



Linx

Revue des linguistes de l'université Paris X Nanterre

54 | 2006

La cause : approche pluridisciplinaire

Unité des sciences, causalité et rôle causal de l'esprit

Max Kistler



Édition électronique

URL : <http://journals.openedition.org/linx/515>

DOI : 10.4000/linx.515

ISSN : 2118-9692

Éditeur

Presses universitaires de Paris Nanterre

Édition imprimée

Date de publication : 1 juin 2006

Pagination : 139-150

ISSN : 0246-8743

Référence électronique

Max Kistler, « Unité des sciences, causalité et rôle causal de l'esprit », *Linx* [En ligne], 54 | 2006, mis en ligne le 01 août 2007, consulté le 06 mai 2019. URL : <http://journals.openedition.org/linx/515> ; DOI : 10.4000/linx.515

Unité des sciences, causalité et rôle causal de l'esprit

*Max Kistler, Université de Paris X-Nanterre
Institut Jean Nicod (CNRS)*

Introduction

Les adversaires de la philosophie peuvent être tentés de voir dans l'histoire de la réflexion sur le concept de causalité une illustration de plus de son impuissance : il semble y avoir autant de conceptions de la causalité qu'il y a de philosophes. Le débat sur la causalité ressemble à un dialogue de sourds où les différentes doctrines se suivent et coexistent sans entrer dans une confrontation fructueuse. Or, ce qui empêche un véritable progrès dans ce domaine, c'est l'absence de réflexion méthodologique : sans une réflexion préalable sur le type d'analyse qu'il s'agit de construire, il ne peut pas y avoir d'entente sur les critères qu'on utilisera pour juger du succès ou de l'échec d'une analyse donnée de la causalité. Or sans accord sur ces critères, les défenseurs de chaque doctrine n'ont aucune raison de remettre en cause leur propres thèses ni d'envisager celle de leurs adversaires.

Je commencerai par présenter l'une parmi les analyses contemporaines les plus importantes du concept de causalité : l'analyse dite « déductive-nomologique ». Je montrerai qu'elle est confrontée à des contre-exemples clairs. Un contre-exemple à l'analyse d'un concept est une situation où l'application du concept selon l'analyse proposée diverge du jugement intuitif : soit, dans le cas d'un « faux positif », l'analyse parvient au résultat que la situation contient une relation causale alors qu'intuitivement elle n'en contient pas ; soit, dans le cas d'un « faux négatif », l'analyse conclut à l'absence de relation causale alors qu'intuitivement, la situation en contient. Faute de place, je ne peux pas présenter ici deux analyses rivales très importantes : la théorie

probabiliste¹ et la théorie contrefactuelle² qui sont-elles aussi, confrontées à des contre-exemples, et essentiellement pour les mêmes raisons.

Face au constat de l'échec de ces tentatives d'analyse conceptuelle, il est primordial de choisir un critère d'adéquation, ce qui présuppose de déterminer quel genre d'analyse on entend construire. Je propose de considérer la causalité comme un concept analogue aux concepts d'*espèces naturelles*, c'est-à-dire en analogie aux concepts d'eau, d'or, de chat, ou d'être humain : la causalité est un genre naturel de relation (j'entends ici « relation » en un sens suffisamment large pour s'appliquer aussi aux *processus* causaux).

Cette conception du but de la recherche philosophique sur la causalité est tributaire de l'attitude *naturaliste*. Selon le naturalisme, la connaissance philosophique contribue à la construction de la connaissance scientifique. Selon l'idéal de l'unité de la connaissance, les théories élaborées dans les différentes sciences ont pour vocation d'être intégrées dans un système unique, articulées autour de *réductions* inter-théoriques. La génétique classique et la biologie moléculaire s'intègrent par exemple au sein de l'édifice de la connaissance scientifique par le biais d'une réduction du concept de gène à un concept chimique complexe construit en termes de parties de la molécule d'ADN et de mécanismes biochimiques qui se situent à l'échelle moléculaire. La psychologie cognitive, pour donner un second exemple, s'intègre dans l'édifice scientifique à mesure que certains de ses concepts se trouvent associés à des concepts neuroscientifiques. Lorsque l'on découvre, par exemple, les mécanismes au niveau neuronal et moléculaire, qui réalisent, dans un organisme donné, l'apprentissage par conditionnement, on est engagé sur la voie de la découverte des articulations susceptibles de faire de ces deux sciences des parties de l'édifice unique de la science.

Le critère naturaliste ne fournit pourtant pas le seul critère d'adéquation du concept de causalité qu'il s'agit d'élaborer. La causalité est après tout un concept issu du sens commun et reste profondément ancré en lui. Toute analyse acceptable de la causalité doit être compatible avec l'intuition, au moins dans une large majorité de cas : lorsqu'on étudie les situations concrètes, la théorie n'est acceptable que si son verdict sur la présence ou absence de relations causales s'accorde avec l'intuition. Il faut préciser qu'il peut être nécessaire de se contenter d'une seule majorité des cas où l'intuition est en accord avec la théorie en question, dans la mesure où il peut s'avérer que la pratique des jugements intuitifs n'est pas cohérente. À ce moment-là, la reconstruction philosophique ne peut pas en rester à la seule clarification du concept intuitif, mais devra procéder à une rectification. Par exemple, il s'avère que les jugements intuitifs que l'on est enclin à faire sur l'effet d'absences, omissions et autres faits *négatifs* conduisent à des contradictions avec des exigences fondamentales imposées au concept de causalité, notamment avec son anti-symétrie : il est vrai *a priori* de toute relation causale que si un événement particulier *c* est la cause d'un second événement *e*, *e* ne peut pas en même temps être cause de *c*. Cependant, faute d'imposer au concept de causalité une contrainte étrangère au sens commun, celui-ci conduit à la fois au résultat que la fonte d'un cube de glace dans un verre d'eau est la cause du

¹ Cf. Eells (1991).

² Cf. Lewis (1986).

refroidissement de l'eau dans le verre, et au résultat que l'énergie contenue dans l'eau fait fondre le glaçon, de sorte que ce même refroidissement est à la fois la cause et l'effet de la fonte du glaçon³.

On peut s'attendre à des controverses sur l'articulation précise entre ces deux contraintes. Cela n'empêche que l'objectif est clair : il s'agit de reconstruire le concept de causalité de manière compatible à la fois avec une contrainte *a priori* et une contrainte *a posteriori*. Selon la première, notre reconstruction ne doit s'éloigner du concept intuitif de causalité que dans la mesure du strict nécessaire pour maintenir la cohérence des pratiques intuitives de jugement dans leur ensemble. Selon la seconde, il s'agit de construire un concept qui s'applique indifféremment à toute l'échelle des phénomènes naturels, des particules subatomiques jusqu'aux agents humains. Il s'agit de trouver l'essence objective d'une espèce naturelle de relation (catégorie dans laquelle nous voudrions inclure les processus) à laquelle appartiennent aussi bien les collisions entre particules élémentaires que les processus cognitifs qui conduisent aux mouvements corporels d'un agent humain.

1. Le problème de la causalité mentale

Les remarques précédentes ont fait apparaître l'importance de la *réduction* pour l'unification progressive des sciences, les réductions de concepts accompagnant la réduction des théories dans lesquelles ils figurent. Mais le concept de causalité est tout aussi indispensable à l'unification de la connaissance.

On peut s'en convaincre en réfléchissant sur le simple fait que l'esprit peut changer le cours du monde. Je suis tranquillement assis dans mon fauteuil en passant en revue dans mon esprit les événements de la journée lorsque je me souviens d'avoir oublié de répondre à une lettre que j'ai reçue le matin. Cette pensée me fait me lever pour me mettre à taper une lettre sur le clavier de mon ordinateur. La pensée, ou plus précisément un ensemble de croyances et de désirs - la croyance d'avoir reçu une lettre le matin, de ne pas lui avoir répondu, la croyance qu'il est bon de répondre, le désir de faire ce que j'estime être bon etc. - a fait bouger les choses. D'après un critère métaphysique traditionnel, les entités réelles sont celles et seulement celles qui sont capables de faire une différence causale dans le monde. Selon ce critère, les entités mentales, et avec elles l'esprit, existent parce qu'elles sont causalement efficaces. Non seulement peuvent-elles causer, notamment lors de l'action, elles peuvent aussi être causées : lorsque je lis, un processus causal qui passe par la réception de lumière par mes rétines ainsi que par des événements neuronaux dans mon cortex visuel, cause la formation de croyances portant sur les idées exprimées dans le texte. Le pouvoir causal des pensées est la raison même de les considérer comme réelles ; mais ce pouvoir pose le problème de sa coexistence avec le pouvoir causal des événements neuronaux sous-jacents à ces pensées.

³ Sur le problème posé par des énoncés désignant un fait négatif comme la cause ou l'effet, cf. Kistler (à paraître).

Reprenons les exemples de l'apprentissage et de la mémoire, puisque ce sont des capacités cognitives qui ont déjà été réduites avec succès à la neurophysiologie. Ce sont des phénomènes réels et scientifiquement importants dans la mesure où ils ont un impact causal. C'est parce que le rat a mémorisé le labyrinthe qu'il trouve la sortie immédiatement, sans s'engager dans des culs de sac. C'est parce que le pigeon est conditionné à la lumière rouge qu'il se met à picorer dès que la lampe rouge s'allume. Dans les deux cas, il s'agit clairement d'explications causales. En effet, ce n'est qu'à condition d'interpréter les explications psychologiques comme causales que la psychologie peut être intégrée à l'ensemble des sciences à côté des neurosciences. En d'autres termes, cette intégration requiert que les concepts psychologiques fassent référence à des phénomènes réels, c'est-à-dire à des phénomènes qui font une différence causale au cours des événements.

Le problème qui se pose dans ce contexte est de comprendre comment la mémoire ou le résultat de l'apprentissage peuvent faire une différence causale, alors que la réduction de ces concepts semble mettre à notre disposition une explication complète du comportement de l'animal en termes neurophysiologiques : selon une conviction répandue, il est possible de fournir une explication complète du comportement du pigeon sans jamais quitter le niveau de la neurophysiologie, en décrivant les effets du conditionnement sur le plan de la modification des synapses⁴, ou même sur le plan des processus moléculaires auxquels la modification des synapses se réduit à son tour. Il semble envisageable de construire une explication causale complète du comportement du pigeon qui ne fait pas référence à des catégories cognitives, telles que la force du lien d'association entre les stimuli ou la représentation des stimuli dans des catégories. Cette explication ne ferait référence qu'à des signaux lumineux, des excitations de neurones, et le renforcement de liens synaptiques.

La question est cependant controversée de savoir s'il est en principe possible d'expliquer le comportement exclusivement en termes neurophysiologiques, sans aucune référence au niveau cognitif. Certains auteurs cherchent à montrer que les concepts cognitifs restent indispensables à l'explication du comportement des animaux et des humains, tout comme les concepts chimiques restent indispensables à l'explication des réactions entre substances chimiques, même lorsque ces concepts et les théories auxquelles ils appartiennent ont fait l'objet d'une réduction achevée⁵. L'analyse de cas concrets de réductions inter-théoriques semble montrer que la réduction ne fait pas disparaître les concepts réduits des explications scientifiques : la température reste un concept explicatif même après sa réduction à la mécanique statistique⁶ ; l'apprentissage par conditionnement continue d'être étudié sur le plan et dans des termes cognitifs même après sa réduction à des mécanismes neuronaux et moléculaires⁷. Il y a donc des raisons de penser que ces concepts identifient des phénomènes réels et différents des phénomènes dans leur base de réduction. D'autres auteurs, comme Jaegwon Kim, font valoir que lorsqu'un processus causal a fait l'objet

⁴ Une synapse est une structure par laquelle un neurone exerce une influence sur l'état d'un autre neurone.

⁵ Cf. Gold et Stoljar (1999).

⁶ Cf. Sklar (1993).

⁷ Cf. Rescorla (1988).

d'une réduction, le seul processus réel au sens de l'efficacité causale est le processus plus fondamental, autrement dit le processus décrit par la théorie réductrice⁸. Cela signifierait que les concepts qui ont fait l'objet d'une réduction n'identifient pas de phénomènes réels différents des phénomènes identifiés par les concepts réducteurs.

La possibilité même de mener ce débat sur l'efficacité causale des phénomènes ayant fait l'objet d'une réduction présuppose un travail de clarification conceptuelle : le défi pour l'analyse philosophique consiste à forger un concept de causalité qui permet d'envisager la possibilité que l'apprentissage soit à la fois un processus qui se réduit à des mécanismes moléculaires et un phénomène réel au sens de contribuer causalement au comportement d'un animal, phénomène réel distinct des mécanismes moléculaires. De cette manière, la reconstruction naturaliste du concept de causalité peut contribuer à la compréhension du processus d'unification des sciences.

En cas de succès, nous aurons fait la preuve que la philosophie est capable de progresser sans pour autant prétendre à la vérité absolue. Le travail conceptuel est complémentaire au travail scientifique : l'unification des différents domaines de connaissance scientifique requiert le recours aux concepts de réduction et de causalité ; ces concepts n'appartiennent à aucune science particulière et relèvent donc de la compétence de la philosophie.

2. La réduction de la causalité à l'explication déductive-nomologique

Selon la tradition de l'empirisme logique, qui fournit l'arrière-plan de la réflexion sur les sciences dans la tradition analytique, la relation causale est équivalente à la relation *explicative*. L'explication scientifique est à son tour comprise comme un argument déductif-nomologique, c'est-à-dire un argument déductif qui contient dans ses prémisses au moins une loi de la nature (d'où l'adjectif « nomologique », d'après le grec « nomos » qui signifie « loi »), ainsi que des « conditions initiales », correspondant à la description de la situation qui constitue l'arrière-plan du phénomène à expliquer. Plus précisément, l'empirisme logique identifie la cause avec les conditions initiales d'un tel argument, et l'effet avec sa conclusion⁹.

Il y a un certain nombre de contre-exemples, désormais bien connus, à cette analyse. Il s'agit d'explications *non-causales*, qui satisfont aux exigences de l'analyse déductive-nomologique, mais où il est clair que les conditions initiales ne désignent pas la cause du phénomène désigné par la conclusion. Cela montre que toutes les explications ne sont pas causales, autrement dit que l'explication causale n'est qu'une *espèce* (importante) à l'intérieur du *genre* des explications déductives-nomologiques. Dans une explication causale, la référence à la cause contribue à l'explication de l'effet. Or, les exemples suivants montrent que l'on peut expliquer un phénomène sans faire référence à sa cause.

⁸ Cf. Kim (1998).

⁹ La présentation classique de cette analyse de l'explication qui est en même temps censée rendre compte de la causalité, se trouve dans l'article de Hempel et Oppenheim (1948). Mais on la trouve également chez Popper (1934), pourtant critique à l'égard de l'empirisme logique à d'autres égards.

(i) Un genre important d'explications utilise des « principes d'extremum », tel que le principe de Fermat selon lequel le chemin d'un rayon de lumière est celui qui minimise le temps du parcours. On utilise ce principe pour expliquer la réfraction d'un rayon de lumière lors d'un passage à travers une interface entre deux milieux à indice de réfraction différent, par exemple lorsque le rayon, venant de l'air, pénètre dans l'eau. Les conditions initiales décrivent les points de départ et d'arrivée du rayon. À partir du principe de Fermat, on peut en déduire le point où le rayon franchit l'interface ; en général, ce point ne se situe pas sur la ligne droite entre le point de départ et le point d'arrivée, ce qui constitue le phénomène de réfraction. Il est clair que cette explication n'est pas causale au sens indiqué : l'événement que l'on explique et qui est mentionné dans la conclusion, à savoir le passage du rayon par l'interface, a lieu *plus tôt* que l'un des événements mentionnés dans les conditions initiales, à savoir l'arrivée du rayon. Le premier événement ne peut pas être l'effet du second, dans la mesure où le concept de causalité exige que la cause précède l'effet dans le temps.

(ii) Sylvain Bromberger (1966) fait remarquer que, dans le cas d'explications qui ne font appel qu'à des lois de la nature qui sont réversibles par rapport au temps, nous pouvons tout aussi bien expliquer la cause par l'effet que l'effet par la cause. Nous pouvons, par exemple, expliquer la longueur de l'ombre projetée par une tour, à partir de la hauteur de la tour, de la position du soleil et de la loi selon laquelle les rayons du soleil se propagent en ligne droite ; mais nous pouvons tout aussi bien expliquer la hauteur de la tour en partant de la même loi, de la position du Soleil et de la longueur de l'ombre. Il est clair que cette dernière explication ne peut pas être causale, dans la mesure où l'événement mentionné dans les conditions initiales, c'est-à-dire l'arrivée du rayon au sol, et *postérieur* à l'événement mentionné dans la conclusion, à savoir le passage du rayon au sommet de la tour¹⁰.

L'inadéquation de l'analyse déductive-nomologique de la causalité apparaît peut-être encore plus clairement lorsque l'on considère des explications qui ne mentionnent pas de loi causale du tout, c'est-à-dire qui ne mentionnent aucune loi qui gouverne l'évolution d'un processus à travers le temps.

On peut expliquer les lois de Kepler concernant la forme des orbites des planètes et la vitesse de leur rotation autour du Soleil, à partir des lois de Newton, des lois de conservation et de l'hypothèse selon laquelle le Soleil exerce une force centrale sur les planètes. Autrement dit, on explique les lois de Kepler à partir d'autres lois. Rien dans une telle explication ne peut tenir lieu de cause ou d'effet puisque l'explication ne fait référence à aucun événement spatio-temporel particulier. Or toutes les causes et tous les effets sont situés dans l'espace et le temps.

Un autre genre d'explication qui ne peut pas être interprété comme causale dans la mesure où il ne concerne pas le déroulement des phénomènes dans le temps, est l'explication « analytique » ou « compositionnelle ». On explique, par exemple, la température d'un gaz à partir de l'énergie cinétique moyenne de ses particules. Une telle explication ne donne pas la cause d'un événement mais répond plutôt à la question : quelle est *la nature* de la température ? Les particules qui constituent un gaz possèdent leur énergie cinétique moyenne au même moment où le gaz possède sa

¹⁰ On trouve d'autres contre-exemples dans Kistler (2004) et Kistler (à paraître).

température. L'explication de la température par l'énergie cinétique moyenne ne peut donc pas être causale, dans la mesure où la cause n'est jamais simultanée à son effet. On peut s'en convaincre ainsi : supposons que l'énergie moléculaire soit la cause de la température. Étant donné que l'on peut aussi bien expliquer l'énergie cinétique des particules par la température du gaz, il s'ensuit que la température est en même temps la cause de l'énergie cinétique des particules. En supposant que la causalité est transitive, cela conduit au résultat absurde que l'énergie cinétique est une cause de soi.

3. La causalité comme transfert

La causalité est essentiellement une relation (ou un processus) qui relie des événements particuliers. Prenons un ballon qui brise une vitre. Avant de poser la question de savoir *pourquoi* cette cause a provoqué cet effet, l'analyse de la relation causale en tant que telle doit tenir compte du fait qu'il s'agit d'un processus qui a lieu autour de l'endroit et de l'instant où la vitre se brise. Dans la mesure où l'analyse déductive-nomologique porte essentiellement sur les rapports entre propriétés ou facteurs *universels*, elle n'est pas adéquate à l'analyse de la relation causale particulière et localisée¹¹. L'analyse déductive-nomologique vise avant tout le fondement de la *régularité* avec laquelle les lancers de ballon causent des bris de vitre, plutôt que le lien singulier entre des événements qui ont lieu à un endroit et moment précis.

Je propose l'hypothèse que les explications causales visent des rapports complexes : ce qui rend vraie une telle explication est bien, d'une part, ce qui fait l'objet de la théorie déductive-nomologique – et de ses successeurs probabiliste et contrefactuelle – à savoir un rapport de dépendance entre facteurs ou propriétés universelles ; mais pour qu'il soit vrai que ce lancer de ballon ait brisé *cette* vitre, une autre condition doit être satisfaite qui porte sur les événements localisés. Il faut qu'il y ait un *transfert* entre l'événement « cause » et l'événement « effet ». Deux événements *c* et *e* – par exemple, le lancer du ballon en direction de la vitre et le bris de la vitre – sont liés comme cause et effet, si et seulement si le premier transfère quelque chose sur le second, en l'occurrence de l'énergie et de la quantité de mouvement. Plus généralement, lorsqu'une quantité d'une *grandeur conservée* est transférée entre deux événements, ces événements sont reliés comme cause et effet¹².

Pour que cette analyse de la causalité soit parfaitement générale et universellement applicable, il faut entendre par « événement » un particulier dont l'identité est déterminée par les limites de la zone spatio-temporelle qu'il occupe. Afin de ne pas rendre l'analyse circulaire, il faut éviter de donner au mot « transfert » un sens qui impliquerait déjà la causalité. Cela est possible en entendant dans ce contexte, par « transfert » simplement l'existence de *la même* quantité de grandeur conservée dans deux événements distincts. Cela rend la relation de transfert symétrique, alors que la

¹¹ Cela vaut autant pour d'autres analyses de la causalité qui ont été développées plus récemment, comme l'analyse probabiliste et l'analyse contrefactuelle. Pour une présentation succincte et critique de ces approches, cf. Kistler (2004) et Kistler (à paraître).

¹² Pour une justification de cette thèse, Cf. Kistler (1999).

relation causale et fondamentalement asymétrique. L'asymétrie de la causalité avec laquelle nous sommes familiers est fondée sur l'existence de processus irréversibles, en particulier de processus accompagnés d'une augmentation d'entropie¹³.

4. Responsabilité causale

Pour comprendre la causalité dans sa réalité complexe, il est utile de l'envisager comme ce qui est visé par les explications causales : les relations et processus causaux réels sont ceux qui rendent vraies les explications causales, comme par exemple : « la vitre s'est brisée parce qu'elle a été heurtée par le ballon ». Ce qui rend vraie cette explication (au cas où elle l'est), c'est un rapport entre deux faits, en l'occurrence le fait que le ballon lancé en direction de la vitre l'a heurtée et le fait que la vitre s'est brisée. Le rapport entre ces deux faits peut être analysé en distinguant les deux composantes fondamentales de la causalité : le transfert entre les événements et l'existence d'une dépendance nomique entre certaines de leurs propriétés. Le ballon lancé transfère de l'énergie et de la quantité de mouvement sur la vitre, et le bris est nomiquement dépendant des propriétés particulières du mouvement du ballon : ce n'est qu'à partir d'un certain poids du ballon et d'une certaine vitesse que sa rencontre avec la vitre entraîne le bris de celle-ci. Je propose d'appeler « responsabilité causale » le rapport entre les faits qui font l'objet d'une explication causale, et de l'analyser de la manière suivante :

une explication causale « e est G parce que c est F » est vraie, dans la mesure où il existe un fondement de sa vérité dans la réalité : le fait que c soit F est causalement responsable du fait que e est G si et seulement si c est cause de e en termes de transfert, c exemplifie F , e exemplifie G , et il existe une loi de la nature en vertu de laquelle les instanciations de F tendent à produire des instanciations de G .

Le succès et la plausibilité des théories traditionnelles de la causalité, telles que l'analyse déductive-nomologique, s'explique par le fait qu'elles visent l'une des composantes du rapport de responsabilité causale : elles peuvent être utilisées pour l'étude de la relation de dépendance nomique entre les propriétés F et G . Cependant, ces théories omettent l'aspect « mécanique » du processus local de transfert qui a lieu au niveau des événements. Sans que je puisse le montrer dans le détail ici, il s'avère que les contre-exemples à ces analyses proviennent de situations où il existe bien d'une dépendance nomique entre propriétés mais pas d'événements liés par un transfert.

5. La possibilité de la causalité mentale

Mettons cette analyse à l'épreuve pour nous aider à comprendre la possibilité de la causalité mentale, c'est-à-dire de l'impact réel de nos pensées sur le déroulement des événements physiques. Rappelons que le problème de l'exclusion causale se pose dans le cadre de la conviction physicaliste que tout événement mental est même temps un événement physique : cela signifie que les pensées, décisions, perceptions, et autres

¹³ Cf. Kistler (1999).

événements mentaux ont toujours pour support le cerveau d'une personne. Qui plus est, les propriétés du cerveau déterminent, en un sens qu'il est difficile à spécifier exactement, ces propriétés mentales. Selon certains auteurs, il s'agit d'une relation de détermination logique¹⁴. Il y a cependant aussi des raisons de penser que cette détermination soit *nomologique*, c'est-à-dire fonction de lois de la nature et non seulement de logique : les réductions inter-théoriques ne sont pas découvertes de manière purement *a priori* mais grâce à la recherche empirique, c'est-à-dire de manière *a posteriori*. Quoi qu'il en soit, que les propriétés mentales soient déterminées par le cerveau de manière logique ou par des lois de la nature, cette détermination garantit dans tous les cas une corrélation systématique entre propriétés cérébrales et mentales. On dit que les propriétés mentales « surviennent » sur les propriétés physiques du cerveau, en définissant le mot *survenance* ainsi : un ensemble de propriétés *A* (par exemple mentales) survient sur un ensemble de propriétés *B* (par exemple, physiques), si et seulement s'il est impossible que deux objets diffèrent à l'égard de leurs propriétés *A* sans différer à l'égard de leurs propriétés *B*¹⁵.

Le problème de la causalité mentale est de comprendre comment une propriété mentale *M* peut avoir un impact causal distinct de la propriété physique *P* sur laquelle elle survient. Certains auteurs, notamment Kim, soutiennent que cela est conceptuellement incohérent¹⁶. Cependant, le concept de responsabilité causale permet de montrer que la causalité mentale est cohérente. La démonstration de son existence réelle appartient aux sciences empiriques, c'est-à-dire à la psychologie et la neurophysiologie. La tâche de la philosophie se limite à la démonstration de sa possibilité ou cohérence conceptuelle.

Imaginons une situation simple où la causalité mentale est à l'œuvre : Jean pense à l'instant *t* que le bruit dans la rue gêne sa concentration (appelons cette propriété¹⁷ mentale « *M* »), *ce qui cause* sa décision de fermer la fenêtre à l'instant *t** (rappelons cette seconde propriété mentale « *M** »).

La base de survenance *P* de la propriété mentale *M* est une « propriété conjonctive », qui peut être décrite par une longue conjonction de descriptions d'états d'activation d'un grand nombre de cellules nerveuses et de synapses. *P* résulte d'une combinaison d'un grand nombre de propriétés locales instanciées dans le cerveau de Jean à l'instant *t*. La propriété mentale *M* peut être conçue comme une structure

¹⁴ Selon Yablo (1992), les propriétés cérébrales déterminent les propriétés mentales, au même sens que l'écarlate détermine le rouge, en d'autres termes, au sens où l'écarlate est une propriété déterminée, relativement à la propriété déterminable d'être rouge. Selon Jackson et Chalmers (2001), il est possible de déduire, de manière *a priori*, toute description d'états de choses mentaux à partir d'une (hypothétique) description complète du monde en termes physiques. Cela présuppose également que les états de choses mentaux soient déterminés par les états de choses physiques de manière *logique*.

¹⁵ Le concept de survenance a été systématiquement étudié par Kim (1993).

¹⁶ Kim (1992 ; 1998) soutient que seule la propriété *P* sous-jacente à *M* a un pouvoir causal et est donc seule « réelle » ; pour Kim, le prédicat « *M* » exprime bien un concept mais ne désigne aucune propriété réelle.

¹⁷ Lorsque je parle de *propriété* dans le présent contexte, j'entends toujours son « instance » à un endroit spatio-temporel, plutôt que la propriété universelle.

globale d'activation nerveuse présente dans l'activité du cerveau dans sa totalité. P détermine M en fonction de lois de la nature.

Il est cohérent d'envisager la possibilité que le fait que Jean ait M soit *causalement responsable* du fait que Jean ait M^* . Il le sera à plusieurs conditions. La condition de transfert est sans doute satisfaite de manière quasiment triviale, étant donné la proximité spatio-temporelle des événements où Jean possède M et M^* . Si notre analyse de la causalité en termes de responsabilité causale est correct, M joue un rôle causal distinct de celui de P , si et seulement si 1) il n'existe aucune loi reliant P à P^* , et 2) il existe une loi qui relie M à M^* .

La loi entre M et M^* est une « loi de système »¹⁸ : son domaine de validité est restreint à un type particulier de système, les humains vivants. C'est en vertu de cette loi et en vertu de l'existence d'une relation causale entre c (l'événement auquel Jean possède les propriétés P et M) et e (l'événement auquel Jean possède les propriétés P^* - M^*) par transfert, que (le fait que c soit) M est causalement responsable de M^* (du fait que e soit M^*).

On pourrait objecter à la cohérence du scénario envisagé qu'il a pour conséquence que le fait que e est P^* n'a pas de cause du tout à l'instant t , au sens où il n'existe aucun fait qui en soit causalement responsable. Or, il s'avère que le fait que e soit P^* a bien une cause à l'instant t ; seulement, il n'a pas de cause complète au niveau neurophysiologique à l'instant t . Voici pourquoi : M aurait pu être réalisée par P_1, P_2, \dots au lieu de P . Mais étant donné que M est actuellement réalisée par P , il est concevable qu'une partie de la conjonction complexe P constitue une contrainte d'invariance du système dont la stabilité permet à la loi de niveau mental M - M^* de mener l'évolution du système de M à M^* . C'est cette contrainte du système qui reste stable pendant l'évolution de M à M^* , qui détermine laquelle parmi de nombreux états physiques P^* , P_1^*, P_2^*, \dots le système occupera à l'instant où il possède la propriété mentale M^* . Dans un tel scénario, P^* est déterminée partiellement par P et partiellement par M , quoiqu'il n'existe ni loi causale « vers le bas » entre M et P^* ni loi purement neurophysiologique entre P et P^* . Étant donné l'état initial P , la loi qui gouverne l'évolution du système au niveau de M contraint le système d'évoluer non seulement vers un état M^* , mais vers un état M^* compatible avec les conditions aux limites P , à savoir vers l'état P^* .

Il est possible que de nombreux états P_1^*, \dots, P_n^* soient compatibles avec ces contraintes. Dans ce cas, même la conjonction de M et de P n'est pas causalement suffisante pour P^* . Alors *aucun* fait portant sur le système à l'instant t de l'événement c (P - M) est causalement responsable du fait que l'effet e (P^* - M^*) possède la propriété neuronale P^* . Tous les faits qui sont causalement responsables de P^* (autrement dit, déterminent P^* causalement) sont des faits portant sur des événements qui précèdent P^* par un laps de temps plus bref. Cela est propre aux systèmes chaotiques en général : il n'existe aucune loi déterministe stricte qui permettrait de prévoir la distribution précise des nuages au-dessus de Paris le 31 octobre 2005, à partir de la distribution des nuages au même endroit un mois auparavant. En d'autres termes, il n'existe aucun fait portant sur le 30 septembre qui est causalement responsable de la distribution précise des nuages un mois plus tard.

¹⁸ L'expression est de Schurz (2002).

Conclusion

Le concept de causalité est indispensable à la poursuite du projet d'unification des connaissances scientifiques. Nous avons vu qu'il est possible de construire un concept adéquat à cet objectif et pourtant ancré dans le sens commun. Cela présuppose la découverte de l'articulation entre deux composantes du concept de causalité : la causalité est, d'une part, un processus qui se déroule au niveau des événements particuliers et qui est fondé sur un transfert matériel. Mais elle est, d'autre part, le fondement des explications causales. Pourquoi la vitre s'est-elle brisée ? En général, pourquoi l'événement e a-t-il la propriété G ? Parfois, il peut être éclairant d'indiquer la cause d'un événement au sens de la source matérielle du processus qui a conduit à cet événement. Mais la plupart du temps, la seule indication du processus de transfert ne suffit pas à expliquer complètement un fait portant sur l'événement « effet ». Il ne suffit pas de mentionner le vol du ballon en direction de la vitre. Il faut en outre désigner la loi (ou les lois) de la nature qui permettent de comprendre que la propriété (ou les propriétés) F possédée par l'événement « cause » rend inévitable que l'événement « effet » e possède G .

Cette analyse nous donne les moyens conceptuels de clarifier l'un des principaux problèmes de la philosophie contemporaine de l'esprit : comment l'esprit peut-il intervenir causalement dans le cours des événements physiques, alors que chaque événement physique semble posséder une cause complète au niveau physique même ? Si l'esprit est réel au sens de pouvoir faire une différence parmi les événements physiques, la physique (et même la neurophysiologie) devrait être incapable de fournir des explications complètes des mouvements corporels réalisant les actions d'agents humains. Or la physique ne semble jamais avoir besoin de recourir à des concepts psychologiques ni même biologiques ou chimiques. Cependant, nous avons vu qu'il est conceptuellement cohérent d'envisager la possibilité que le cerveau humain soit un système chaotique où il n'existe aucune cause complète d'un fait portant sur son cerveau à un instant donné, à un instant suffisamment lointain dans le passé. La compétence de la philosophie s'arrête à la démonstration qu'il est cohérent d'envisager que cette indétermination sur le plan neurophysiologique s'accompagne d'une détermination sur le plan cognitif ou mental. Nous devons laisser aux sciences empiriques, à savoir à la psychologie et à la neurophysiologie, le soin de découvrir si cette possibilité correspond à la réalité, autrement dit si c'est bien nos pensées, plutôt que l'état d'activation de nos neurones, qui déterminent nos actes.

RÉFÉRENCES BIBLIOGRAPHIQUES

- BROMBERGER, S., 1966, « Why-Questions », in R.G. Colodny (ed.), *Mind and Cosmos : Essays in Contemporary Science and Philosophy*, Vol 3, Pittsburgh, The Center for Philosophy of Science; repr. in S. Bromberger, *On What We Know We Don't Know*, Chicago, University of Chicago Press, et Stanford, CSLI, 1992 : 75-100.
- CHALMERS, D., JACKSON F., 2001, « Conceptual Analysis and Reductive Explanation », *Philosophical Review*, 110 : 315-360.
- DAVIDSON, D., 1970, « Mental Events », in D. Davidson, *Essays on Actions and Events*, Oxford, Clarendon Press 1980 (trad. fr. de P. Engel : *Actions et événements*, Paris, PUF, 1993).
- EELLS, E., 1991, *Probabilistic Causality*, Cambridge University Press.
- GOLD, I. et STOLJAR D., 1999, « A Neuron Doctrine in the Philosophy of Neuroscience », *Behavioral and Brain Sciences*, vol. 22 (5) : 585 – 642.
- HEMPEL, C. G., OPPENHEIM, P., 1948, « Studies in the Logic of Explanation », in Hempel, C. G., *Aspects of Scientific Explanation*, New York, The Free Press, 1965.
- JAEGWON, K., 1993b, *Supervenience and mind*, Cambridge, Cambridge University Press.
- JAEGWON, K., 1998, *Mind in a Physical World*, Cambridge (Mass.), MIT Press.
- KISTLER, M., 1999, *Causalité et lois de la nature*, Paris, Vrin, Collection Mathesis, 1999. (Trad. anglaise, à paraître, Londres, Routledge, 2006).
- KISTLER, M., 2004, « La causalité dans la philosophie contemporaine », *Intellectica*, 38 : 139-185.
- KISTLER, M., à paraître, « La causalité comme transfert et dépendance nomique », *Philosophie*.
- LEWIS, D., 1986, « Causation », in D. Lewis, *Philosophical Papers, vol. II*, New York, Oxford University Press.
- POPPER, K., 1934, *Logik der Forschung*, 3^e éd. augmentée, Tübingen, J.C.B. Mohr, 1969 (trad. par N. Thyssen-Rutten et P. Devaux : *La logique de la découverte scientifique*, Paris, Payot, 1973).
- RESCORLA, R. A., 1988, « Pavlovian conditioning: It's not what you think it is », *American Psychologist* 43 : 151-160.
- SCHURZ, G., 2002, « Ceteris paribus Laws : Classification and Destruction », *Erkenntnis* 57 : 351-372.
- SKLAR, L., 1993, *Physics and Chance*, Cambridge, Cambridge University Press.
- YABLO, S., 1992, « Mental Causation », *Philosophical Review* 101 : 245-280.