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

**Artikel:** Actions, raisons, et causes mentales
**Autor:** Engel, Pascal
**DOI:** https://doi.org/10.5169/seals-381487

**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:** 13.01.2026

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

REVUE DE THÉOLOGIE ET DE PHILOSOPHIE, 124 (1992), P. 305-321

# ACTIONS, RAISONS, ET CAUSES MENTALES

PASCAL ENGEL

*Résumé*

*L'une des difficultés principales du matérialisme contemporain est la menace d'épiphénoménalisme: si les propriétés mentales dépendent systématiquement des propriétés physiques, comment peuvent-elles avoir une efficacité causale? Le 'monisme anomal' de Davidson ne résout ce problème qu'au prix d'une conception «lâche» de l'individuation des événements, et d'une relative imprécision de la pertinence des explications causales formulées en termes psychologiques. D'autres conceptions de l'individuation des événements et du pouvoir causal des états mentaux, comme celles de Kim et de Jackson et Pettit ne résolvent pas mieux le problème. Ce n'est donc pas en modifiant la théorie de l'individuation des événements qu'on le résoudra.*

## I

Mon imperméable gris souris me protège de la pluie. Il a cet effet parce qu'il a la propriété d'être imperméable, et non pas parce qu'il est gris souris. Cette dernière propriété n'a aucune pertinence causale dans le fait que mon imperméable me protège de la pluie (un imperméable vert pomme aurait le même effet). Mais si mon imperméable a cette propriété, c'est parce que les fibres du tissu ont une certaine composition physico-chimique. Donc en ce sens cette propriété n'a elle aussi aucune pertinence causale par rapport au fait que ce tissu me protège de la pluie. Ou tout au moins c'est ce que nous avons tendance à dire quand nous souscrivons à l'idée que les relations causales que nous observons au niveau macroscopique sont sous-tendues par des relations causales au niveau microscopique — que nous identifions couramment avec le niveau physique de description. Essayons maintenant de transposer ces remarques à un autre type de relations causales sanctionnées par le sens commun, les relations causales entre des événements ou états mentaux et des événements ou états physiques. Mon désir de me rendre à un colloque sur l'action, ma croyance qu'il y a un colloque sur l'action à Neuchâtel, ainsi que ma croyance qu'en me rendant à Neuchâtel je satisferai ce désir, causent mon voyage à Neuchâtel.

Certaines des propriétés de mes états mentaux, comme le fait que cette croyance soit la première que j'aie eue après le petit déjeuner, ne sont pas causalement pertinentes. Mais le fait que j'aie eu *cette* croyance et *ce* désir le sont (une autre croyance et un autre désir n'auraient pas nécessairement produit cette action). Mais on peut dire aussi que cette croyance et ce désir se sont produits dans mon esprit parce qu'ils sont sous-tendus par certaines configurations de neurones dans mon cerveau, et qu'en ce sens le fait qu'elles soient une croyance *qu'il y a un colloque à Neuchâtel sur l'action et un désir d'aller à un colloque sur l'action* n'ont pas de pertinence causale. Il semble donc bien possible de faire un parallèle entre la causalité physique et la causalité mentale. Tout comme on peut dire que l'imperméabilité d'un tissu a le pouvoir causal de me protéger de la pluie, tout en admettant que ce pouvoir causal provient avant tout de la composition physico-chimique des fibres du tissu, on peut dire que des états mentaux doués de contenu ont le pouvoir causal de produire mon action, tout en admettant que ce pouvoir causal appartient avant tout aux états de mon cerveau[1]. Mais cette impression est trompeuse, et le parallèle douteux. Car s'il est plausible de soutenir que les états physico-chimiques des fibres du tissu sous-tendent la propriété d'imperméabilité et lui confèrent son pouvoir causal, c'est parce que nous avons de bonnes raisons de supposer qu'il y a une relation nomologique entre des propriétés comme l'imperméabilité et les propriétés physico-chimiques des fibres du tissu, alors que nous n'avons aucune raison de supposer qu'il existe une telle relation entre les contenus mentaux et les états neuronaux qui les sous-tendent. Il existe pourtant une manière de sauver le parallèle tout en reconnaissant ce fait. Elle consiste à dire que la propriété mentale — par exemple la croyance *qu'il y a un colloque sur l'action à Neuchâtel* — a un rôle causal, mais qu'elle a ce rôle non pas *en tant que* propriété mentale, mais en tant qu'elle est instanciée dans un état physique du cerveau qui tombe lui-même sous une loi causale. C'est notoirement la solution adoptée par Davidson. Selon Davidson on peut sauver le parallèle entre causalité physique et causalité mentale tout en reconnaissant le fait qu'il n'y a pas de lois reliant le mental et le physique, parce qu'on peut soutenir que les trois principes suivants (que je reformule ici légèrement) sont compatibles: (1) il y a de la causalité mentale, c'est-à-dire des relations causales entre états mentaux et états physiques, (2) toute relation causale implique l'existence d'une loi causale,

[1] «De mon cerveau» est ici légèrement trompeur (comme me l'a fait remarquer P. Jacob) pour caractériser la position de Davidson: en toute rigueur il suffit de parler pour lui d'une propriété *physique* sans spécifier qu'il s'agit d'une propriété interne à l'individu, ou de son cerveau. Mais dans la mesure où je ne discute pas dans cet article la pertinence du monisme anomal pour le problème de l'externalisme du mental, cette caractérisation peut être gardée telle quelle.

et (3) il n'y a pas de lois du mental ni des relations du mental au physique. Ces trois principes sont, selon Davidson, compatibles, parce que bien que les états mentaux, en tant qu'ils sont décrits comme mentaux, n'instancient aucune loi, ils instancient des lois quand on les décrit en termes des événements physiques auxquels ils sont corrélés. Davidson en tire aussi argument pour identifier les événements mentaux et les événements physiques, puisque les premiers ont des descriptions physiques en tant qu'ils instancient des lois physiques. Cette solution — le «monisme anomal» — est ingénieuse parce qu'elle permet d'assigner un rôle causal aux événements mentaux, tout en reconnaissant que c'est le cerveau qui joue le rôle essentiel dans la causalité mentale. Mais elle est loin d'être satisfaisante. L'objection la plus fréquente est qu'elle rend le mental inerte causalement ou épiphénoménal, puisque ce n'est pas en tant que mental, mais en tant que physique que le mental joue son rôle causal. Une objection voisine consiste à relever que cette conception semble contredire la thèse célèbre de Davidson selon laquelle les raisons sont des causes, puisque ce ne sont pas en tant que décrits comme raisons, mais en tant que décrits physiquement, que les événements mentaux peuvent causer des actions[2].

Le problème soulevé par ces objections — que l'on peut appeler le problème de la pertinence causale du mental — semble trouver sa source dans deux thèses centrales et étroitement liées de la philosophie de l'esprit et de l'action de Davidson. La première est que les actions sont des événements, et que les événements sont des entités dont l'individuation est lâche, et non pas fine, c'es-à-dire qui existent indépendamment des descriptions que l'on peut en donner. Ceci se traduit immédiatement quand il s'agit d'examiner les relations du mental et du physique. Selon Davidson les événements sont des particuliers, au même titre que les tables, les chaises et les girafes. Ils peuvent avoir différents types de propriétés, et en particulier des propriétés mentales et des propriétés physiques. Qu'on les décrive mentalement ou physiquement est indifférent. La seconde thèse est que la relation de causalité est une relation entre événements, par conséquent elle aussi indépendante de la manière dont ces événements sont décrits. La conséquence directe de cette thèse est que c'est une chose que d'énoncer l'existence d'une relation causale entre deux événements, et une autre chose que d'expliquer l'occurrence de ces événements, en les décrivant comme physiques ou comme mentaux. Il s'ensuit que la relation de causalité n'est pas sensible à la description, alors que la relation d'explication l'est. La première est extensionnelle, la seconde intensionnelle. En particulier une description d'un événement comme instanciant une certaine loi physique ne peut pas être substituée *salva veritate* à une autre description, pas plus que

[2] La littérature est abondante. Cf. notamment Honderich 1982, Stoutland 1985, Kim 1984 et 1990, Laurier 1988 et 1989, McDonald 1989.

ne peut l'être une description mentale de cet événement. On comprend dès lors pourquoi le problème de la pertinence causale du mental peut se poser, et comment on peut, au moins en partie, le résoudre. Quand on s'inquiète du fait qu'un certain événement *e*, décrit sous une description mentale, ne puisse pas causer *en tant que mental* un événement *f* décrit sous une description physique, on oublie que, pour Davidson *e* cause *f* quelle que soit la manière dont ils sont décrits, et par conséquent que le soupçon que *e* n'ait pas de pouvoir causal est sans fondement, puisque par définition on a supposé que *e* et *f* entraient dans une relation causale. L'objection a simplement ignoré la distinction entre causalité et explication. Mais même si l'on admet cette distinction et si l'on souscrit à cette conception de la causalité, l'objection a encore sa force. Elle revient à se demander en quoi peut bien consister, dans cette conception une explication causale en général, et en quoi peut consister une explication causale par les raisons, c'est-à-dire une explication qui rende compte de la pertinence de la cause mentale par rapport à son effet. Le principe (2) du caractère nomologique de la causalité semble impliquer qu'une explication causale est une explication nomologique. Mais dans ce cas le principe (3) d'anomie du mental semble impliquer que les explications mentales ou par les raisons ne sont pas des explications causales. Comment Davidson peut-il alors dire que «causalité et rationalité vont main dans la main» et que les raisons sont des «causes rationnelles»?[3] Le même problème peut se formuler en termes de la notion d'événement: quelle forme d'individuation des événements reliés causalement a le plus de chances d'expliquer la causalité mentale? Leur individuation physique ou leur individuation mentale? La tension vient du fait que bien que Davidson semble admettre qu'il n'y ait pas en droit de prééminence des descriptions physiques des événements sur leurs descriptions mentales, ce sont bien les descriptions physiques, et donc les explications causales *physiques* qui sont pertinentes parce que seules ces dernières sont nomologiques, ce qui a pour effet de priver les explications mentales du titre d'explications causales *bona fide*. En d'autres termes, Davidson semble souscrire à une doctrine que l'on appelle le *réductionnisme causal*, selon laquelle les relations causales observables à un niveau supérieur donné — ici mental — sont réductibles à, ou tout au moins dépendent, de relations causales à un niveau inférieur - ici physique. Cette doctrine est plus générale que celle qui dit que les relations causales impliquent l'existence de lois, mais puisque Davidson soutient que seule la physique contient des «lois strictes», cette dernière thèse implique chez lui la première.

Ce que l'on a appelé, dans la littérature récente, le problème de «l'épiphénoménisme du contenu (mental)» a pris d'autres formes que celles qui

[3] Cf. par exemple Davidson 1982.

sont spécifiques au monisme anomal[4]. Je choisis cependant de l'examiner dans ce cadre parce qu'il me semble suffisamment général pour poser le problème de la causalité mentale. Mon but n'est pas ici d'analyser cette question comme une question seulement interne à la doctrine de Davidson, mais je partirai des deux prémisses qui apparaissent chez lui, d'après ce qui précède, problématiques: la conception «lâche» de l'individuation des événements d'une part, et le réductionnisme causal de l'autre. J'envisagerai seulement trois solutions possibles du problème, en ayant bien conscience qu'il y en a d'autres. La première consiste à rejeter l'individuation lâche des événements sans renoncer au réductionnisme causal. La seconde consiste à garder l'individuation lâche mais à rejeter le réductionnisme causal. La troisième rejette aussi le réductionnisme causal, mais entend rester agnostique quant à l'individuation des événements. Je soutiendrai que c'est la seconde version qui a le plus de chances d'être correcte.

## II

La première solution du problème de la pertinence causale du mental que je voudrais examiner consiste à rejeter la conception davidsonienne des événements comme entités dont l'individuation est lâche, et par là-même la théorie de l'identité des événements mentaux et des événements physiques qu'elle implique. Les problèmes soulevés par cette conception sont bien connus. Davidson nous dit que quand je tourne le commutateur, illumine la pièce, et alerte un rôdeur, ce sont des descriptions du même événement. Mais cela a des conséquences problématiques. Si je tourne le commutateur avec la main, et si l'événement de tourner le commutateur est identique à l'événement d'alerter le rôdeur, il semble s'ensuivre que j'alerte le rôdeur avec la main, ce qui semble être une action très différente de celle décrite initialement. Davidson propose bien un critère d'identité des événements: deux événements *e* et *e'* sont identiques si et seulement si ils ont les mêmes causes et les mêmes effets. Mais comme on l'a fait remarquer[5], d'une part la clause «seulement si» de cette définition ne dit rien d'autre que la discernabilité des non identiques: des entités qui diffèrent par leurs causes et leurs effets sont distinctes parce que toutes les entités qui diffèrent par une propriété ou une autre sont distinctes. Quand à la clause «si» de la définition, elle ne semble pas être correcte parce que les descriptions que nous donnons de ce que Davidson traite comme un même événement peuvent ne pas avoir les mêmes effets ni les mêmes causes: par exemple tourner le

[4] Cf. Dretske 1989 par ex., Fodor 1990, ch. 5 («Making mind matter more»), Jacob 1991a et 1991b.

[5] Mellor 1987. Davidson 1985 admet lui-même la faiblesse de son critère, mais n'en propose pas d'autre.

commutateur cause l'envoi d'électrons du fil électrique à l'ampoule, mais alerter le rôdeur ne cause rien de tel. Il existe plusieurs manières de réviser le critère, soit en identifiant les événements à des régions spatio-temporelles, soit en diversifiant les types d'entités tenues comme *relata* de la relation causale. Je ne les examinerai pas, et ne considérerai que la version proposée par Kim[6]. Selon Kim les événements sont les référents d'expressions nominales de la forme «le fait que *a* ait la propriété *F* en *t*», où l'expression peut renvoyer aussi bien à un état de choses qu'à un état ou une condition d'un objet. En ce sens un événement est pour Kim le fait qu'un objet exemplifie une propriété à un temps donné. Deux événements sont identiques quand ils impliquent les mêmes objets, et instancient les mêmes propriétés au même temps. En ce sens allumer la lumière et alerter le rôdeur ne sont pas le même événement, parce qu'ils n'impliquent pas les mêmes objets, n'exemplifient pas la même propriété et sans doute ne se produisent pas en même temps. Mais ceci ne veut pas dire qu'il n'y ait pas une relation causale entre ces événements. Cette relation est, selon Kim, la relation de dépendance, ou de survenance (*supervenience*) causale[7]. D'une manière générale, une famille *A* de propriétés ou de relations survient sur une famille *B* de propriétés ou de relations si et seulement si, pour toute propriété *F* dans *A*, si un objet *x* a *F*, alors il existe une propriété *G* dans *B* telle que *x* a *G*, et nécessairement tout ce qui a *G* a *F*. Un événement *e* survient sur un autre *e*' si et seulement si *e* instancie une propriété *G* qui survient sur une propriété *F* que *e*' instancie. On peut alors définir la relation de *causalité survenante.* Il y a une telle relation, selon Kim, entre le fait qu'un événement *e* a une propriété *F* et le fait qu'un événement *e*' a *G* si et seulement si il y a deux événements, *e* qui a *F** et *e*' qui a *G**, tels que le fait que *e* a *F* survient sur le fait que *e* a *F**, et le fait que *e*' a *G* survient sur le fait que *e*' a *G**, et si le fait que *e* a *F** cause le fait que *e*' a *G**. L'application intéressante de cette définition concerne le cas où *e* a une propriété mentale *F* et cause *e*' qui a une propriété physique *G*: il y a une telle relation de causalité mentale seulement si cette relation survient sur la relation [*F*(e)* cause *G*(e')*]. Cette dernière relation, selon Kim, est une relation entre des événements qui instancient des propriétés physiques à un «micro-niveau», c'est-à-dire à un niveau inférieur au niveau «macro» des propriétés décrites par le langage commun. Ainsi les relations causales au niveau supérieur sont-elles survenantes sur les relations causales au

[6] Kim 1975, et pour une analyse générale, McDonald 1989. On peut considérer que Goldman 1970 donne une analyse comparable à celle de Kim.

[7] Kim 1986, 1990. On me presse, de tous côtés, de cesser de traduire *supervenience* par «survenance». Mais les termes proposés, «superposition» (Jacob), «surimposition» (Gayon), «dépendance» (Laurier), me paraissent également inadéquats.

niveau inférieur. La relation de survenance est plus faible que la relation d'identité des propriétés mentales et des propriétés physiques — que Kim rejette, mais elle est plus forte que la relation d'identité entre des *tokens* événements proposée par Davidson. Cette affirmation peut paraître surprenante, parce que Davidson lui-même invoque l'idée d'une survenance des propriétés mentales sur les propriétés physiques. Il veut dire par là que tout changement dans une propriété mentale *P* d'un événement *e* doit être accompagné d'un changement dans les propriétés physiques de *e* (pas de différences mentales sans différences physiques). Mais cette forme de survenance n'est pas celle que vise Davidson, parce qu'elle n'implique pas qu'un changement dans les propriétés mentales d'*autres* événements que *e* sera accompagné d'un changement identique dans les propriétés physiques de ces autres événements, alors que la conception de Kim implique que les *mêmes* propriétés physiques doivent changer avec les *mêmes* propriétés mentales (cependant sans que les secondes se réduisent aux premières). Autrement dit, la relation de survenance visée par Kim est une survenance *forte* (comme l'indique l'adverbe «nécessairement» qui figure dans la définition ci-dessus), alors que la survenance visée par Davidson est la survenance *faible*. Cette relation est, pour Kim, trop faible parce qu'elle ne nous dit pas *quelles* propriétés physiques doivent changer avec les propriétés mentales, et c'est pourquoi elle ne peut pas nous dire *en vertu de quoi* les causes mentales causent leurs effets physiques. Selon Kim, c'est en vertu du fait qu'elles réalisent, au niveau inférieur, des ensembles spécifiés de propriétés physiques[8].

La conception de la causalité survenante de Kim implique le principe suivant: quand deux événements sont unis par une relation de cause à effet, il ne peut pas y avoir deux explications causales distinctes de cette relation, mais une seule. En particulier, si l'on dit qu'un événement mental cause un événement physique, on ne peut pas à la fois en donner une explication «mentale» ordinaire au niveau «macro» en termes de contenus («son désir de vengeance l'a conduit jusqu'en Birmanie») et une explication physique au niveau «micro». Si l'une des explications est la bonne, elle doit exclure l'autre et être indépendante de toute autre explication. C'est ce que Kim (1990) appelle le *Principe d'Exclusion Causale-Explicative*. Puisque, selon Kim, les relations causales au niveau supérieur (c'est-à-dire intentionnel ou

[8] La différence est, je crois, bien exprimée par un parallèle proposé par Peacocke (1979, p. 122). Supposons que l'on dise que les propriétés (ou les vérités, ou les lois sur) des nations surviennent sur les propriétés, les vérités ou les lois des individus qui les composent. On peut soutenir cela sans dire de *quelles* propriétés des individus dépendent les généralisations portant sur les nations, c'est-à-dire sans soutenir que les propriétés de la «base de survenance» *déterminent* les propriétés survenantes énoncées au sujet des nations. Dans la première hypothèse, la survenance est faible, dans la seconde elle est forte

en termes de psychologie populaire) surviennent sur les relations causales au niveau inférieur, en vertu du Principe d'Exclusion, seules les secondes sont réellement explicatives. Au contraire la théorie de Davidson, avec sa distinction entre causalité et explication, semble autoriser un «double aspect» des explications mentales (par les raisons) et des explications physiques. C'est précisément parce qu'il y a cette tension entre deux types d'explications, alors même que Davidson semble ramener toute causalité à une causalité physique, que l'on a l'impression que chez lui le mental ne joue aucun rôle causal. La seule manière d'éliminer cette tension est, selon Kim, d'admettre l'idée qu'il n'y a qu'un seul type d'explications causales véritables, ce qui revient à souscrire pleinement à la thèse qu'on a appelée ci-dessus la thèse du réductionnisme causal.

## III

Nous avons examiné deux réponses générales possibles au problème de la pertinence causale du mental, que nous pouvons appeler respectivement le modèle de l'identité et le modèle de la survenance (forte, mais j'omettrai cette spécification). Le modèle de l'identité proposé par Davidson semble impliquer que des événements mentaux n'ont de pertinence causale que s'ils sont identiques à des événements physiques qui sont eux-mêmes causalement pertinents. Le modèle de la survenance de Kim dit que les événements mentaux ne sont pertinents que s'ils surviennent sur des événements physiques. Le premier modèle repose sur une individuation lâche des événements, alors que le second repose sur une individuation fine. Le lien, signalé ci-dessus, qui existe entre la conception davidsonienne de l'individuation des événements et le fait que le monisme anomal ne semble pas apporter de solution au problème de la pertinence causale du mental, paraît rendre plus favorable une conception plus fine de l'individuation des événements et le second modèle. Mais en fait ce dernier modèle n'est pas plus satisfaisant que le premier, comme le montrent les considérations suivantes.

L'objection la plus répandue contre le modèle de la survenance consiste à faire remarquer qu'il n'est pas vrai, en général, que les états mentaux, en tant qu'individualisés par leurs contenus, surviennent sur des états physiques ou du cerveau des individus, en invoquant des expériences de pensée du type de la Terre-Jumelle de Putnam où des jumeaux physiquement identiques peuvent différer dans leurs états mentaux en vertu des environnements physiques distincts qu'ils habitent, ou des expériences de pensée du type de celles de Burge où des états mentaux différents peuvent être attribués au même individu si l'on suppose des variations dans les normes de sa communauté linguistique. Ces arguments sont complexes et controversés, et bien qu'ils menacent fortement le modèle de la survenance, ils peuvent

appeler en réponse, de la part du partisan de ce modèle, une distinction entre deux types d'états mentaux, ceux qui sont «larges» et dont l'individuation dépend de l'environnement et du contexte, et ceux qui sont étroits et dont l'individuation dépend des seuls états physiques internes des individus, les premiers pouvant ainsi être survenants sur les premiers au sens désiré. Je ne chercherai pas à examiner ces arguments, non seulement parce qu'ils me semblent problématiques tout comme la distinction entre contenu étroit et contenu large, mais aussi et surtout parce qu'il me semble qu'il y a une manière plus simple — bien que rappelant par certains côtés la stratégie de ces arguments — de montrer l'inadéquation du modèle de la survenance. En fait, pour établir la fausseté de ce modèle, il suffit de montrer que deux événements survenant sur le même événement peuvent différer dans leurs relations causales avec d'autres événements. Dans ce cas, les deux événements ne peuvent pas dériver leur rôle causal de l'événement sur lequel ils sont supposés survenir. Peter Menzies (1988, cf. aussi Pettit 1991) a donné des contre-exemples de ce genre. Il n'est pas évident que tous ces exemples soient des exemples où une relation *causale* est en jeu, mais ce sont tous des cas où une certaine relation *explicative* est présente (je reviendrai sur ce point plus bas).

Le premier exemple est celui de Ben, qui court un cent mètres en 10 secondes, et qui ce faisant bat l'adversaire qui est dans le couloir de droite, mais est battu par l'adversaire qui est dans le couloir de gauche. Dans ce cas, le fait que Ben batte le coureur du couloir de droite et le fait qu'il soit battu par le coureur du couloir de gauche survient sur le fait que Ben court son cent mètres en 10 secondes en vertu du fait que les deux autres concurrents ont couru le cent mètres en un certain temps (disons respectivement 10''1 et 9''8). Et pourtant nous pouvons dire que la cause du fait que Ben perd la course est le fait qu'il a été battu par le coureur de gauche, pas le fait qu'il ait battu le coureur de droite. Bien que les deux événements surviennent sur le même événement — que Ben a couru en 10'' — ils ont des rôles causaux différents.

Deuxième exemple. Pierre me dit «Bonjour» d'une voix forte et il dit «Bonjour» sur un ton rogue. Les deux événements surviennent sur le même événement, à savoir le fait que Pierre ait dit «Bonjour» d'une voix forte et sur un ton rogue. Mais il se peut que ces événements n'aient pas les mêmes causes ni les mêmes effets: par exemple Pierre a pu dire «bonjour» d'une voix forte parce qu'il était énervé, et le fait qu'il ait dit «bonjour» sur un ton rogue a pu causer ma gêne.

Ces exemples peuvent ne pas être convaincants contre un modèle comme celui de Kim, parce qu'ils ne supposent pas que la «base de survenance» soit constituée d'événements microphysiques. Mais le troisième exemple satisfait cette condition. La conductivité d'un métal et son opacité sont causées par le nuage d'électrons libres qui emplit tout le métal et qui

donne à ses atomes un état solide. Mais quand un courant électrique passe à travers un morceau de cuivre, nous n'expliquons pas cet effet en disant qu'il est produit par le fait que le morceau de cuivre est opaque; nous l'expliquons par le fait qu'il est un bon conducteur. On peut voir au passage que le modèle de l'identité ne permet pas mieux d'expliquer, puisque selon ce modèle, la conductivité du métal est simplement identique au fait qu'il est empli d'un nuage d'électrons libres; mais l'opacité du métal est aussi identique à ce fait.

On pourrait essayer de transposer ce type d'exemples au cas de la causalité mentale. C'est plus difficile, parce qu'il faut pour cela envisager qu'il existe certaines relations psychophysiques plus ou moins établies, ce qui risque d'entraîner une pétition de principe relativement à la question qui nous occupe. Supposons que la peur de l'alpiniste mais aussi son courage pour continuer l'escalade surviennent sur une poussée d'adrénaline dans son organisme. Mais si nous voulons expliquer pourquoi l'alpiniste a eu peur pendant son ascension de la paroi, nous dirons que c'est parce qu'il a eu une poussée d'adrénaline, pas parce qu'il a eu du courage.

Il est assez clair, d'après ces exemples, que la relation de survenance, au sens où Kim la comprend — mais pas plus celle d'identité — ne permettent de résoudre le problème de la pertinence causale. Ce sont tous des exemples où il y a une certaine corrélation et une certaine dépendance de certains événements à un niveau supérieur par rapport à des événements à un niveau inférieur, mais où la corrélation n'est, pour ainsi dire, pas sensible à nos intérêts explicatifs dans les cas considérés, parce qu'elle introduit des causes non pertinentes. Or le modèle de la survenance repose, comme on l'a vu, sur la thèse du réductionnisme causal et sur le principe d'exclusion causale-explicative, c'est-à-dire sur l'idée que toutes les relations causales surviennent sur des relations causales microphysiques et sur l'idée qu'il ne peut pas y avoir d'autre explication causale possible. Ce sont donc ces principes eux-mêmes que les exemples nous conduisent à mettre en doute.

Comme on l'a noté plus haut, le monisme anomal semble impliquer lui aussi le principe du réductionnisme causal dans la mesure où le principe (2) du caractère nomologique de la causalité énonce que toute relation causale implique l'existence d'une loi causale, et d'une loi causale «stricte», c'est-à-dire d'une loi du type de celles qu'on peut trouver, selon Davidson, seulement en physique. Dans la mesure où la nomologicalité d'une relation causale est pour Davidson un trait de l'*explication* et non pas la causalité comme telle (qui n'est qu'une relation entre événements), il semble naturel d'en conclure que selon Davidson il n'y a qu'un seul type d'explications véritables, les explications au niveau physique, et que c'est cette incapacité de concevoir un autre type d'explication qui est à l'origine de ce que l'on

appelle son épiphénoménalisme. Plusieurs auteurs[9] ont ainsi suggéré que le monisme anomal reposait sur un physicalisme outrancier, lui-même basé sur une conception totalement erronée des lois physiques et du déterminisme. Abandonnez ces conceptions, suggèrent-ils, et les apories de la causalité mentale disparaîtront. Je n'examinerai pas ce point, mais je ne suis pas sûr que ces apories disparaîtront du seul fait que l'on aura adopté une autre conception des lois physiques et du déterminisme, car, comme on l'a vu, la thèse de la survenance de Kim ne fait pas explicitement allusion à la notion de loi physique, mais rencontre les mêmes apories: comme le fait remarquer Pettit (1991), la difficulté générale vient du fait que l'on postule que des antécédents causaux non physiques qui ont un pouvoir causal doivent être identiques ou survenants sur des antécédents causaux *physiques*, qu'on considère où non que ceux-ci soient soumis à des lois strictes. Une stratégie plus effective semble être celle qui consiste à abandonner l'idée qu'il n'y a qu'un seul type d'explication causale possible, c'est-à-dire à abandonner à la fois le principe d'exclusion causale explicative et le principe du réductionnisme causal. Et en fait c'est ce que fait Davidson lui-même. Il admet explicitement que bien que seules les descriptions physiques des événements mentaux puissent être authentiquement légales, cela n'interdit pas des descriptions mentales légales qui ne reposent pas sur des lois strictes, mais qui sont néanmoins des explications causales[10]. En d'autres termes, les événements mentaux selon lui peuvent tomber sous des lois, mais ce sont des lois non strictes, hétéronomiques ou *ceteris paribus.*

Toute la question est alors de savoir ce que peuvent être ce genre d'explications causales non strictement légales, et comment elles peuvent se rattacher à des explications causales physiques. Plusieurs auteurs ont développé cette stratégie, et en particulier Fodor (1987, 1989), McLaughlin (1989), Horgan (1989), et Le Pore et Lœwer (1987). Je ne mentionnerai que les trois derniers, parce qu'ils ont en commun une analyse de la relation causale en termes de contrefactuels et entendent conserver l'ontologie davidsonienne. Horgan commence par admettre que toute explication causale doit nous dire en quoi une cause est pertinente pour l'occurrence de son effet (ceci est supposé capturer le sens de la relation *en tant que* ou *en vertu de* (ou «*qua*») employée quand on dit que tel événement mental cause qua mental un événement physique, relation que Horgan appelle joliment la *quausation*). Il soutient ensuite que pour que deux propriétés ou types *F* et *G* sous lesquels tombent deux événements reliés causalement soient causalement pertinentes, il faut une condition contrefactuelle de la forme: «*S'il*

[9] Mellor & Crane 1990, Van Fraassen 1989, p. 23-25.

[10] Ceci est clair d'après la distinction davidsonienne entre lois strictes et lois non strictes admise dans *Mental Events* (Davidson 1980), et explicite dans Davidson 1990.

*n'y avait pas eu un événement de type F, alors il n'y aurait pas eu un événement de type G*». Mais cette condition n'est que nécessaire. Si l'on dit par exemple que dans le cas où une vitre s'est brisée parce que la Castafiore a chanté l'air des Bijoux avec une voix stridente, la vitre ne se serait pas brisée si la Castafiore n'avait pas chanté l'air des Bijoux avec une voix stridente, ce contrefactuel est nécessaire, mais pas suffisant si l'on ne précise pas quelle sorte de hauteur du son peut avoir pour effet de briser la vitre. Horgan propose, dans le style de David Lewis (1973), une «Condition de Pertinence» d'après laquelle l'événement *e* qui est la cause de l'événement *e'* doit être *similairement pertinent* à un événement qui dans un autre monde possible suffisamment semblable aurait pu causer un événement du même type. Pour établir les mesures de similarité nécessaires, Horgan propose alors de combiner cette condition avec des méthodes d'accord et de différence du type de celles de Mill. Horgan ne donne pas de précisions sur ce point, mais il est clair que la condition de pertinence qu'il propose est distincte d'une relation nomologique entre événements, et que les explications à ce niveau sont supposées être distinctes de celles qui interviennent au niveau physique. Le Pore et Lœwer (1987) distinguent eux aussi deux sortes de pertinence causale, l'une selon laquelle:

a) des propriétés *F* et *G* sont pertinentes-1 pour faire en sorte que *e* cause *e'*,

et l'autre selon laquelle:

b) le fait que *e* possède la propriété *F* est causalement pertinent-2 pour le fait que *e'* possède *G*.

La pertinence-1 correspond au cas où il y a une loi stricte impliquant que les *F* causent des *G*, alors que la pertinence-2 n'implique que l'existence de lois *ceteris paribus* et de contrefactuels qui supportent les énoncés causaux correspondants.

Chacune de ces conceptions est supposée être consistante avec le physicalisme *token-token* et avec l'individuation lâche des événements à la Davidson. Toutes deux rejettent le réductionnisme causal et le principe d'exclusion, et divisent les explications en deux types, les unes légales strictes, les autres contrefactuelles. Ces dernières font appel à des régularités, mais pas à des lois, et peuvent aussi consister en des explications causales singulières du type «si *c* n'avait pas été présent *e* n'aurait pas eu lieu». Cette stratégie, cependant, pose deux problèmes. Le premier est que si l'on admet une distinction entre deux types d'explications causales, les unes légales-physiques et les autres pas, que devient l'inférence des principes (1) et (2) à l'identité des événements mentaux et physiques qu'autorisait le monisme anomal? Comme le fait remarquer Fodor (1989, p. 153) à partir du moment où nous admettons que les explications causales peuvent être de deux sortes, les unes propres aux événements en tant que physiques,

les autres propres aux événements en tant que mentaux, nous perdons toute justification du physicalisme. Le second problème est lié au précédent: quelle est la relation entre les deux types d'explications causales proposées? Si nous disons, par exemple dans la conception de Le Pore et Lœwer, que la pertinence-1 implique la pertinence-2, alors nous risquons de retomber dans le modèle de la survenance (forte). Et si nous disons que la pertinence 2 implique la pertinence 1, nous risquons de ramener les explications causales contrefactuelles à des explications via des régularités nomiques. Mais si nous séparons les deux types d'explication radicalement, en disant que l'une n'implique pas l'autre, nous revenons à une bifurcation des deux types d'explications — les unes mentales, les autres physiques, qui ressemble fortement à une forme de dualisme (cf. Kim 1990, p. 44).

## IV

La difficulté qui vient d'être énoncée est celle de savoir en quoi une relation causale énoncée au niveau «macro» ou supérieur — comme le sont typiquement les explications causales en termes de contenus intentionnels de la psychologie populaire — peut être causalement pertinente sans que pour autant on exclue la pertinence causale des explications énoncées au niveau «micro» ou inférieur, et vice versa. Rejeter radicalement le réductionnisme causal, c'est admettre que les deux types d'explications puissent éventuellement ne pas fonctionner de concert. Mais comment, dans ces conditions, peut-on soutenir que les explications par les raisons sont des explications causales, comme le font Davidson et la plupart des auteurs cités ici, sans faire une équivocation sur le terme de «cause»? Il existe cependant plusieurs tentatives pour réinstaurer le parallélisme entre les deux types d'explications. L'une est celle de Fodor, pour qui les lois *ceteris paribus* de la psychologie populaire doivent être «implémentées» dans des mécanismes physiques qui assurent la médiation entre le «supérieur» et «l'inférieur» (Fodor 1987, 1989). Une autre solution est celle de Pettit et Jackson (1988, Pettit 1991). C'est celle que j'examinerai pour finir.

Pettit et Jackson partent d'exemples semblables à ceux qui ont été énoncés au début de III ci-dessus. Dans le cas du métal qui est à la fois conducteur et opaque en vertu du nuage d'électrons libres qui l'emplit, ce qui est recherché, pour qu'il y ait une pertinence causale (relativement au fait à expliquer, par exemple que quelqu'un s'électrocute en touchant le métal), c'est l'existence d'un facteur causal au niveau supérieur qui soit indépendant de, mais qui rende certain, ou tout au moins probable, l'existence d'un facteur causal au niveau inférieur. Ici la conductivité du métal rend certaine l'existence d'un nuage l'électrons, alors que son opacité ne la rend pas certaine. De même, pour prendre un exemple célèbre invoqué

aussi par Pettit, quand j'essaie de faire passer un morceau de bois carré dans un trou circulaire dont le diamètre est égal au côté du morceau de bois, et que le morceau se bloque, je peux dire d'une part que cet effet intervient en raison de la forme carrée du morceau de bois, et d'autre part en raison de l'impénétrabilité de la partie du morceau de bois qui ne passe pas par le trou. Quelle explication choisir? Toutes deux sont bonnes, mais il y a une relation entre les deux: la forme carrée du morceau de bois rend certaine la présence d'une partie impénétrable du morceau. Dans le vocabulaire de Jackson et Pettit, le premier type d'explication *programme* la seconde, et c'est en ce sens qu'elle est pertinente, alors que la seconde est proprement responsable de l'événement à expliquer. Toutes deux sont pertinentes. Jackson et Pettit appellent les explications du premier type *programmatiques* et les explications du second type des explications par *processus.* Le trait intéressant de la conception de Jackson et Pettit est que ces deux types d'explications doivent fonctionner de concert pour qu'on ait une véritable explication causale pertinente: l'une seulement des deux types d'explication n'est pas suffisante, et en particulier les explications programmatiques ne mentionnent pas de causes (Jacob 1990). Appliquée au problème de la causalité mentale, la distinction devient celle entre des explications fonctionnelles d'une part, spécifiant des états mentaux par leurs rôles fonctionnels, qui programment les états physiques satisfaisant ces rôles, et les explications neurophysiologiques qui analysent les processus physiques sur lesquels reposent ces états. La solution de Jackson et Pettit est donc compatible avec une conception fonctionnaliste des états mentaux. L'autre aspect intéressant de ce modèle est le diagnostic qu'il nous fournit sur la manière d'individualiser les événements. Le modèle implique le rejet de l'identité des événements décrits au niveau supérieur programmatique avec les événements décrits au niveau inférieur: la conductivité du métal, par exemple, n'est pas identique avec le fait qu'il existe un nuage d'électrons, puisque ce dernier fait est aussi causalement responsable de son opacité. Cela nous conduit donc, dans ce cas, à favoriser une individuation fine des événements à la Kim, puisque ce ne sont pas les mêmes propriétés qui sont exemplifiées. Mais Pettit (1991) admet aussi que rien ne nous empêche de dire que la conductivité est identique à la présence du nuage d'électrons. Il suffit alors de dire que la conductivité programme, bien que l'opacité ne programme pas, l'existence du nuage d'électrons. Dans ce cas, l'événement est causalement pertinent *en tant que (qua)* conductivité et *en tant que (qua)* présence d'un nuage d'électrons libres, mais pas *en tant qu'*opacité. Pettit en conclut que le modèle «programme-processus» laisse ouverte la question de l'individuation des événements: elle peut-être, selon nos intérêts explicatifs, fine ou lâche.

V

Mais cette dernière remarque est assez surprenante. Car admettre qu'il peut y avoir un seul et même événement qui soit un facteur causal en tant que décrit comme un *F* mais pas en tant que décrit comme un *G* c'est souscrire à la distinction davidsonienne entre la relation extensionnelle de causalité et la relation intensionnelle d'explication. Au contraire, soutenir que les propriétés *G* et *G* sont des propriétés exemplifiées dans deux événements distincts c'est confondre ces deux relations. Mais le modèle de Jackson et Pettit ne nous offre aucune raison de justifier cette assimilation. Au contraire, il propose de distinguer deux types d'explications, programmatiques et processuelles, d'un seul événement. En d'autres termes, il confirme plutôt qu'il n'infirme la métaphysique qui sous-tend le monisme anomal, et nous autorise à dire que l'individuation de ces entités ne peut être que lâche.

Ce que ce modèle ajoute — ou permet de rendre explicite — c'est le type d'épistémologie de l'explication causale que doit adopter le monisme anomal, en distinguant les explications causales au niveau supérieur — «contentuel» ou intentionnel ou par les raisons — des explications causales au niveau inférieur - physique ou microphysique, ce qui implique, comme on l'a vu, un rejet du principe d'exclusion et du réductionnisme causal. A partir du moment où l'on rejette ce principe, la question de savoir *comment* les propriétés mentales sont réalisées dans des propriétés physiques perd son importance. La seule condition requise est qu'elles soient réalisées, d'une manière ou d'une autre, dans des propriétés physiques, si nous ne voulons pas souscrire à une ontologie dualiste. Si nous savions comment ces propriétés sont réalisées, nous aurions résolu le problème esprit-corps. Mais ainsi formulé, je ne vois pas comment ce problème peut être résolu (cf. McGinn 1989). Car le physicalisme, pas plus que le dualisme, ne résout le problème épistémologique, et c'est en ce sens qu'« il n'y a pas de question du physicalisme» (Mellor & Crane 1990).

J'ai, dans cet article, voulu soulever une difficulté qui semble commune à diverses formes de matérialisme quant à la pertinence causale du mental. Parce que cette difficulté semble affecter le monisme anomal de Davidson, on a essayé de proposer des conceptions de la pertinence des propriétés mentales et de l'individuation des événements qui permettraient de rendre compte du pouvoir explicatif de ces propriétés. Mais il est apparu que ces conceptions ne permettent pas mieux que celle de Davidson de résoudre la difficulté. J'en ai conclu que rien ne nous autorisait à renoncer à l'ontologie davidsonienne et à l'individuation lâche des événements. A partir du moment où l'on renonce au réductionnisme causal et au principe d'exclusion explicative, le problème pendant est celui de savoir quels *types* d'explications causales la psychologie usuelle peut nous fournir. La taxinomie de Jackson et Pettit est un premier pas en ce sens, mais il semble clair qu'il

ne va guère plus loin que certaines formulations davidsoniennes. Le problème posé demeure donc, en l'occurrence, ouvert[11].

## REFERENCES

DAVIDSON, D. (1980), *Essays on Actions and Events*, Oxford, Oxford University Press, tr. fr. à paraître, Paris, P.U.F.

DAVIDSON, D. (1982), «Paradoxes of Irrationality», in R. WOLLHEIM, & J. HOPKINS (éds), *Philosophical Essays on Freud*, Cambridge, tr. fr. in *Paradoxes de l'irrationalité*, Nîmes, l'Eclat, 1991.

DAVIDSON, D. (1985), «Reply to Quine on Events», in LEPORE & MCLAUGHLIN (éds), *Actions and Events, Perspectives on the Philosophy of Donald Davidson*, Oxford, Blackwell.

DAVIDSON, D. (1990), «Thinking Causes», inédit.

DRETSKE, F. (1989), «Reasons and causes», in Tomberlin, 1989, pp. 1-15.

ENGEL, P. (1986), «Le mental dépend-il du physique?», inédit.

EDWARDS, J. (1990), «Functional Support For Anomalous Monism», in D. KNOWLES (éd.), *Explanation and Its Limits*, Royal Institute of Philosophy Lectures Series: 27, Cambridge, Cambridge University Press, pp. 45-64.

FODOR, J. (1987), *Psychosemantics*, Cambridge Mass, MIT Press.

FODOR, J. (1989), *A Theory of Content and Other Essays*, Cambridge Mass, MIT Press.

GOLDMAN, A. (1970), *A Theory of Human Action*, Englewood Cliffs, New Jersey, Prentice Hall.

HORGAN, T. (1989), «Mental Quausation» in TOMBERLIN, 1989, pp. 47-76.

HONDERICH, T. (1982), «The Argument for Anomalous Monism», *Analysis*.

JACOB, P. (1991a), «Are Mental Properties Causally Efficacious?», *Grazer Philosophische Studien*, vol. 39, pp. 51-73.

JACOB, P. (1991b), «Externalism and the Explanatory Relevance of Broad Content», inédit.

JACKSON, F. and PETTIT, P. (1988), «Functionalism and Broad Content», *Mind*, 97, pp. 381-400.

KIM, J. (1975) «Events as Property Exemplifications», in M. BRAND & D. WALTON (eds), *Action Theory*, Dordrecht, Reidel, pp. 159-177.

KIM, J. (1984), «Epiphenomenal and Supervenient Causation», *Midwest Studies in Philosophy*, 9, pp. 257-70.

[11] Des versions assez différentes de ce texte ont été lues en janvier 1991 à la Société Romande de Philosophie à Genève et à l'Université de Montréal en février 1991. Je remercie Kevin Mulligan, Jerôme Dokic, Daniel Laurier, Michel Seymour, pour leurs remarques en ces occasions, ainsi que, pour la version présente, Kevin Mulligan à nouveau, et Filip Buekens pour de longs commentaires écrits, ainsi que tous les participants au colloque à Neuchâtel.

KIM, J. (1990), «Explanatory Exclusion and Mental Causation» in VILLANEUVA, E. (ed.), *Information, Semantics and Epistemology*, Oxford, Blackwell, pp. 36-56.

LAURIER, D. (1988), «Le mental a-t-il des effets physiques?», in *Hermès, Psychologie ordinaire et sciences cognitives*, 3, Paris, Presses du CNRS.

LAURIER, D. (1989), «Le monisme anomal et la dépendance psychophysique», Université de Montréal.

LEPORE, E. and LOEWER, B. (1987), «Mind Matters», *Journal of Philosophy*, 84, pp. 630-642.

LEWIS, D. (1973), «Causation», *Journal of Philosophy*, 70, pp. 556-567.

MC LAUGHLIN, B. (1989), «Type Epiphenomenalism, Type Dualism, and the Causal Priority of the Physical», in TOMBERLIN, 1989, pp. 109-136.

MCDONALD, C. (1989), *Mind-Body Identity Theories*, Londres, Routledge.

MELLOR, D. H. and CRANE, T. (1990), «There Is No Question of Physicalism», *Mind*, 99, pp. 185-206.

MENZIES, P. (1988), «Against Causal Reductionism», *Mind*, 97, pp. 551-574.

MOYA, C. (1990), *The Philosophy of Action*, Oxford, Blackwell.

PEACOCKE,C. (1979), *Holistic Explanation*, Oxford, Oxford University Press.

PETTIT, P. (1991), «Pertinence causale et identité des événements», in PETIT, J. L, *Qu'est-ce qu'un événement?*, Paris, EHESS, Raisons pratiques.

TOMBERLIN, J. (éd.) (1989), *Philosophical Perspectives, 3, Philosophy of Mind and Action Theory*, Atascadero, Cal., Ridgeview.

VAN FRAASSEN, B. (1989), *Laws and Symetries*, Oxford, Oxford University Press.