à le faire fumer et bourdonner. Supposons que vous découvriez qu'un fil bleu a été branché par erreur sur le plot P_1 alors qu'il devrait être branché sur le plot P_2 . Si vous aviez branché le fil bleu sur le plot P_2 , l'appareil ne fumerait ni ne bourdonnerait plus. Or vous pourriez avoir branché le fil bleu sur le plot P_1 pour au moins deux raisons différentes: vous pourriez avoir incorrectement appliqué un manuel d'instruction correct ou vous pourriez avoir correctement appliqué un manuel d'instruction incorrect.

Imaginons maintenant deux appareils A et B microphysiquement indiscernables qui ont le même comportement dès que l'électricité est branchée. A fume et bourdonne parce que lorsque vous l'avez assemblé, vous avez incorrectement appliqué un manuel correct d'instruction. B fume et bourdonne parce que lorsque je l'ai assemblé, j'ai correctement appliqué un manuel incorrect d'instruction. Manifestement, la relation entre A ou B et le manuel d'instruction — propriété relationnelle historique des appareils — ne covarie pas avec les propriétés physiques de A ou de B . Cette relation n'est-elle pas pertinente dans l'explication causale du comportement de A ou de B ?

Pour répondre à cette question, il convient, selon moi, de scinder l'explication du comportement d'un appareil en deux étapes explicatives. Premièrement, nous pouvons nous demander pourquoi A et B fument et

bourdonnent. A cette question, il y a une seule et même réponse pour *A* et *B*: parce que le fil bleu est branché sur le plot P_1 . Cette première explication — identique pour *A* et *B* — ne fait référence qu'à des propriétés physiques des deux appareils. Deuxièmement, nous pouvons nous demander pourquoi le fil bleu est branché sur le plot P_1 . Nous obtenons alors deux explications différentes pour les deux appareils: l'un des facteurs pertinents est une erreur commise par la personne qui a assemblé l'appareil; l'autre est une erreur commise par l'auteur du manuel d'instruction. La raison pour laquelle il me paraît préférable de scinder l'explication en deux étapes tient à l'asymétrie suivante: le fait que l'électricité soit branchée est pertinent pour expliquer pourquoi l'appareil fume et bourdonne, non pour expliquer pourquoi le fil bleu est branché sur le mauvais plot.

Si la relation entre un appareil et un manuel d'instruction explique pourquoi le fil bleu est branché sur le plot P_1 , et si le fait que le fil bleu est branché sur le plot P_1 contribue à expliquer pourquoi l'appareil fume et bourdonne, cette relation n'explique-t-elle pas pourquoi l'appareil fume et bourdonne? Si c'est le cas, nous avons un exemple d'une propriété historique relationnelle d'un appareil électromécanique qui explique causalement le comportement de l'appareil sans covarier avec les propriétés physiques de l'appareil. Mais cette relation explique-t-elle causalement pourquoi le fil bleu est branché sur le plot P_1 ?

En réalité, pour expliquer causalement pourquoi le fil bleu est branché sur le plot P_1 , il ne suffit pas de faire référence à l'information contenue dans le manuel (ou à la relation entre cette information et un appareil). Si le manuel, correct ou incorrect, gît non consulté sur une étagère, l'information qu'il contient sera causalement inerte. Pour entrer dans une explication causale, l'information contenue dans le manuel doit être activement consultée par une personne ayant l'intention d'assembler l'appareil. L'explication causale dans laquelle entrera cette information doit au surplus faire référence aux intentions de celle ou de celui qui a assemblé l'appareil, à sa capacité de comprendre le manuel d'instruction, de former des croyances à partir de sa compréhension du manuel et ainsi de suite. Si le manuel d'instruction est incorrect, ce fait ne sera pertinent que *si* celle ou celui qui a assemblé l'appareil a des mécanismes corrects de formation de croyance et forme correctement une croyance qui se révélera être incorrecte. Si cette personne avait, en traitant incorrectement l'information contenue dans le manuel, formé incorrectement (et peut-être fortuitement) la croyance correcte que le fil bleu doit être branché sur le plot P_2 , et non sur le plot P_1 , elle n'aurait vraisemblablement pas branché le fil bleu sur le plot P_1 .

A la différence donc de l'explication du fait que *A* et *B* fument et bourdonnent, qui ne fait référence qu'à des propriétés physiques des deux appareils, l'explication causale du fait que le fil bleu est branché sur le plot P_1 ne fait pas référence à des propriétés physiques des appareils. Elle ne fait

pas non plus simplement référence à l'information abstraite contenue dans le manuel d'instruction; elle fait aussi référence aux intentions et aux croyances de l'auteur du manuel et aux capacités de compréhension et de formation de croyance de la personne qui a assemblé l'appareil en consultant le manuel. Manifestement, de telles propriétés ne sont pas des propriétés des appareils, mais des propriétés des cerveaux d'au moins deux personnes.

Le contre-exemple de Baker à la condition nécessaire de Fodor sur la pertinence causale

Fodor (1991) a récemment proposé une condition *nécessaire* pour qu'une propriété d'une cause soit pertinente dans une explication causale de son effet: si le fait que c ayant la propriété F est une cause du fait que e a la propriété G , alors le lien entre le fait que c soit F et le fait que e soit G doit être contingent, non conceptuel ou non analytique. Considérons une paire de causes microphysiquement indiscernables et leurs effets respectifs: c_1 ayant la propriété F_1 est une cause de e_1 ayant la propriété G_1 et c_2 ayant la propriété F_2 est une cause de e_2 ayant la propriété G_2 «seulement si ce n'est pas une vérité conceptuelle que des causes qui diffèrent en ce que l'une est F_1 et l'autre est F_2 ont des effets qui diffèrent en ce que l'un est G_1 et l'autre est G_2 » (Fodor 1991: 19).

Selon Fodor, le contenu *large*² — la condition de vérité — des croyances de deux jumeaux microphysiquement identiques (Jules et Gilles) est causalement inerte: il ne joue aucun rôle dans une explication causale car, selon lui, le comportement de chaque membre d'une paire de jumeaux microphysiquement identiques est identique au comportement de l'autre. Il n'est donc nul besoin d'invoquer des contenus *distincts* pour expliquer des actions *différentes*. Quoiqu'elles aient des contenus *larges* distincts, leurs croyances ont un seul et même contenu *étroit* qui contribue à expliquer l'identité des comportements.

Contrairement à Fodor, Baker (1991), qui nie l'existence d'un contenu *étroit*, veut défendre la thèse de la pertinence du contenu que Fodor nomme *large* dans une explication causale et entend démontrer que le contenu qualifié de *large* peut satisfaire la condition nécessaire de Fodor. Considérons avec elle

² Il y a deux façons de décrire le contenu de ce que croient respectivement Jules et Gilles en présence de leurs tasses de café perceptuellement indiscernables. On peut dire que Jules croit que t_1 contient du café et que Gilles croit que t_2 contient du café. On leur attribue des croyances ayant des contenus *larges* distincts. Mais on peut aussi attribuer à Jules et Gilles la croyance commune que la tasse placée devant chacun d'entre eux contient du café. On caractérise alors le contenu *étroit* commun de leurs croyances.

deux jumeaux microphysiques *A* et *B*. *A* vit dans une communauté franco-phone «normale» dans laquelle le mot «jade» se dit indifféremment de néphrites ou de jadéites. *B* vit dans une communauté burgienne³ contrefactuelle dans laquelle le mot “jade” s’applique exclusivement aux jadéites (qui sont en réalité des pierres plus précieuses que les néphrites). Supposons que *A* et *B*, chacun dans sa communauté respective, soient candidats à un concours télévisé — chaque concours étant qualitativement identique à l’autre. Pour gagner le grand prix, il est demandé d’identifier une pierre. Les pierres sont deux morceaux identiques de néphrite. Chaque animateur dit: «Voici une jolie pierre verte. Pouvez-vous l’identifier?» A cette question, *A* et *B* donnent deux réponses acoustiquement identiques: «C’est de la jade». A partir de cet instant, les deux histoires divergent car *A* a fourni une réponse correcte et *B* une réponse incorrecte.

A et *B* sont dans le même état cérébral ou neural type mais leurs croyances respectives ont des propriétés intentionnelles distinctes. Dans le cadre de Fodor, les croyances c_1 de *A* et c_2 de *B* sont respectivement la cause des énoncés e_1 de *A* et e_2 de *B*. c_1 a la propriété mentale M_1 d’être une croyance ayant la même condition de vérité qu’un énoncé de la phrase française «C’est de la jadéite ou de la néphrite». c_2 a la propriété mentale M_2 d’être une croyance ayant la même condition de vérité qu’un énoncé de la phrase française «C’est de la jadéite». c_1 est une cause de l’énonciation par *A* de l’énoncé e_1 de la phrase type «C’est de la jade» qui est vrai si et seulement si la pierre désignée par «c» est de la jadéite ou de la néphrite. c_2 est une cause de l’énonciation par *B* de l’énoncé e_2 de la même phrase type qui est vrai si et seulement si la pierre désignée par «c» est de la jadéite. e_1 a la propriété d’être une réponse gagnante et e_2 a la propriété d’être une réponse perdante. Baker conclut que la différence de contenu — la différence entre la propriété vériconditionnelle M_1 de c_1 et la propriété vériconditionnelle M_2 de c_2 — est pertinente pour expliquer causalement

³ Tyler Burge (1979) a imaginé une fameuse expérience de pensée destinée à rendre plausible une doctrine qu’il nomme «l’anti-individualisme» et qui affirme que le contenu des pensées d’un individu est pénétré par le (ou dépend du) contenu des pensées des membres de sa communauté. Victor vit dans une communauté francophone. Il est atteint d’arthrite aux doigts, aux poignets et aux genoux. Il a des croyances correctes sur sa propre condition. Mais un matin, il se réveille avec des douleurs à la fesse et forme la croyance erronée qu’il a de l’arthrite à la fesse. Il consulte un médecin qui le détrompe. Imaginons que Victor soit exactement dans la même condition physique interne et vive dans une communauté contrefactuelle dans laquelle le mot «arthrite» peut désigner des inflammations des tendons et des muscles. Dans cette communauté contrefactuelle, lorsque Victor consulte un médecin après avoir formé la croyance qu’il a maintenant de l’arthrite dans la fesse, le médecin corrobore sa croyance. Selon Burge, dans la communauté francophone et dans la communauté contrefactuelle, Victor ne croit pas la même chose lorsqu’il énonce en présence du médecin la phrase: «Docteur, j’ai de l’arthrite dans la fesse».

pourquoi les effets e_1 et e_2 ont respectivement les propriétés d'être une réponse gagnante et une réponse perdante.

Selon la condition nécessaire de Fodor, pour qu'une propriété d'une cause soit pertinente dans une explication causale, il faut que la relation entre la propriété de la cause et celle de son effet ne soit pas une vérité conceptuelle ou nécessaire. Pour démontrer que la relation entre le contenu de la croyance et la propriété de l'énoncé d'être une réponse gagnante (ou perdante) n'est ni conceptuelle ni nécessaire, Baker imagine un monde dans lequel les mêmes conditions linguistiques prévalent mais il n'y a pas de concours télévisé. Dans un tel monde, c_1 et c_2 gardent leurs propriétés sémantiques, mais e_1 et e_2 perdent leurs propriétés d'être respectivement une réponse gagnante et une réponse perdante.

L'argument de Baker est confronté, selon moi, au dilemme suivant. Ou bien Baker traite la propriété d'être une réponse gagnante sur le modèle de la vérité d'un énoncé ou non. Concentrons-nous sur la relation entre c_1 et e_1 .

Première branche du dilemme: elle traite la propriété d'être une réponse gagnante sur le modèle de la propriété sémantique d'un énoncé d'être une réponse correcte ou vraie. Il est plausible de supposer qu'un énoncé est vrai (ou faux) indépendamment de ce qu'en pense un juge — indépendamment du fait qu'un juge perçoive l'énoncé, l'interprète et forme la croyance que l'énoncé est vrai. Dans cette branche du dilemme, être une réponse gagnante est une propriété qu'a l'énoncé e_1 en vertu du fait qu'il est vrai. Il est tentant de penser qu'il existe un lien conceptuel entre la propriété de c_1 d'avoir la même condition de vérité qu'un énoncé de la phrase française «C'est de la jadéite ou de la néphrite» et la propriété de e_1 d'être une réponse gagnante. Mais c'est une erreur: la *condition de vérité* de la pensée (qui cause l'énoncé) n'*implique* pas la *valeur de vérité* de l'énoncé. Celle-ci dépend de la signification linguistique de la phrase énoncée et du *monde*. En revanche, si la propriété de l'énoncé d'être une réponse gagnante est assimilée à la vérité, alors le prédicat «être une réponse gagnante» n'exprime pas la propriété que le sens commun entend par ces mots, c'est-à-dire le fait de remporter le grand prix. Dans cette interprétation, être une réponse gagnante est en effet une propriété conférée à e_1 par les règles abstraites d'un jeu qui demeurent causalement inertes jusqu'à ce que le jeu soit joué en présence d'un animateur. Enoncer une réponse gagnante ne garantit pas à son auteur de remporter le grand prix. Il en va de même *mutatis mutandis* pour la relation entre la propriété de c_2 d'avoir la même condition de vérité qu'un énoncé de la phrase française «C'est de la jadéite» et la propriété de e_2 d'être une réponse perdante.

Seconde branche de l'alternative: supposons que, à la différence de la vérité d'un énoncé, être une réponse gagnante soit une propriété qu'un énoncé ne possède qu'à la condition qu'un juge tienne l'énoncé pour une réponse correcte ou vraie. Dans cette hypothèse, le fait que c_1 soit M_1 n'est

ni une condition suffisante ni même une condition nécessaire pour que e_1 soit une réponse gagnante. Pour rapporter le prix à la maison, il n'est ni nécessaire ni suffisant que c_1 soit M_1 .

Il n'est pas suffisant que e_1 soit correct ou vrai pour être une réponse gagnante. L'animateur de télévision doit au surplus former correctement la croyance correcte que e_1 est une réponse gagnante ou correcte. Si l'animateur croit incorrectement que e_1 est incorrect, alors l'énoncé ne sera pas une réponse gagnante. Selon cette branche de l'alternative donc, l'animateur doit avoir une capacité appropriée à former des croyances correctes — capacité qui est une propriété de son cerveau. Si le fait que e_1 soit une réponse gagnante dépend du jugement de l'animateur, alors il n'est même plus nécessaire que e_1 soit une réponse correcte pour être une réponse gagnante: si e_1 était une réponse incorrecte que l'animateur prenait pour une réponse correcte, e_1 serait une réponse gagnante. *Mutatis mutandis*, il en va de même de la relation entre la propriété sémantique de c_2 et la propriété de e_2 d'être une réponse perdante.

Dans la présente étude, j'ai soutenu que le sens commun erre en attribuant au contenu des attitudes propositionnelles une efficacité causale mais qu'il suppose à bon droit que le contenu joue un rôle dans les explications causales. J'ai examiné et rejeté la possibilité que le contenu entre dans des explications non causales. J'ai examiné trois propriétés relationnelles d'un système qui, à la différence d'une simple propriété fonctionnelle du système (comme la propriété d'être un antalgique d'un comprimé d'aspirine), entrent dans des explications causales du comportement du système sous réserve d'être détectées, traitées et interprétées par les activités cognitives d'un système de traitement de l'information extérieur au système dont le comportement est expliqué.

RÉFÉRENCES

- BAKER, L. R. (1991), «Belief in Explanation», mimeo, Middlebury College, lu au CREA, Paris, mars 1991.
- BLOCK, N. (1990), «Can the Mind Change the World», in G. BOOLOS (éd.) *Mind and Method, Essays in Honor of Hilary Putnam*, Londres, Cambridge University Press.
- BURGE, T. (1979), «Individualism and the Mental», in P. A. FRENCH, T. E. UEHLING & H. K. WETTSTEIN (éds), *Midwest Studies in Philosophy*, IV, Minneapolis, Minnesota UP.
- DAVIDSON, D. (1963), «Actions, Reasons and Causes», in D. DAVIDSON, *Essays on Actions and Events*, Oxford, Clarendon Press, 1980.
- DENNNETT, D. (1983), «Three Kinds of Intentional Psychology», in D. Dennett, *The Intentional Stance*, MIT Press, 1987.

- DRETSKE, F. (1988a), «The Explanatory Role of Content», in R. H. GRIMM & D. D. MERRILL (éds), *Contents of Thought*, Tucson, Arizona Press.
- DRETSKE, F. (1988b), *Explaining Behavior*, Cambridge, MIT Press.
- DRETSKE, F. (1990), «Does Meaning Matter?», in E. VILLANUEVA (éd.), *Information, Semantics & Epistemology*, Oxford, Blackwell.
- DRETSKE, F. (1991), «What Isn't Wrong With Folk Psychology», à paraître dans *Metaphilosophy*, lu au CREA, Paris, avril 1991.
- FODOR, J. A. (1991), «A Modal Argument for Narrow Content», *The Journal of Philosophy*, lxxxviii, 1, 5-26.
- GEACH, P. (1969), *God and the Soul*, Londres, Routledge & Kegan Paul.
- HORGAN, T. (1989), «Mental Quausation», in J. TOMBERLIN (éd.), *Philosophical Perspectives*, 3, «Philosophy of Mind and Action Theory», Atascadero, Ridgeview.
- HORNSBY, J. (1986), «Physicalist Thinking and Behavior», in P. PETTIT & J. McDowell (éds), *Subject, Thought and Context*, Oxford, Clarendon Press.
- JACKSON, F. & P. PETTIT (1989), «Functionalism and Broad Content», *Mind*, XCVII, 387, 381-400.
- JACKSON, F. & P. PETTIT (1990a), «Programme Explanation: a General Perspective», *Analysis*, 50, 2.
- JACKSON, F. & P. PETTIT (1990b), «In Defense of Explanatory Ecumenism», mimeo, The Australian National University.
- KIM, J. (1973), «Causes and Counterfactuals», in E. SOSA (éd.), *Causation and Conditionals*, Oxford, Oxford University Press.
- KIM, J. (1974), «Non Causal Relations», *Noûs*, 8, 41-52.
- KIM, J. (1984), «Epiphenomenal and Supervenient Causation», in P. A. FRENCH, T. E. UEHLING & H. K. WETTSTEIN (éds), *Midwest Studies in Philosophy*, IX, Minneapolis, Minnesota UP.
- KIM, J. (1989), «Mechanism, Purpose, and Explanatory Exclusion», in J. TOMBERLIN (éd.), *Philosophical Perspective*, 3, *op. cit.*
- KIM, J. (1990), «Explanatory Exclusion and Mental Causation», in E. VILLANUEVA (éd.), *Information, Semantics and Epistemology*, Oxford, Blackwell.
- LEPORE, E. & B. LOEWER (1987), «Mind Matters», *The Journal of Philosophy*, 93, 630-42.
- MCGINN, C. (1989), *Mental Content*, Oxford, Blackwell.
- PUTNAM, H. (1975a), «The Meaning of 'Meaning'», in H. PUTNAM, *Philosophical Papers*, II, Cambridge, Cambridge UP.
- PUTNAM, H. (1975b), «Philosophy and Our Mental Life», in H. PUTNAM, *Philosophical Papers*, *op. cit.*
- SOSA, E. (1984), «Mind-Body Interaction and Supervenient Causation», in P. A. FRENCH, T. E. UEHLING & H. K. WETTSTEIN (éds), *Midwest Studies in Philosophy*, IX, Minneapolis, Minnesota UP.
- STICH, S. (1978), «Autonomous Psychology and the Belief-Desire Thesis», *The Monist*, 61, 4, 573-91.
- STICH, S. (1983), *From Folk Psychology to Cognitive Science*, Cambridge, Mass., MIT Press.