

Ce que pense un individu peut-il expliquer ce qu'il fait?

Pierre Jacob

Volume 27, Number 1, Spring 2000

Le matérialisme contemporain

URI: <https://id.erudit.org/iderudit/004936ar>

DOI: <https://doi.org/10.7202/004936ar>

[See table of contents](#)

Publisher(s)

Société de philosophie du Québec

ISSN

0316-2923 (print)

1492-1391 (digital)

[Explore this journal](#)

Cite this article

Jacob, P. (2000). Ce que pense un individu peut-il expliquer ce qu'il fait? *Philosophiques*, 27(1), 115–138. <https://doi.org/10.7202/004936ar>

Article abstract

A philosopher who subscribes jointly to physicalism and to intentional realism is bound to grant content (or semantic properties) causal efficacy in the causation of physical effects. He must then overcome two problems. First, the extrinsicness of semantic properties seems to deprive them of causal power (the externalist threat). Secondly, on minimalist physicalist assumptions, no entity can instantiate a semantic property unless it simultaneously instantiates some physical properties. A representation's semantic property risks therefore to be deprived of causal efficacy by its physical properties (the problem of causal-explanatory exclusion). I deal with the first problem in the first section of this paper and with the second problem in the second section.

Ce que pense un individu peut-il expliquer ce qu'il fait?¹

PIERRE JACOB

jacob@poly.polytechnique.fr

Institut des sciences cognitives

RÉSUMÉ. — Un partisan du monisme matérialiste qui souscrit au réalisme intentionnel s'engage à octroyer aux contenus (ou aux propriétés sémantiques) des représentations mentales une efficacité causale dans la production d'effets physiques. Deux difficultés se présentent : d'une part, le caractère extrinsèque des propriétés sémantiques paraît inconciliable avec leur efficacité causale (c'est le défi de l'externalisme) ; d'autre part, comme un matérialiste suppose que toute entité qui exemplifie une propriété sémantique exemplifie simultanément des propriétés physiques, celles-ci risquent de priver celle-là d'efficacité causale (c'est la menace de l'exclusion explicative). La première section est consacrée au premier problème, et la seconde section au second problème.

ABSTRACT. — A philosopher who subscribes jointly to physicalism and to intentional realism is bound to grant content (or semantic properties) causal efficacy in the causation of physical effects. He must then overcome two problems. First, the extrinsicness of semantic properties seems to deprive them of causal power (the externalist threat). Secondly, on minimalist physicalist assumptions, no entity can instantiate a semantic property unless it simultaneously instantiates some physical properties. A representation's semantic property risks therefore to be deprived of causal efficacy by its physical properties (the problem of causal-explanatory exclusion). I deal with the first problem in the first section of this paper and with the second problem in the second section.

1. Introduction

La psychologie partage avec les autres sciences (dont la cosmologie, l'astronomie, la chimie, la biologie moléculaire ou la géologie) l'ambition de fournir des explications causales. On soutient généralement que les relations causales unissent des événements : l'événement *c* produit l'événement *e*². Dans les explications non psychologiques, on dit volontiers qu'un court-circuit a provoqué un incendie ou que la collision entre une météorite et la Terre a provoqué l'extinction des dinosaures. Toute explication n'est pas causale³.

1. Merci à deux lecteurs anonymes de la revue pour leurs remarques.

2. Soit que comme Davidson, 1967, on tienne la relation causale pour une relation purement extensionnelle entre deux individus, soit que comme Kim, 1973, on la tienne pour une relation intensionnelle entre deux exemplifications de propriétés à deux instants. Dans la suite, je supposerai qu'on peut faire référence à une cause et à un effet — les *relata* de la relation causale — sans mentionner aucune de leurs propriétés respectives. Mais je supposerai que dans une explication causale, on ne peut mentionner les faits constitutifs de l'*explanans* et de l'*explanandum* sans faire référence à certaines propriétés de la cause et de l'effet.

3. On peut, par exemple, tenir la démonstration d'un théorème mathématique à partir d'axiomes pour une explication de la vérité du théorème. On peut également tenir la traduction

Mais toute explication causale ou non causale relie des faits. Pour qu'une explication soit causale, il faut, d'une part, qu'il existe une relation causale entre l'événement mentionné dans l'*explanandum* (la cause *c*) et l'événement mentionné dans l'*explanans* (l'événement *e*). Il faut, d'autre part, que le fait que *c* exemplifie la propriété *F* soit causalement responsable du fait que *e* exemplifie la propriété *G*. Nombre de philosophes supposent de surcroît que le fait que *c* est *F* ne peut être causalement responsable du fait que *e* est *G* à moins qu'il existe une corrélation nomique entre les exemplifications de la propriété *F* et les exemplifications de la propriété *G*⁴.

À la différence d'une explication non psychologique, une explication psychologique fait souvent (sinon toujours) référence au contenu (ou à la propriété sémantique) d'au moins une représentation mentale d'un individu⁵. Dans le domaine psychologique, les relations causales pertinentes tiennent tantôt entre deux représentations mentales — lorsqu'une pensée produit une autre pensée —, tantôt entre une représentation mentale et un geste corporel.

Il existe une distinction entre ce qui arrive à un individu et ce qu'il fait (c'est-à-dire son comportement) : si un kinésithérapeute prend ma tête entre ses mains et lui fait subir une rotation de 90° vers la droite, le mouvement de ma tête est quelque chose qui m'arrive. Si j'effectue spontanément une rotation de la tête de 90° vers la droite, ce mouvement de ma tête est quelque chose que je fais. Cette distinction se fonde sur le contraste entre une cause interne et une cause externe du mouvement corporel : dans le second cas, mais non dans le premier, la cause de la rotation de ma tête est dans mon cerveau. Parmi les comportements d'un individu, certains sont intentionnels et d'autres ne le sont pas : sont intentionnels les comportements impliquant des gestes corporels guidés par des états internes doués de contenus représentationnels. Lorsque je ronfle et lorsque je vomis, j'émetts des sons et je contracte mes muscles abdominaux. Mais ni les sons que j'émetts ni mes contractions abdominales n'ont pour causes des représentations internes douées de contenus ou de propriétés sémantiques. Un comportement intentionnel peut être volontaire ou involontaire selon que le contenu de la représentation mentale qui guide le geste corporel est conceptuel ou non conceptuel : un geste corporel guidé par le contenu conceptuel d'une croyance fait partie d'un comportement intentionnel volontaire ; un geste corporel guidé par le contenu

d'une expression *e* d'une langue *L* par une expression *e'* appartenant à une langue *L'* différente de *L* pour une explication de la signification de la première expression. Dans ce cas, la démonstration du théorème et la traduction seront des explications non causales.

4. La corrélation nomique sera généralement supposée admettre des exceptions.

5. Je suppose que les croyances d'un individu sont des représentations mentales. Toute représentation n'est pas mentale : un énoncé, une photographie ou une carte géographique sont des représentations non mentales. Je suppose qu'une représentation mentale est un état physique du cerveau d'un individu. À la différence d'une représentation non mentale, une représentation mentale a des propriétés neurophysiologiques.

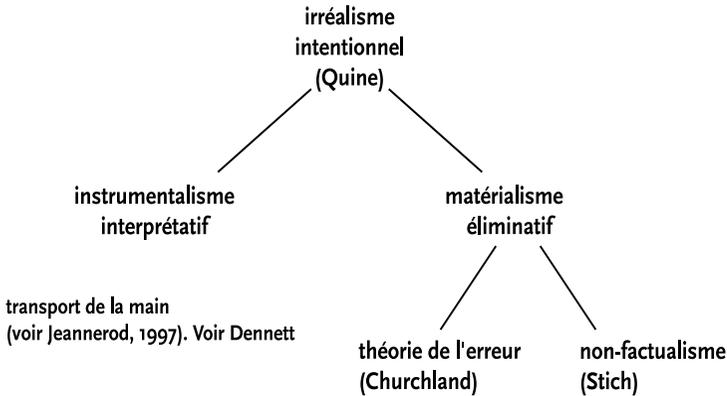
non conceptuel d'une représentation sensorielle fait partie d'un comportement intentionnel non volontaire⁶.

Imaginons que je transporte ma main droite à quarante centimètres de l'axe de mon buste et que, pendant le transport, j'ajuste (non consciemment) l'écart entre mon pouce et mon majeur au diamètre d'un verre contenant de l'eau ; je saisis le verre et bois une gorgée d'eau. La séquence ordonnée des gestes moteurs de mon bras et de ma main fait partie d'un comportement intentionnel volontaire⁷ : mes gestes ont pour causes ma croyance que le verre placé devant moi contient de l'eau et mon désir d'une gorgée d'eau.

Supposons que ma croyance *c* — une représentation mentale — est un état physique de mon cerveau (ou de mon système nerveux central) : l'exemplaire particulier de ma croyance qui a contribué à produire causalement la série de mes gestes corporels *m* n'est autre qu'un état concret de mon cerveau. Ma croyance *c* exemplifie donc plusieurs propriétés non sémantiques — des propriétés physiques, chimiques et neurophysiologiques. Ma croyance *c* a aussi un contenu (ou une propriété sémantique) : elle représente le fait que le verre placé devant moi contient de l'eau. La propriété sémantique de ma croyance possède-t-elle une efficacité causale dans le processus au cours duquel j'ai accompli la série de gestes corporels que j'ai décrits? Un matérialiste peut incontestablement tenir une représentation mentale pour la cause d'effets physiques grâce à l'identification entre un exemplaire concret de la représentation mentale et un état cérébral concret de l'individu. Mais pour assurer le transport de ma main droite à quarante centimètres de l'axe de mon buste et pour refermer la pince de mes doigts efficacement sur deux points opposés de la surface du verre, d'innombrables signaux électriques et chimiques ont dû être envoyés par mon cerveau le long des axones et des fibres nerveuses de mon

6. Malgré les critiques qui lui ont été adressées par McDowell, 1994, j'admets la distinction entre le contenu conceptuel et le contenu non conceptuel d'une représentation. La typologie des comportements de ce paragraphe est fidèle à la conception développée par Dretske, 1988, à ceci près que Dretske, 1988, ne reconnaît pas la distinction entre les comportements intentionnels volontaires et involontaires. Pour lui, un geste corporel produit par une représentation sensorielle douée d'un contenu non conceptuel fait partie d'un comportement instinctif non intentionnel. Selon lui, lorsqu'une représentation sensorielle contribue à la production d'un geste corporel, et quoique la représentation soit douée d'un contenu non conceptuel, ce contenu est néanmoins dépourvu d'efficacité causale dans la production du geste. Dans le chapitre VIII de Jacob, 1997, et dans Jacob, 1998, j'ai expliqué en détail pourquoi, contrairement à Dretske, j'admets que le contenu non conceptuel d'une représentation sensorielle peut être causalement efficace dans la production d'un comportement intentionnel non volontaire.

7. Ce qui ne veut pas dire que chaque étape motrice est planifiée consciemment. Les neurosciences cognitives montrent notamment que, dans la saisie d'un objet perçu visuellement, l'écart entre le pouce et l'index est calculé automatiquement pendant le transport de la main (voir Jeannerod, 1997).



système nerveux central aux muscles de ma main et de mon bras droits afin de commander leur contraction et leur décontraction successives.

La plupart des philosophes matérialistes qui souscrivent à ce qu'on peut appeler l'« irréalisme intentionnel » sont enclins à conférer aux propriétés non sémantiques (physiques, chimiques et neurophysiologiques) de mon état cérébral une efficacité causale dans ce processus et ils sont prêts à renoncer à la thèse de l'efficacité causale des propriétés sémantiques des représentations mentales. On peut représenter ainsi les principales versions de l'irréalisme intentionnel :

Pour un partisan de l'instrumentalisme interprétatif⁸, en attribuant une attitude propositionnelle douée d'une propriété sémantique, on peut prédire mais non expliquer causalement le comportement d'un dispositif physique. Pour un partisan de la théorie de l'erreur⁹, toute attribution d'attitude propositionnelle douée de propriété sémantique à un dispositif physique est fautive. Pour un partisan du non-factualisme¹⁰, une telle attribution est dénuée de condition de vérité¹¹.

Mais un matérialiste qui souscrit au réalisme intentionnel ne peut admettre la thèse de l'inefficacité causale des propriétés sémantiques des représentations mentales. Il doit leur reconnaître une efficacité causale dans le processus comportemental par lequel la représentation mentale concourt à la production d'un geste corporel. Le test de l'authenticité (ou de la réalité) d'une propriété, c'est son efficacité causale. Ce qui atteste la réalité d'une propriété, c'est sa capacité à produire des effets décelables : si la propriété sémantique d'une représentation mentale est une propriété authentique, alors elle doit concourir à produire des effets dans le processus comporte-

8. Voir Dennett, 1971, 1987.

9. Voir Churchland, 1981.

10. Voir Stich, 1983.

11. Pour une discussion détaillée des versions de l'irréalisme intentionnel, voir Jacob, 1997, ch. I.

mental. Dans la première partie de ce texte, je vais examiner un argument sur lequel s'appuient les adversaires matérialistes du réalisme intentionnel pour rejeter la thèse de l'efficacité causale des propriétés sémantiques des représentations mentales. Ayant indiqué la voie que, selon moi, doit suivre un matérialiste qui veut souscrire au réalisme intentionnel, j'examinerai dans la dernière partie le dilemme auquel se trouve confronté, selon Kim¹², le matérialiste qui souscrit au réalisme intentionnel.

2. La menace externaliste sur l'efficacité causale du contenu mental

On peut reconstruire le raisonnement qui conduit des philosophes matérialistes à adopter l'une des versions de l'irréalisme intentionnel de la manière suivante.

L'argument en faveur de l'irréalisme intentionnel :

1) Une explication causale doit faire référence à une propriété causalement efficace de la cause *c* dans le processus par lequel elle a produit son effet *e*.

2) Les seules propriétés causalement efficaces d'une cause *c* dans la production de son effet *e* sont des propriétés intrinsèques de *c* ou des propriétés de *c* qui dépendent systématiquement de propriétés sous-jacentes intrinsèques de *c*¹³.

3) Or, le contenu est une propriété historique et extrinsèque d'une représentation (mentale ou non mentale) qui ne dépend pas systématiquement de propriétés sous-jacentes intrinsèques de la représentation.

4) Donc : le contenu d'une représentation *R* d'un individu n'est pas une propriété causalement efficace dans le processus par lequel *R* contribue à produire causalement un comportement intentionnel de l'individu.

J'accepte les trois prémisses mais je rejette la conclusion (4). Je prétends que (4) n'est pas une conséquence des prémisses (1)-(3). Mais avant d'expliquer pourquoi je maintiens que (4) ne suit pas des prémisses (1)-(3), je vais d'abord motiver mon adhésion à la vérité des prémisses (1)-(3). Autrement dit, je vais préciser la nature de la menace externaliste sur la thèse de l'efficacité causale des propriétés sémantiques des représentations mentales.

12. Voir Kim, 1993a.

13. Par « dépendance systématique », j'entends ce que les anglophones nomment *supervenience* : une propriété *B* dépend systématiquement d'une propriété sous-jacente *A* si, nécessairement, chaque fois que *B* est exemplifiée, *B* est exemplifiée et (nécessairement) si *B* est exemplifiée, alors *A* est exemplifiée. Je mets la seconde occurrence de l'adverbe de modalité « nécessairement » entre parenthèses, car je ne veux pas trancher entre une interprétation faible ou forte de la notion de dépendance systématique requise ici. En qualifiant les propriétés physiques sous-jacentes d'« intrinsèques », je ne considère que les versions locales de la dépendance systématique et j'écarte l'interprétation globale. Voir Kim, 1987.

La vérité des prémisses (1) et (2) repose sur un principe que je tiens pour hautement plausible et qu'on pourrait appeler le « principe de l'inéquité de l'efficacité causale des propriétés d'une cause » et selon lequel l'efficacité causale n'est pas distribuée équitablement parmi les propriétés d'une cause. Pour corroborer ce principe, je gloserai un célèbre exemple dû à Dretske¹⁴.

a) Le sens du son

En chantant la Traviata, la Callas émet un son aigu c à l'instant t . L'émission du son à l'instant t brise le verre d'une vitre à l'instant $t + 1$. Appelons c l'émission du son à l'instant t et e le bris du verre à $t + 1$. c produit e . c exemplifie, d'une part, une propriété acoustique K et, d'autre part, un sens (ou un contenu) F . Posons la question suivante : c a-t-il produit e en vertu du fait qu'il exemplifiait la propriété F ? Le fait que c était F est-il causalement efficace dans le processus par lequel c a produit e ? Intuitivement, la réponse semble clairement négative. Si le son émis c avait eu un autre sens F^* (ou si c avait été dénué de sens), c aurait néanmoins pu produire e . La propriété causalement efficace de c dans le processus par lequel la vitre a été brisée (et e est advenu) est la propriété acoustique K : c'est parce que c exemplifiait la propriété acoustique K que c a produit e . Si c n'avait pas exemplifié K , c aurait pu ne pas produire e — le son aurait pu ne pas briser la vitre. La raison pour laquelle le fait que c était K a eu une efficacité causale dans le processus par lequel la vitre a été brisée est qu'il existe une corrélation nomique entre la fragilité du verre de la vitre et la propriété acoustique K du son et que la relation causale entre c et e peut être subsumée sous cette corrélation nomique. En revanche, il n'existe pas de corrélation nomique entre la fragilité du verre de la vitre et le fait que c exemplifie la propriété sémantique F . La propriété causalement efficace du son c est sa propriété nomique — la propriété grâce à laquelle l'interaction causale singulière entre le son et le verre de la vitre exemplifie une corrélation nomique.

Premièrement, cet exemple corrobore le principe de l'inéquité de l'efficacité causale des propriétés d'une cause présupposé par la première prémisse de l'argument en faveur de l'irréalisme intentionnel. Deuxièmement, cet exemple corrobore la vérité de la prémisse (2) de l'argument pour l'irréalisme intentionnel. Cette prémisse se sert du contraste entre les propriétés intrinsèques et les propriétés extrinsèques d'une entité x . Quoique ce contraste soit difficile à définir avec précision, il n'en est pas moins utile. Intuitivement, une propriété d'une entité x peut être qualifiée d'« intrinsèque » si son exemplification à l'instant t ne dépend de rien d'autre que de x à l'instant t . Une propriété est extrinsèque si elle n'est pas intrinsèque. En un sens, contrairement à la propriété acoustique K du son c , le sens (ou la propriété sémantique) F n'est pas une propriété intrinsèque de c : en effet, le son

14. Voir Dretske, 1988, p. 79.

c n'aurait pas été le même son s'il n'exemplifiait pas la propriété *K*. Mais le son *c* pourrait demeurer le même son s'il n'avait pas la propriété sémantique *F*. Certes, on pourrait faire valoir que l'événement consistant à produire le son *c* ne serait pas le même événement si la propriété sémantique *F* faisait défaut au son *c*. Mais il y a un sens où deux émissions sonores pourraient être deux exemplaires (ou deux occurrences) du même son-type même si l'une des deux émissions seulement exemplifiait la propriété sémantique *F*. En revanche, deux émissions sonores ne pourraient pas être deux exemplaires du même son-type à moins qu'elles n'exemplifient toutes les deux la même propriété acoustique *K*. Un son qui aurait la même propriété sémantique *F* et exemplifierait une propriété acoustique *K** distincte de *K* ne serait pas le même son. Contrairement à sa propriété acoustique *K*, le son tient sa propriété sémantique *F* des intentions du compositeur, Giuseppe Verdi. Certes, le son a été choisi par Verdi. Mais contrairement au fait que le son exemplifie le sens *F*, le fait que le son exemplifie la propriété acoustique *K* est indépendant des intentions de Verdi. Verdi a donné au son son sens. Il n'a pas créé la gamme de toutes pièces. La propriété sémantique *F* du son est donc une propriété historique extrinsèque du son.

Selon la distinction entre l'intentionnalité primitive et l'intentionnalité dérivée (à laquelle souscrivent les partisans du réalisme intentionnel), il existe évidemment une différence entre le contenu (linguistique) d'un son émis par une chanteuse et le contenu d'une représentation mentale (ou d'un état cérébral) d'un être humain. Le second est primitif. Le premier est dérivé du second¹⁵. D'une manière générale, la signification d'une représentation non mentale dépend du contenu d'une représentation mentale. Mais le contenu mental lui aussi est une propriété historique et extrinsèque d'un état cérébral de l'individu. Il dépend de certaines corrélations entre le cerveau de l'individu et les propriétés exemplifiées dans son environnement¹⁶. Ce qui nous conduit tout droit à la prémisse (3) de l'argument en faveur de l'irréalisme intentionnel. Pour corroborer la prémisse (3), j'emprunterai un exemple à Dennett, que j'exploiterai dans un style dretskéen¹⁷.

b) Le distributeur automatique de boissons

Soit un distributeur automatique de boissons *S* situé en France. Il émet une boisson non alcoolisée chaque fois qu'il reçoit une pièce de cinq francs. Appelons *m* l'émission d'une boisson : *S* produit *m* grâce à l'ouverture d'un clapet. Pour simplifier, je traiterai la pièce de cinq francs *c* insérée dans la fente du dispositif *S* comme la cause de *m*, l'émission d'une boisson.

15. Pour une critique de cette distinction, voir Dennett, 1987, ch. 8. Pour une critique des critiques de Dennett, voir Jacob, 1997, ch. I.

16. Je ne fais ici qu'ébruiter mon adhésion à une composante informationnelle dans la conception du contenu mental.

17. Voir Dennett, 1987, ch. 8 et Dretske, 1998.

Aussi, c a des propriétés physiques intrinsèques P (une masse, une forme géométrique, une composition physique caractéristique, etc.). Et c a aussi une valeur monétaire V . V est manifestement une propriété historique extrinsèque de c parce que ce qui confère V à c , c'est le fait historique d'avoir été estampillée par une personne travaillant pour la Banque de France, et non par une autre personne — en particulier, non par un faussaire. Ce qui confère V à c , c'est le fait d'être le relatum d'une relation historique appropriée avec un agent de la Banque de France. Supposons qu'un disque métallique c^* soit physiquement et chimiquement indiscernable de c . Si, contrairement à c , c^* n'a pas bénéficié d'une relation historique et extrinsèque avec un agent de la Banque de France, alors en dépit du fait que c^* exemplifie les propriétés physiques intrinsèques P , V fait pourtant défaut à c^* . Dire de V que c'est une propriété historique et extrinsèque de c , c'est donc dire que le fait que c est V ne dépend pas systématiquement du fait que c exemplifie la propriété physique intrinsèque sous-jacente P ¹⁸.

L'insertion de c à l'instant t a incontestablement produit m (l'émission d'une boisson) à l'instant $t + 1$. L'insertion d'un objet physique qui n'aurait pas les propriétés physiques intrinsèques d'une pièce de cinq francs dans la fente de S ne provoquerait pas m : S n'émettrait pas de boisson. Mais posons à présent la question suivante : la propriété V de c est-elle causalement efficace dans le processus par lequel S produit m ? Le fait que c est V est-il causalement responsable du fait qu'un clapet dans S libère une boisson?

Intuitivement, la réponse est : Non. Le fait que c est V n'est pas responsable du fait que S produit m . Si c contribue à produire m , c'est parce que c est P et non pas parce que c est V . Ce n'est pas parce que S « croit » avoir reçu une pièce de cinq francs qu'il émet une boisson ; c'est parce qu'il détecte un disque métallique ayant la propriété physique intrinsèque P . L'ouverture du clapet dans S est reliée nomologiquement à la propriété P de c , non à sa propriété V . Le distributeur automatique de boissons ne détecte pas la valeur monétaire V ; il réagit aux propriétés physiques des objets insérés dans sa fente. La propriété V de c est donc « épiphénoménale » dans le mécanisme de production de m par c . Si la valeur monétaire V du disque métallique inséré dans la fente du distributeur est un modèle du contenu mental, alors ce que je crois quand je crois que le verre placé devant moi contient de l'eau n'explique pas du tout ce que je fais lorsque je transporte ma main droite au voisinage du verre. Les prémisses (2)-(3) de l'argument en faveur de l'irréalisme intentionnel semblent donc acceptables. La conclusion (4) n'est-elle pas inévitable? Comment puis-je accepter les prémisses et rejeter la conclusion?

18. Du moins, c'est dire que V ne dépend pas systématiquement de P dans la version forte de la dépendance systématique, c'est-à-dire qu'il n'est pas vrai que dans tous les mondes possibles, si x est P , alors x est V .

Le distributeur automatique de boissons est un dispositif physique qui admet en « entrée » des disques métalliques ayant la propriété physique intrinsèque P et la propriété extrinsèque V et répond en émettant une boisson. La question que nous nous sommes posée était la suivante : le fait que c est V est-il causalement responsable dans le mécanisme par lequel un clapet s'ouvre dans S lorsque c est insérée dans la fente de S et S émet une boisson? La propriété V de c est-elle causalement efficace dans le mécanisme grâce auquel S produit m (l'émission d'une boisson)? Nous venons de conclure que la réponse à cette question est négative. Appelons cette question la question 1. Je maintiens que la réponse à la question 1 n'a aucune raison de mentionner la propriété V de c . Autrement dit, la propriété de c qui explique la production du mouvement physique m est la propriété P .

Or, cette question porte sur le mécanisme de la production du mouvement physique, m ¹⁹. Comprendre le mécanisme de production du mouvement physique m , n'est-ce pas expliquer le comportement du distributeur de boissons? Suivant les indications de Dretske²⁰, je voudrais suggérer une réponse négative à cette question en considérant une question différente de la question 1. Appelons cette question la question 2 : pourquoi chaque fois qu'un disque métallique ayant la propriété P est inséré dans la fente de S , S produit-il m , l'émission d'une boisson? Pourquoi l'insertion d'un disque métallique ayant la propriété P dans une fente de S produit-elle régulièrement m , la chute d'une boisson? En un mot, pourquoi le comportement de S est-il constitué par (ou composé de) la coordination entre c et m ?

En adoptant la conception « componentielle » du comportement défendue par Dretske²¹, disons que le comportement de S est le processus par lequel (ou au cours duquel) l'insertion d'un disque métallique c produit l'émission d'une boisson, m . Selon cette conception, ce serait une confusion conceptuelle que de confondre la « sortie » m avec le comportement. L'émission d'une boisson, m , est l'un des constituants du comportement de S . Ce n'est pas le comportement lui-même²². L'insertion du disque métallique c est l'autre constituant du comportement de S . Ces deux constituants sont unis par la relation causale : l'insertion de c à l'instant t est la cause de m , l'émission d'une boisson à l'instant $t + 1$. L'insertion de c à l'instant t n'est donc pas, contrairement à ce qu'affirme la conclusion (4) de l'argument en faveur de l'irréalisme intentionnel, une cause du comportement de S . C'est un constituant du comportement de S . C'est aussi une cause du second constituant du comportement de S : m , l'émission d'une boisson. Voilà pourquoi je peux

19. L'émission d'une boisson, m , est, pour le distributeur, ce qu'est le geste corporel pour un être humain.

20. Voir Dretske, 1988, ch. 1-2.

21. Voir Dretske, 1988, ch. 1-2, et Jacob, 1997, ch. VIII.

22. Contrairement à la conception du comportement qu'on trouve dans le fonctionnalisme et dans le monisme anomal, dans la conception componentielle, le geste corporel n'est pas identique au comportement.

admettre la vérité des trois prémisses de l'argument en faveur de l'irréalisme intentionnel sans accepter sa conclusion (4). Appelons (4*) la conclusion qui est une conséquence valide des prémisses (1)-(3) de l'argument en faveur de l'irréalisme intentionnel. La conclusion (4*) est la suivante :

(4*) Le contenu d'une représentation d'un individu n'est pas une propriété causalement efficace dans le processus par lequel une représentation contribue à produire un mouvement corporel de l'individu.

Ce que démontre l'exemple du distributeur automatique de boissons, c'est, comme l'affirme la conclusion (4*), que la propriété extrinsèque *V* du disque métallique *c* n'est pas causalement efficace dans la production de *m*, l'émission d'une boisson. La valeur économique « authentique » de la pièce de cinq francs n'est pas causalement efficace dans le processus d'ouverture du clapet dans *S*. Seule l'exemplification de la propriété physique intrinsèque *P* par le disque métallique *c* est causalement efficace dans ce mécanisme.

Comme j'ai soutenu que le comportement est un processus qui ne s'identifie à aucun de ses constituants et comme le mouvement physique (ou le geste corporel) *m* est l'un de ces constituants, je vais à présent faire valoir qu'il existe un sens dans lequel l'explication causale du comportement n'est pas la même chose que l'explication de la production du mouvement physique *m*. En réponse à la question 1 sur le mécanisme, j'ai fait valoir que seule la propriété physique intrinsèque *P* de *c* était causalement efficace pour assurer l'ouverture du clapet dans *S*. Mais la question 2 est distincte de la question 1. En posant la question 2, on demande : pourquoi chaque insertion d'un disque métallique *c* dans la fente de *S* provoque-t-elle régulièrement *m*, la chute d'une boisson (grâce à l'ouverture du clapet)? Pourquoi le distributeur automatique est-il un dispositif physique dont le comportement est constitué par la coordination entre deux événements : l'insertion de *c* et la chute d'une boisson?

Contrairement à la question 1, la question 2 est une question historique : poser la question 2, c'est se demander quel processus a donné naissance à la coordination entre *c* et *m* dans *S*. La réponse à la question 2 aura la forme générale suivante : comme *S* est un artefact fabriqué par un ingénieur humain, la coordination entre *c* et *m* a été conçue et réalisée par un agent intentionnel lui-même doué d'attitudes propositionnelles²³. De surcroît, je veux à présent suggérer qu'il existe un sens dans lequel le fait que le disque métallique *c* possède une valeur monétaire (la propriété extrinsèque *V*) contribue à répondre à la question 2, c'est-à-dire à expliquer la structure du comportement de *S*. En un mot, le fait que *c* est *V* contribue à expliquer la coordination entre *c* et *m* dans *S*.

23. Les activités par lesquelles un ingénieur a assuré la coordination entre *c* et *m* sont ce que Dretske, 1988, nomme les « causes structurantes » du comportement du distributeur.

Supposons en effet que, grâce aux efforts des agents de la Banque de France, il existe une corrélation fiable entre les exemplifications de la valeur monétaire V et les exemplifications des propriétés physiques par ces disques métalliques que sont les pièces de cinq francs²⁴. Compte tenu, d'une part, du prix des boissons et, d'autre part, de la corrélation fiable entre la valeur monétaire V et les propriétés physiques intrinsèques P de certains objets métalliques, les objets métalliques ayant la propriété P ont été choisis, recrutés ou sélectionnés par un ingénieur humain pour être régulièrement les causes de m parce que ces objets métalliques exemplifient V . C'est bien parce que les disques métalliques sont P qu'ils provoquent l'ouverture du clapet dans S . Mais c'est parce qu'un objet qui est P est V qu'il a été choisi par un ingénieur humain comme cause de l'ouverture du clapet dans S , c'est-à-dire comme cause de m . Sans doute la corrélation entre P et V n'est-elle pas parfaite : à l'occasion, l'insertion d'une « fausse » pièce de cinq francs — un disque métallique qui est P mais non V — provoque m , l'émission d'une boisson. Mais vraisemblablement, le bénéfice qu'entraînerait une amélioration de la sensibilité du distributeur ne compense pas le coût d'un tel effort.

Supposons à présent qu'à la date t , la Banque de France décide de changer les propriétés intrinsèques des pièces de cinq francs. Désormais, il n'y a plus de corrélation fiable entre les exemplifications de la propriété P et les exemplifications de la propriété V : après t , ce ne sont plus les disques métalliques ayant la propriété P qui exemplifient la propriété V ; les disques métalliques ayant la propriété V ont dorénavant la propriété physique intrinsèque P^* . Dans cette hypothèse et en supposant que le prix des boissons non alcoolisées reste constant, l'insertion d'un disque métallique c ayant la propriété intrinsèque P dans la fente de S ne contribue plus à la production de m . Avant la décision de la Banque de France, l'ingénieur humain avait sélectionné des objets métalliques c ayant la propriété P comme cause de m . Après la décision de la Banque de France — en admettant la fixité du prix des boissons —, les ingénieurs français doivent fabriquer de nouveaux distributeurs automatiques : ils sélectionnent désormais des objets métalliques c^* ayant la propriété P^* (différente de P) comme cause de m .

Dans cette hypothèse, on peut donc dire que de nouveaux disques métalliques c^* ayant certaines propriétés physiques intrinsèques P^* ont été choisis pour être introduites dans la fente d'un nouveau distributeur S^* parce que ces disques exemplifient la propriété extrinsèque V . Par parité, c'est parce que les disques c ayant la propriété P exemplifiaient la propriété V qu'ils avaient été choisis initialement comme cause de m dans S . C'est donc parce que certains objets métalliques ont la propriété V qu'ils ont été sélectionnés dans le processus de production de m . Si les objets métalliques c ayant la propriété intrinsèque P n'exemplifiaient pas en général

24. La corrélation fiable entre les exemplifications de P et de V correspond à l'interprétation faible de la dépendance systématique entre V et P (au sens de Kim).

la propriété V , ces objets n'auraient pas été sélectionnés comme cause de m . Le fait qu'un disque métallique exemplifie la propriété extrinsèque V contribue à expliquer la coordination entre c et m dans le comportement de S . L'exemplification de la propriété extrinsèque V contribue donc à expliquer la structure du comportement de S .

c) La proie

Dans le cas du distributeur automatique de boissons, l'explication historique de la structure du comportement (ou de la coordination entre c et m) fait référence à un processus intentionnel de sélection : c'est un agent humain doué de désirs et de croyances qui a assuré cette coordination. Je veux à présent considérer brièvement le cas hypothétique dans lequel une représentation sensorielle d'un animal serait recrutée comme cause de mouvements corporels par un processus non intentionnel de sélection : la sélection naturelle à l'œuvre dans l'évolution des espèces. Soit une proie B (pour biche). B exécute des mouvements de fuite m à l'instant t . Je supposerai que le comportement de fuite de la proie B est un comportement intentionnel non volontaire. Nous pouvons poser au moins trois questions distinctes au sujet des mouvements de fuite m de B .

Premièrement, nous pouvons demander : pourquoi B accomplit-elle m à l'instant t ? À cette question, on répondra : parce qu'à l'instant $t - 1$, B a perçu le prédateur P dans le voisinage. Appelons c l'état interne grâce auquel B a détecté P . La présence du prédateur P est la cause de l'occurrence de l'état interne c de B . Comme l'état c de B est lui-même la cause du mouvement de fuite m , Dretske qualifie la présence du prédateur de « cause déclenchante » du comportement de fuite de B ²⁵.

Deuxièmement, nous pouvons poser la question non historique suivante : quel est le mécanisme interne grâce auquel l'occurrence de l'état interne c dans B produit les mouvements corporels m ? C'est l'équivalent de la question 1 dans le cas du comportement du distributeur automatique de boissons. À cette question, la réponse pertinente sera la réponse suivante : les neurones moteurs dans le cortex moteur de B sont reliés aux neurones sensoriels du système visuel (ou olfactif) de B . Lorsque les neurones sensoriels déchargent, ils envoient un signal aux neurones moteurs. La décharge de ces derniers gouverne la contraction des muscles des membres inférieurs de B .

Enfin, nous pouvons poser la question historique suivante : comment se fait-il que chaque fois que B est dans l'état c , B produit m ? Pourquoi l'état c de B a-t-il été recruté comme cause de m , des mouvements de fuite, plutôt que comme cause d'autre chose? Lorsque B est dans l'état c , B s'enfuit. Pourquoi B ne combat-elle pas son prédateur? Ou pourquoi ne reste-t-elle pas immobile? La réponse à cette question historique doit venir en

25. Voir Dretske, 1988.

deux étapes complémentaires. Premièrement, *c* a été recruté comme cause de *m* (des mouvements de fuite) par la sélection naturelle qui, comme je l'ai dit, est un processus sélectif non intentionnel. Dans la compétition pour la survie entre des ancêtres de la proie *B* et d'autres animaux qui étaient aussi les proies des ancêtres du prédateur *P*, la sélection naturelle a favorisé les ancêtres de la proie *B* qui avaient une coordination entre *c* et *m*. Les proies des ancêtres de *P* qui étaient dépourvues de cette coordination entre *c* et *m* n'ont pas survécu à la sélection et furent dépourvus de descendance²⁶. Deuxièmement, la sélection naturelle n'aurait pas recruté *c* comme cause de *m* chez les ancêtres de *B* si *c* n'avait pas été un indicateur fiable de la présence du prédateur *P*. La propriété d'être un indicateur fiable n'est sans doute pas une propriété pleinement sémantique ; c'est une propriété proto-sémantique. En tout cas, c'est une propriété historique et extrinsèque d'un état interne de la proie qui suppose l'existence de corrélations nomiques entre des états du cerveau de la proie et l'exemplification de la propriété d'être un prédateur dans son environnement ou d'une propriété coextensionnelle avec cette propriété²⁷.

Un matérialiste peut donc souscrire au réalisme intentionnel et tenir la propriété sémantique d'une représentation mentale pour causalement efficace dans le processus comportemental par lequel la représentation contribue à produire un geste corporel. Il le peut, à condition d'adopter conjointement la conception componentielle du comportement et ce que, suivant Kim, j'ai appelé ailleurs « la stratégie de la dualité des *explananda*²⁸ », c'est-à-dire à condition de distinguer les questions 1 et 2 : la question du mécanisme de la production d'un mouvement physique et la question de la sélection de la coordination grâce à laquelle un état interne est recruté comme cause du mouvement physique²⁹.

26. Suivant Sober, 1984, Dretske, 1988, ch. 4, admet une distinction tranchée entre l'explication causale du comportement intentionnel d'un individu et une explication sélectionniste de la distribution d'un trait dans une population. Dans Jacob, 1997, ch. VIII, section 6, et dans Jacob, 1998, j'ai longuement expliqué pourquoi, à la suite de Neander, 1988, 1995, je pense qu'une explication historique de la coordination entre un état interne *c* et un mouvement physique *m* est conjointement une explication causale et une explication sélectionniste.

27. Pour le partisan que je suis d'une conception téléosémantique à base informationnelle du contenu mental — contenu conceptuel et contenu non conceptuel —, le fait qu'un état interne soit un indicateur fiable est une condition nécessaire mais non suffisante du fait que cet état est une représentation mentale douée d'une propriété sémantique authentique. Voir Jacob, 1997, ch. IV et VIII.

28. Voir Kim, 1990, 1991, et Jacob, 1997, ch. VIII.

29. Comme me l'a fait remarquer à juste titre un lecteur de la revue, je n'ai considéré que le rôle causal du contenu d'une représentation sensorielle innée, et non le contenu d'une représentation conceptuelle acquise. Comme je l'indique à la note 5, à la différence de Dretske, je soutiens que le contenu d'une représentation innée peut contribuer à expliquer un comportement intentionnel non volontaire. Pour la contribution du contenu d'une représentation conceptuelle acquise au comportement intentionnel volontaire, voir Jacob, 1997, ch. VIII, et Jacob, 1998.

3. Le réalisme intentionnel face à la menace de l'exclusion explicative

Comme je l'ai soutenu ailleurs, un matérialiste qui souscrit au réalisme intentionnel est confronté à ce que j'ai appelé « le dilemme du réalisme intentionnel³⁰ ». Ce dilemme surgit de la tension entre deux desiderata : une tendance réductionniste et une tendance antiréductionniste.

D'une part, comme nous venons de le voir dans la section précédente, un partisan du réalisme intentionnel se sent tenu de conférer une efficacité causale originale aux propriétés sémantiques des états cérébraux d'un individu. Son réalisme le pousse à accorder aux faits sémantiques une responsabilité causale particulière dans la production de faits physiques : un système physique capable de former des représentations douées de propriétés sémantiques doit pouvoir produire des effets physiques originaux. D'autre part, un matérialiste ne peut admettre que les faits sémantiques soient des faits métaphysiques bruts ou ultimes. Le fait qu'un état cérébral possède un contenu (ou une propriété sémantique) ne peut pas être une donnée métaphysique irréductible. C'est ce qu'exprime Fodor lorsqu'il dit qu'un réaliste intentionnel ne peut que souscrire au réductionnisme et que les propriétés sémantiques n'ont aucune chance de figurer dans le « catalogue ultime » des propriétés répertoriées par la physique fondamentale³¹. C'est aussi ce qu'exprime Dennett lorsqu'il compare toute explication qui fait référence au contenu à un emprunt : quiconque fait un emprunt sémantique est tenu de le rembourser en termes non sémantiques³². C'est aussi ce que dit Field qui tient la doctrine qu'il nomme le « sémanticalisme » pour incompatible avec le matérialisme³³. Un matérialiste partisan du réalisme intentionnel est donc tiraillé entre deux projets : il cherche à résoudre le problème de « la causalité mentale » et il cherche à combler le fossé entre les propriétés sémantiques et les propriétés non sémantiques d'un dispositif physique (c'est le projet de la « naturalisation de l'intentionnalité »).

Dans une série d'écrits importants³⁴, Kim a mis le matérialiste qui est aussi un partisan du réalisme intentionnel au défi de répondre à l'argument de « l'exclusion explicative ». Cet argument est destiné à convaincre le matérialiste de renoncer à la composante antiréductionniste du réalisme intentionnel. L'argument de l'exclusion explicative procède en deux étapes dont la première consiste à montrer au partisan de ce que Kim nomme le « matérialisme non réductionniste » — par exemple, le partisan du physicalisme des exemplaires, par opposition au physicalisme des types — que toute chaîne causale « horizontale » impliquant l'exemplification de deux propriétés

30. Voir Jacob, 1997, ch. I.

31. Voir Fodor, 1987, p. 97.

32. Voir Dennett, 1971.

33. Voir Field, 1972.

34. Dont bon nombre sont recueillis dans Kim, 1993b. Voir aussi Kim, 1998.

mentales (ou sémantiques) suppose que l'exemplification de la propriété mentale qui joue le rôle de cause est inévitablement impliquée dans une chaîne causale « descendante³⁵ ». Un matérialiste non réductionniste soutient que la propriété mentale d'un état cérébral (par exemple, la propriété sémantique ou le contenu) n'est pas identique ou réductible à une propriété physique intrinsèque de l'état cérébral³⁶. Un matérialiste non réductionniste peut donc souscrire au physicalisme des exemplaires et identifier un exemplaire d'une représentation mentale (d'une attitude propositionnelle, par exemple) à un état cérébral concret de l'individu.

L'exemplification d'une propriété mentale (ou sémantique) *M* peut être un maillon dans une chaîne causale « horizontale » lorsqu'elle provoque l'exemplification d'une autre propriété mentale (ou sémantique) *M**. Lorsqu'une pensée produit une autre pensée, on est en présence d'une chaîne causale « horizontale » de ce genre. L'exemplification d'une propriété mentale (ou sémantique) *m* peut être un maillon dans une chaîne causale « descendante » lorsqu'elle provoque l'exemplification d'une propriété physique *P**. Lorsqu'une représentation mentale produit un geste corporel, on est en présence d'une chaîne causale « descendante » de ce genre³⁷.

Supposons donc que l'exemplification de la propriété sémantique *M* provoque l'exemplification d'une autre propriété sémantique *M**. Dans ce qu'on pourrait appeler « le carré de Kim », il est évidemment crucial de supposer que les propriétés sémantiques (ou mentales) sont distinctes des propriétés physiques sous-jacentes dont elles dépendent. Un partisan du matérialisme non réductionniste supposera naturellement que la propriété *M** dépend systématiquement d'une propriété physique non sémantique sous-jacente *P** ou — ce qui revient au même — que *P** sert de base physique à la réalisation de *M**. Selon les hypothèses du matérialiste non réductionniste, nous nous trouvons en présence d'un cas illustré par la Figure 1.

À cette étape de l'argument, Kim fait valoir que l'exemplification de *M** semble donc provoquée conjointement par l'exemplification de la propriété de *M* et par l'exemplification de *P** : la propriété sémantique *M** est exemplifiée à la fois parce qu'une autre propriété sémantique *M* a été exemplifiée et parce que sa base physique *P** est exemplifiée. Kim maintient qu'on ne peut ni envisager que l'exemplification de *M* et l'exemplification de *P** collaborent causalement à l'exemplification de *M** en faisant valoir que, du

35. Voir Kim, 1993a.

36. Un psychofonctionnaliste ou un partisan du monisme anomal est un représentant du matérialisme non réductionniste.

37. Par souci de complétude, on peut mentionner l'existence hypothétique de chaînes causales « ascendantes » dans lesquelles peuvent entrer des propriétés sémantiques : tels seraient les cas où l'exemplification de *M* provoquerait l'exemplification d'une propriété « émergente » sociale. Voir Kim, 1993a, p. 351. Ces cas ne joueront aucun rôle dans la suite.

Figure 1

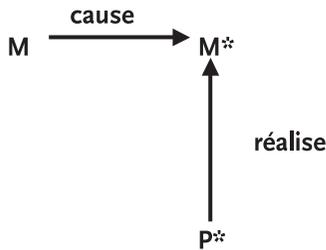
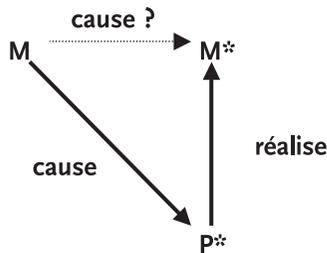


Figure 2



point de vue du matérialisme non réductionniste, l'exemplification de P^* à elle seule est une condition suffisante de l'exemplification de M^* ³⁸. Pour la même raison, Kim rejette l'interprétation selon laquelle l'exemplification de M^* serait « surdéterminée » par deux facteurs causaux. Il conclut la première étape de l'argument en indiquant que la seule voie ouverte au matérialiste non réductionniste est de supposer que l'exemplification de M contribue à produire l'exemplification de M^* grâce au fait qu'elle provoque l'exemplification de P^* , comme dans la Figure 2. Cette figure montre que, du point de vue du matérialisme non réductionniste, la causalité « horizontale » entre une exemplification de M et une exemplification de M^* dépend de (ou suppose) la causalité « descendante » entre une exemplification de M et une exemplification de P^* , la base physique de M^* .

La seconde étape de l'argument de l'exclusion explicative consiste à faire valoir que M elle-même doit aussi dépendre systématiquement d'une propriété physique non sémantique sous-jacente P , comme l'indique la Figure 3.

Dans la production de l'exemplification de P^* , l'exemplification de M entre donc en compétition avec l'exemplification de sa propre base physique P . Selon le « principe de la clôture causale du domaine physique »,

38. Si on suppose que M^* dépend de P^* dans l'interprétation forte de la dépendance systématique, cela va de soi puisque, dans cette interprétation, il est nécessaire (ou vrai dans tous les mondes) que si P^* est exemplifiée, alors M^* est exemplifiée aussi. Voir Kim, 1984.

Figure 3

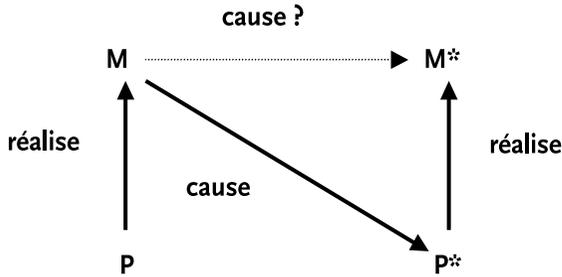
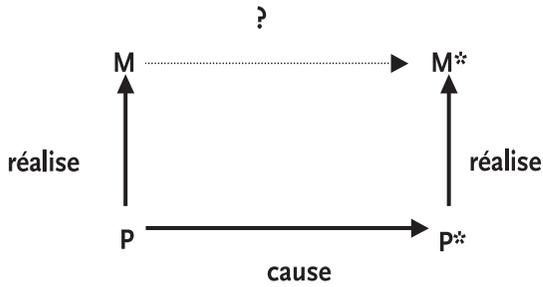


Figure 4



tout ancêtre causal ou tout héritier causal d'une exemplification d'une propriété physique doit être lui-même une exemplification de propriété physique. Dans la compétition pour produire l'exemplification de P^* , l'exemplification de M semble inévitablement devoir être privée d'efficacité causale par l'efficacité de l'exemplification de P . On se retrouve dans la situation illustrée par la Figure 4 dans laquelle la propriété P réalise la propriété M (ou réciproquement M dépend systématiquement de P), P^* réalise M^* et l'exemplification de P est la cause de l'exemplification de P^* . L'exemplification de M semble dépourvue de tout rôle causal. Comme le dit Kim³⁹, « dans le cas de la relation causale alléguée entre M et M^* , la situation est comparable à une série d'ombres projetées par une voiture en mouvement : il n'y a pas de relation causale entre l'ombre de la voiture à l'instant t et son ombre à l'instant suivant ; l'une et l'autre sont l'effet du mouvement de la voiture. Le mouvement de la voiture est un processus causal authentique, mais la séquence des ombres projetées, aussi régulière et nomologique qu'elle soit, n'est pas un processus causal⁴⁰ ».

39. Kim, 1998, p. 45.

40. La distinction entre processus causal authentique et pseudo-processus vient de Salmon, 1984. Voir Jacob, 1997, ch. VII.

S'il veut éviter la conclusion de Kim, le partisan du matérialisme non réductionniste peut sans doute résister à la première étape de l'argument qui vise à démontrer que la causalité « horizontale » entre l'exemplification de M et l'exemplification de M^* implique inévitablement la causalité « descendante » entre l'exemplification de M et l'exemplification de P^* , la base physique de M^* . Du moins peut-il rejeter le modèle de la causalité « descendante » que lui prête Kim. Kim traite en effet la contribution de P^* à l'exemplification de M^* comme un cas de relation causale en tous points comparable à la contribution supposée de M^{41} . Mais comme Kim lui-même l'a fait remarquer à juste titre à Searle⁴², c'est manifestement une erreur de traiter la relation de dépendance systématique entre la propriété M^* et sa base physique sous-jacente P^* comme une relation causale : même si l'exemplification de P^* est une condition suffisante de l'exemplification de M^* , la première n'est pas la cause de la seconde⁴³. Premièrement, dans une relation causale, la cause doit normalement précéder son effet. Or, l'exemplification de M^* et l'exemplification de sa base physique P^* sont simultanées. De surcroît, si on admet la version forte de la dépendance systématique entre l'exemplification de M^* et celle de sa base physique P^* , non seulement les deux exemplifications sont simultanées, mais la relation entre les deux exemplifications ne peut pas, comme une relation causale singulière, être une relation contingente : tout ce qui est P^* à l'instant t est nécessairement M^* à cet instant. Dans tous les mondes possibles, si quelque chose exemplifie P^* à l'instant t , cette chose doit exemplifier M^* à cet instant. Les deux exemplifications sont reliées de manière nomologique, mais la loi qui les relie n'est pas une loi causale. Cette loi ne peut pas être une loi causale en raison de la simultanéité entre l'exemplification respective des deux propriétés impliquées. C'est une identité nomique du type « eau = H₂O ».

L'exemplification de M est dans une relation causale avec l'exemplification de M^* . Quoique l'exemplification de P^* soit une condition suffisante de l'exemplification de M^* , la première n'est pas une cause de la seconde. L'exemplification de M et l'exemplification de P^* font donc une contribution différente à l'exemplification de M^* . L'efficacité causale de l'exemplification de M dans la production causale de l'exemplification de M^* n'est donc pas menacée par l'exemplification de P^* . Le partisan du matérialisme non réductionniste peut rejeter la thèse de Kim selon laquelle toute relation causale « horizontale » entre M et M^* implique une relation causale « descendante » entre M et P^* , la base physique de M^* . Ce que le partisan du matérialisme non

41. Dans la note 30, p. 351, Kim, 1993a, écrit de la rivalité entre les rôles respectifs de P^* et de M à l'exemplification de M^* « qu'elle est essentiellement identique à la situation dans laquelle nous sommes lorsque nous sommes en présence de deux causes distinctes d'un seul et même événement, chacune pouvant prétendre être une cause suffisante ».

42. Voir Kim, 1998, p. 47-50. La critique de Kim est dirigée contre Searle, 1992.

43. Slors, 1998, fait la même observation.

réductionniste objecte au modèle kimien de la causalité « descendante », c'est la supposition selon laquelle la chaîne causale reliant l'exemplification de M à l'exemplification de M^* passe par deux relations causales intermédiaires : premièrement, la relation causale entre l'exemplification de M et l'exemplification de P^* et deuxièmement, la relation causale entre l'exemplification de P^* et l'exemplification de M^* . Le partisan du matérialisme non réductionniste peut, d'une part, soutenir que produire l'exemplification de M^* , ce n'est rien d'autre que produire l'exemplification de P^* , puisque les deux propriétés sont co-exemplifiées simultanément. En produisant l'une, l'exemplification de M produit ipso facto l'autre. Il peut, d'autre part, soutenir que l'exemplification de P^* est nomologiquement constitutive de l'exemplification de M^* : M^* est exemplifiée en vertu du fait que P^* est exemplifiée ; M^* ne serait pas exemplifiée si P^* n'était pas exemplifiée au sens où un liquide ne serait pas de l'eau s'il n'était pas composé de molécules de H_2O . Mais l'exemplification de P^* n'est pas la cause de l'exemplification de M^* . L'exemplification de M peut donc rester la cause « horizontale » de l'exemplification de M^* sans que son efficacité causale soit menacée par celle de P^* .

Ce que répudie le partisan du matérialisme non réductionniste dans le modèle de la causalité « descendante » que lui impute Kim, c'est l'hypothèse d'une sous-relation causale « ascendante » entre l'exemplification de M^* et l'exemplification de sa base physique, P^* . Mais, même si le partisan du matérialisme non réductionniste peut désarmer la première étape de l'argument de l'exclusion explicative et, non sans bonnes raisons, refuser d'admettre que la causalité mentale « horizontale » implique le modèle de la causalité « descendante » entre une exemplification de propriété sémantique et une exemplification de propriété physique sous-jacente que lui impute Kim, il n'est pourtant pas au bout de sa peine⁴⁴. En effet, comme l'indique la seconde étape de l'argument, lorsque l'exemplification d'une propriété sémantique contribue à produire l'exemplification d'une propriété physique, la seule application du principe de la clôture causale du domaine physique suffit à faire surgir la menace de l'exclusion explicative.

Comme l'a montré la section précédente, la conception componentielle du comportement et la stratégie de la dualité des *explananda* semblent permettre au matérialiste qui souscrit au réalisme intentionnel de parer à la menace externaliste et de conférer une efficacité causale aux propriétés sémantiques dans l'explication de la structure du comportement, c'est-à-dire de la coordination entre un état interne et un mouvement physique. Mais la stratégie de la dualité des *explananda* est-elle vraiment compatible avec le matérialisme? Cette stratégie s'appuie sur un contraste entre deux sortes de faits à expliquer : le mouvement physique et le comportement. L'explication du

44. Devant l'acte d'accusation de Kim, j'ai voulu défendre le dossier du matérialiste anti-réductionniste aussi loin que possible. Mais, comme l'indique la suite du texte, je reconnais la validité de la seconde étape de l'argument de Kim.

mécanisme de la production du mouvement physique (la réponse à ce que j'ai appelé la question 1) fait référence aux propriétés non sémantiques de l'état interne *c* — aux propriétés physiques, chimiques et neurophysiologiques intrinsèques de l'état *c* (dans le cas où *c* est l'état interne d'un animal). L'explication de la structure du comportement ou de la coordination entre *c* et *m* (la réponse à la question 2) fait référence à une propriété sémantique ou proto-sémantique de *c*. Mais comme l'a fait valoir Kim⁴⁵, le matérialiste partisan du réalisme intentionnel ne peut pas se contenter de faire valoir que les propriétés sémantiques d'un état interne peuvent contribuer à expliquer l'un des deux *explananda*. Le fait qu'il existe une « explication par des raisons » pour la structure du comportement ne suffit pas à écarter la menace de l'exclusion explicative par une explication physique ou neurophysiologique⁴⁶. Il doit de surcroît montrer pour chacun des deux *explananda* qu'il n'est pas susceptible de recevoir deux explications : une « explication par des raisons » et une « explication physique ». La seule voie qui paraît ouverte au partisan de la stratégie de la dualité des *explananda* est de faire valoir que la structure du comportement au sens de la conception componentielle ne peut pas recevoir d'explication physique ou neurophysiologique.

Celui qui soutient que la structure du comportement au sens de la conception componentielle n'est pas susceptible de recevoir une explication physique ou neurophysiologique semble confronté au dilemme suivant : soit il suppose que les faits et les phénomènes qu'expliquent les propriétés sémantiques — à savoir la structure du comportement — ne sont pas des faits et des phénomènes physiques ou neurophysiologiques mais sont des faits non physiques. Auquel cas, il semble condamné à répudier purement et simplement le monisme matérialiste et à admettre une forme de dualisme des substances de type cartésien. Soit il reconnaît que les faits constitutifs du comportement au sens de la conception componentielle sont des faits physiques ou neurophysiologiques. Cette seconde option paraît parfaitement plausible car même si j'ai souligné que, selon la conception componentielle, il convient de distinguer le processus comportemental du mouvement physique qui est un constituant du comportement, le comportement n'en est pas moins un processus physique complexe composé d'un état physique interne et d'un mouvement physique. Mais dans ce cas, ne sommes-nous pas conduits à admettre qu'il existe des faits physiques ou neurophysiologiques qui n'ont pas de causes physiques ou neurophysiologiques?

Si on écarte la première option (c'est-à-dire si on rejette l'idée que les faits constitutifs du comportement sont des faits non physiques ou non neurophy-

45. Voir Kim, 1990, 1991.

46. Je suis ici la terminologie de Kim (1990, 1991) qui est elle-même fidèle au sous-titre de l'ouvrage de Dretske, 1988, « Les raisons dans un monde de causes » : une explication par des raisons est une explication qui invoque la propriété sémantique d'un état interne d'un organisme.

siologiques) et si on retient la seconde option, on se retrouve confronté au dilemme auquel était confronté, selon Kim⁴⁷, le partisan du matérialisme non réductionniste, à la seconde étape de l'argument de l'exclusion explicative, lorsqu'il est mis en présence d'un cas de causalité « descendante ». Dans ce cas, l'exemplification d'une propriété sémantique *M* est reliée par une chaîne causale à l'exemplification d'une propriété physique *P**. Comme le partisan du matérialisme non réductionniste doit admettre l'existence d'une base physique ou neurophysiologique *P* de *M* (dont *M* dépend systématiquement), il ne peut octroyer une efficacité causale à la propriété sémantique *M* que s'il consent à enfreindre le principe de la clôture causale du domaine physique et à supposer qu'un phénomène physique ou neurophysiologique peut avoir une cause non physique ou non neurophysiologique.

Le partisan du dualisme des *explananda* et de la conception componentielle du comportement doit donc choisir entre deux options : enfreindre le principe de la clôture causale du domaine physique ou renoncer à la composante antiréductionniste du réalisme intentionnel. Comme le suggère Kim⁴⁸, et comme le concède Dretske⁴⁹, la première option semble franchement inadmissible et, pour un partisan du monisme matérialiste, la meilleure stratégie est sans doute la seconde. Mais peut-il choisir la seconde option sans renoncer du coup à la doctrine externaliste? Choisir la seconde option, c'est affirmer que les propriétés sémantiques d'un état interne du cerveau d'un individu sont des propriétés physiques au sens large. Or, qui admet la vérité de l'externalisme reconnaît que les propriétés sémantiques d'une représentation mentale ne dépendent pas systématiquement des propriétés physiques, chimiques et neurophysiologiques intrinsèques du cerveau d'un individu. *A fortiori*, les propriétés sémantiques d'une représentation mentale ne peuvent-elles pas être identiques aux propriétés physiques, chimiques et neurophysiologiques intrinsèques du cerveau d'un individu⁵⁰. L'externalisme n'interdit-il pas la seconde option? Un matérialiste peut-il souscrire à l'externalisme?

Il me semble qu'il peut souscrire à une version « large » de la thèse de l'identité. Les propriétés sémantiques ne sont en effet ni des propriétés intrinsèques du cerveau d'un individu ni des propriétés non physiques. Ce sont des relations entre des états ou des propriétés physiques du cerveau de l'individu et des propriétés physiques exemplifiées dans son environnement. Ce sont donc des relations physiques. S'il admet l'externalisme, un partisan du monisme matérialiste doit concéder que les propriétés sémantiques d'une attitude propositionnelle d'un individu ne sont pas identifiables ou réductibles

47. Voir Kim, 1993a.

48. Voir Kim, 1990, 1991.

49. Voir Dretske, 1991, p. 211, et Dretske, 1995, p. 147-148.

50. Il fait peu de doutes que lorsqu'ils envisageaient (pour la rejeter) l'identification entre les propriétés sémantiques d'une représentation mentale et des propriétés physiques, les adversaires fonctionnalistes du matérialisme réductionniste (dont Putnam et Fodor) pensaient à des propriétés physiques intrinsèques du cerveau d'un individu.

aux propriétés physiques, chimiques ou neurophysiologiques intrinsèques d'un état cérébral de l'individu. Les premières ne dépendent pas systématiquement des secondes. Mais elles dépendent des relations entre les premières et des propriétés physiques exemplifiées dans l'environnement de l'individu. Aussi extrinsèques soient-elles, ces relations n'en sont pas moins physiques pour autant⁵¹.

S'il admet qu'une propriété sémantique n'est autre qu'une relation entre des propriétés neurophysiologiques du cerveau d'un individu et des propriétés physiques exemplifiées dans son environnement, un matérialiste externaliste peut-il encore revendiquer une efficacité causale particulière pour les propriétés sémantiques dans l'explication de la structure du comportement? Je crois que pour répondre au défi de l'exclusion explicative de Kim, le partisan de la dualité des *explananda* et de la conception componentielle du comportement doit recourir à ce que Kim nomme l'« argument de la généralisation⁵² ». Il doit soutenir que les propriétés sémantiques sont dans le même cas que les propriétés non sémantiques des « sciences spéciales » — a propriété chimique d'être une molécule de H₂O, la propriété géologique d'être une montagne ou la propriété biologique d'être un éléphant⁵³. Quoiqu'elles ne soient pas exemplifiées par les particules physiques élémentaires qui constituent le domaine de la physique fondamentale, ces propriétés n'en sont pas moins exemplifiées par des agrégats complexes de particules physiques élémentaires. Ces propriétés sont donc elles-mêmes des propriétés physiques complexes. Comme l'attestent les explications des sciences spéciales, l'efficacité causale de ces propriétés physiques complexes dans des processus chimiques, géologiques et biologiques est indispensable⁵⁴.

Bibliographie

- Churchland, Paul M., 1981. « Eliminative Materialism and the Propositional Attitudes », dans Churchland, Paul, M., *A Neurocomputational Perspective, the Nature of Mind and the Structure of Science*, Cambridge, (Mass.), MIT Press, 1989.
- Davidson, Donald, 1967. « Causal Relations », dans Davidson, Donald, *Essays on Actions and Events*, Oxford, Clarendon Press, 1980.
- Dennett, Daniel, 1971. « Intentional Systems », dans Dennett, Daniel, *Brainstorms: Philosophical Essays on Mind and Psychology*, Cambridge (Mass.), MIT Press, 1978.

51. Kim, 1987, p. 88, soutient la même chose.

52. Voir Kim, 1998, p. 77. C'est notamment la voie choisie par Fodor, 1989.

53. Comme l'a souligné Millikan, 1993, une entité ne peut être un éléphant à moins d'avoir un pedigree historique phylogénétique. Comme les propriétés sémantiques, la propriété d'être un éléphant est donc une propriété historique.

54. Je dois laisser pour une autre occasion une justification appropriée de ce que Kim, 1998, nomme « l'argument de la généralisation ».

- Dennett, Daniel, 1987. *The Intentional Stance*, Cambridge (Mass.), MIT Press.
- Dretske, Fred, 1988. *Explaining Behavior*, Cambridge (Mass.), MIT Press.
- Dretske, Fred, 1991. « Replies », dans McLaughlin, Brian, dir., *Dretske and his Critics*, Oxford, Blackwell.
- Dretske, Fred, 1995. « Reply : Causal Relevance and Explanatory Exclusion », dans MacDonald, Cynthia & MacDonald, Graham, dir., *Philosophy of Psychology*, Oxford, Blackwell.
- Dretske, Fred, 1998. « Minds, Machines, and Money : What Really Explains Behavior », dans Bransen, Jan & Cuypers, Stefaan, E., dir., *Human Action, Deliberation and Causation*, Dordrecht, Kluwer Academic Publishers.
- Field, Hartry, 1972. « Tarski's Theory of Truth », dans Platts, Mark, dir., *Reference, Truth and Reality*, Londres, Routledge & Kegan Paul.
- Fodor, Jerry, 1987. *Psychosemantics*, Cambridge (Mass.), MIT Press.
- Fodor, Jerry, 1989. « Making Mind Matter More », dans Fodor, Jerry, A., *A Theory of Content and Other Essays*, Cambridge (Mass.), MIT Press, 1990.
- Jacob, Pierre, 1997. *What Minds Can Do*, Cambridge, Cambridge University Press, et *Pourquoi les choses ont-elles un sens?*, Paris, Odile Jacob.
- Jacob, Pierre, 1998. « What Can the Semantic Properties of Innate Representations Explain? », dans Bransen, Jan & Cuypers, Stefaan, E., dir., *Human Action, Deliberation and Causation*, Dordrecht, Kluwer Academic Publishers, 1998.
- Jeannerod, Marc, 1997. *The Cognitive Neuroscience of Action*, Oxford, Blackwell.
- Kim, Jaegwon, 1973. « Causation, Nomic Subsumption, and the Concept of Event », dans Kim, Jaegwon (1993b).
- Kim, Jaegwon, 1984. « Concepts of Supervenience », dans Kim, Jaegwon (1993b).
- Kim, Jaegwon, 1987. « “ Strong ” and “ Global ” Supervenience revisited », dans Kim, Jaegwon (1993b).
- Kim, Jaegwon, 1990. « Explanatory Exclusion and the Problem of Mental Causation », dans MacDonald, Cynthia & MacDonald, Graham, dir., *Philosophy of Psychology*, Oxford, Blackwell, 1995.
- Kim, Jaegwon, 1991. « Dretske on How Reasons Explain Behavior », dans Kim, Jaegwon (1993b).
- Kim, Jaegwon, 1993a. « The Nonreductivist's Troubles With Mental Causation », dans Kim (1993b).
- Kim, Jaegwon, 1993b. *Supervenience and Mind, Selected Philosophical Essays*, Cambridge, Cambridge University Press.
- Kim, Jaegwon, 1998. *Mind in a Physical World : An Essay on the Mind-Body Problem and Mental Causation*, Cambridge (Mass.), MIT Press.
- McDowell, John, 1994. *Mind and World* : Cambridge (Mass.), Harvard University Press.

- Millikan, Ruth Garrett, 1993. *White Queen Psychology and Other Essays for Alice*, Cambridge (Mass.), MIT Press.
- Neander, Karen, 1988. « What Does Natural Selection Explain? Correction to Sober », *Philosophy of Science*, 55, p. 422-426.
- Neander, Karen, 1995. « Pruning the Tree of Life », *The British Journal for the Philosophy of Science*, 46, (1), p. 59-80.
- Salmon, Wesley, 1984. *Scientific Explanation and the Causal Structure of the World*, Princeton, Princeton University Press.
- Searle, John, 1992. *The Rediscovery of the Mind*, Cambridge (Mass.), MIT Press.
- Slors, Marc, 1998. « Two Claims That Can Save a Nonreductive Account of Mental Causation », in Bransen, Jan & Cuypers, Stefaan, E., dir., *Human Action, Deliberation and Causation*, Dordrecht, Kluwer Academic Publishers.
- Sober, Elliott, 1984. *The Nature of Selection*, Cambridge (Mass.), MIT Press.
- Stich, Steven, 1983. *From Folk Psychology to Cognitive Science*, Cambridge (Mass.), MIT Press.