

Max Kistler

LA CAUSALITE COMME TRANSFERT ET DEPENDANCE NOMIQUE

in : *Philosophie* 89 (2006), p. 53-77.

La causalité est l'un de ces concepts qui semblent conforter dans leur opinion ceux qui doutent de la possibilité d'accomplir de réels progrès en philosophie : le débat sur la nature de la causalité semble pouvoir illustrer l'une de ces batailles stériles qui caractérisent selon certains la métaphysique, où plusieurs doctrines campent sur leur positions, sans que leur confrontation puisse jamais devenir constructive et donner lieu à un progrès.

Il y a en effet, dans l'histoire de la philosophie mais aussi dans le débat contemporain, une telle pléthore de théories sur la causalité, et si peu de consensus, que l'on a, semble-t-il, toutes les raisons d'adopter à son égard une position « éliminativiste » ou relativiste. Selon la position *éliminativiste* dont la longue histoire a connu un point culminant au siècle passé chez Russell¹, rien ne correspond aux jugements causaux du sens commun, dans la réalité telle qu'elle est décrite par la science fondamentale ; par conséquent, le concept de causalité, s'il a sa place dans le schème conceptuel superficiel du sens commun, doit être banni d'une conception du monde qui se veut respectueuse des connaissances scientifiques.

Selon une position *relativiste*, il existe des relations causales, mais il y a autant de genres de causalité que de théories sur la causalité. Parmi les analyses qui font actuellement l'objet des recherches les plus intenses, mentionnons : 1) la doctrine de l'empirisme logique classique selon laquelle la relation causale est équivalente à une relation d'explication ayant une forme déductive² ; 2) l'analyse contrefactuelle selon laquelle deux événements *c* et *e* sont reliés comme cause à effet si et seulement si : si *c* n'avait pas eu lieu, *e* n'aurait pas eu lieu³ ; 3) l'analyse probabiliste selon laquelle un premier événement de type A est la cause d'un second événement de type B si la probabilité conditionnelle que B ait lieu étant donné que A ait eu lieu est plus grande que la probabilité que B ait lieu étant donné que A n'ait pas eu lieu⁴ ; 4) l'analyse de l'agentivité selon laquelle les causes sont les événements qui peuvent servir de moyens pour obtenir leurs effets : un événement de type A est cause d'un événement de type B dans la mesure où, pour un agent qui a l'intention d'obtenir B, il est rationnel de faire en sorte que A arrive⁵ ; 5) l'analyse selon laquelle cause et effet sont reliés par un processus qui est une chaîne d'événements capable de transmettre un « caractère »⁶ ; 6) l'analyse mécaniste selon laquelle les causes sont reliées à leurs effets par un mécanisme qui

¹ Cf. l'article traduit dans ce volume, *supra*, p. XX-XX.

² Carl Gustav Hempel and Paul Oppenheim, « Studies in the Logic of Explanation », repr. in Carl Gustav Hempel, *Aspects of Scientific Explanation*, The Free Press, New York, 1965, p. 245-290. Selon cette conception, la relation causale est équivalente à la relation entre la prémisse particulière d'une explication qui mentionne les « conditions initiales » et l'*explanandum*, c'est-à-dire ce qu'il s'agit d'expliquer.

³ David Lewis, « Causation », in : *Philosophical Papers, Vol. II*, Oxford University Press, New York, 1986, p. 159-213.

⁴ Ellery Eells, *Probabilistic Causality*, Cambridge University Press, Cambridge 1991.

⁵ Douglas Gasking, « Causation and Recipes », *Mind* 64 (1955), 479-487; Georg Henrik von Wright, *Explanation and Understanding*, Cornell University Press, Ithaca, N.Y. 1971; Huw Price, « Agency and Causal Asymmetry », *Mind* 101 (1992), 501-520; Peter Menzies and Huw Price, « Causation as a Secondary Quality », *British J. Phil. of Science* 44 (1993), 187-203.

⁶ Hans Reichenbach, *The Direction of Time* (1956), University of California Press, Berkeley, 1991, p. 198 sq.; Wesley Salmon, *Scientific Explanation and the Causal Structure of the World*, Princeton University Press, Princeton 1984. Le terme « caractère » correspond à l'anglais « mark » et à l'allemand « Merkmal » : une propriété permanente et reconnaissable qui peut être utilisée comme signal.

produit l'effet étant donné la cause grâce à l'articulation de ses parties et des lois qui gouvernent les interactions entre ces parties⁷ ; 7) l'analyse de la continuité qualitative selon laquelle cause et effet sont les termes d'un processus qui est une « ligne de monde » caractérisée par la constance ou le changement continu de toutes ses propriétés⁸ ; 8) l'analyse du transfert selon laquelle les causes transfèrent sur leurs effets des quantités individuelles de grandeurs conservées, telle que l'énergie⁹.

Face à cette multitude d'approches, la position que j'ai plus haut qualifiée de « relativiste » consiste à admettre que toutes ces théories, ainsi que leurs variantes, « ont raison » en ce sens que chacune exprime *une facette* de la réalité de la causalité. Cependant, selon cette position relativiste, aucune théorie n'a pour vocation d'être une analyse universelle qui s'applique à *toutes* les relations causales : il serait naïf de penser qu'il existe une essence de la causalité.

On peut distinguer deux variantes d'un tel relativisme par rapport à la causalité : la première variante conteste l'idée que les analyses de la causalité ont pour but de décrire une réalité indépendante, alors que la seconde accepte l'idée qu'il s'agisse d'hypothèses descriptives, mais conteste que leur champ d'application soit universel. Selon la première variante, les différentes théories de la causalité sont autant de « schèmes » ou « instruments conceptuels » qui donnent un sens aux jugements causaux, ou encore simplement différentes « façons de parler » : à ce titre, elles n'ont pas de conditions littérales de vérité, et l'on trahit une confusion conceptuelle lorsqu'on demande laquelle « a raison » ou laquelle décrit correctement « la réalité ». L'idée de vouloir comparer ces théories pour déterminer celle qui est correcte révèle une confusion quant au type de discours dont il s'agit : la question de la vérité ne peut être posée qu'à propos d'hypothèses descriptives, mais non à propos d'instruments ou de schèmes conceptuels.

Selon la seconde variante, ce sont bien des théories descriptives, mais chacune a son domaine d'application propre : la conception de la causalité en tant que processus matériel « s'appliquerait » par exemple aux relations causales entre événements *physiques*, alors que la causalité probabiliste s'appliquerait (entre autres) aux relations causales *économiques* et *épidémiologiques*, et la théorie de la manipulabilité s'appliquerait aux relations causales pertinentes pour nos délibérations *pratiques* et dans des contextes *juridiques*.

Dans cet article, je voudrais défendre une forme d'« œcuménisme » bien distincte de ces formes de relativisme. Je commencerai par constater que la prétention de chacune des conceptions majeures de la causalité mentionnées plus haut à la validité universelle se heurte à des contre-exemples. Cependant, il ne me semble pas pour autant nécessaire d'abandonner le projet de construire un concept universel de causalité qui permette de rendre compte des conditions de vérité de nos jugements causaux dans tous les domaines.

Contre-exemples

⁷ Stuart S. Glennan, « Mechanisms and the Nature of Causation », *Erkenntnis* 44 (1996), p. 49-71.

⁸ Bertrand Russell, *Human Knowledge : Its Scope and Limits* (1948), London, Allan and Unwin, 5e édition, 1966; Phil Dowe, « Wesley Salmon's Process Theory of Causality and the Conserved Quantity Theory », *Phil. of Science* 59 (1992), p. 195-216; Phil Dowe, *Physical Causation*, Cambridge University Press, 2000.

⁹ Max Kistler, « Reducing Causality to Transmission », *Erkenntnis* 48 (1998), pp. 1-24; Max Kistler, *Causalité et lois de la nature*, Paris, Vrin, Collection Mathesis, 1999. Le fait même d'énumérer ces approches, sans essayer d'y déceler le résultat de choix stratégiques, reflète le point de vue relativiste : plus loin, j'essaierai de les classer en fonction de tels choix, tout en fournissant des raisons d'en préférer certaines à d'autres.

Strawson¹⁰ a distingué deux conceptions de la métaphysique. Selon la conception *descriptive*, la métaphysique a pour tâche de révéler les structures les plus fondamentales de notre conception naïve de la réalité. L'analyse métaphysique a pour but de mettre à jour des présupposés sous-jacents à nos intuitions et à nos jugements. En revanche, la conception *révisionniste* se propose de développer une conception de la structure fondamentale de la réalité qui ne se plie pas nécessairement à nos jugements intuitifs : la métaphysique révisionniste choisit, comme critères d'adéquation, d'autres contraintes que l'accord avec le sens commun, en particulier l'accord avec les théories scientifiques les plus avancées. Aussi claire que soit la distinction strawsonienne, il n'est pas aisé de classer les recherches sur le concept de causalité selon ces deux catégories. En un sens, le débat sur la nature de la causalité dans la philosophie contemporaine ne peut être compris que sur l'arrière-plan du présupposé que la tâche est descriptive au sens de Strawson. C'est la seule manière de comprendre que les défenseurs des diverses doctrines s'accordent en général sur le critère d'adéquation qui permet de trancher, à savoir l'accord du résultat de l'analyse avec les jugements « naïfs » ou « pré-théoriques » déterminant quelle relation est causale et laquelle ne l'est pas. Dans la mesure où l'on accepte ce critère, on accepte l'idée que la relation causale a une nature objective qu'il s'agit de découvrir ; on suppose, en acceptant la réfutation de certaines analyses par les jugements du sens commun, que le concept de causalité n'est pas un simple « instrument » dont on serait libre de choisir le contenu par stipulation.

Cependant, ce présupposé - que nous assumons comme choix méthodologique délibéré - ne garantit pas qu'il soit possible de développer l'analyse de la causalité sur ce seul niveau descriptif, pour la simple raison qu'elle peut révéler que certaines intuitions ont des conséquences incompatibles entre elles. L'analyse se trouve dans ce cas dans la nécessité de restaurer la cohérence de notre conception globale du monde, au prix d'une certaine dose de révisionnisme : on cherche à construire des concepts qui s'accordent le plus possible avec l'intuition tout en nous permettant de préserver la cohérence de l'ensemble de nos jugements sur les relations causales.

Je propose de montrer d'abord que toutes les conceptions qui ont été proposées pour analyser la relation causale sont confrontées à des contre-exemples : pour chacune d'entre elles, il y a ou bien des relations (ou processus) qui sont intuitivement causales mais que la théorie en question n'autorise pas à considérer comme telles, ou bien des relations qui ne sont intuitivement pas causales mais que la théorie conduit à considérer comme telles. Reconnaître ces cas comme des contre-exemples à ces théories, cela signifie que l'on accepte le sens commun comme critère d'adéquation, et que l'on poursuit le projet d'analyse dans le cadre de la métaphysique descriptive.

Cependant, je voudrais anticiper sur un résultat de l'analyse des contraintes qu'impose le sens commun aux théories de la causalité : les contre-exemples montrent que l'on peut extraire des jugements du sens commun deux manières fondamentalement différentes de concevoir la causalité, que j'appellerai la conception « explicative » (ou « aristotélicienne ») et la conception « mécaniste ». Selon la première, la relation causale est essentiellement une relation *explicative*, alors que selon la seconde, elle est essentiellement fondée sur un processus de transmission matérielle. Il s'avérera que les contre-exemples qui semblent réfuter les conceptions explicatives (les conceptions déductive-nomologique, contrefactuelle et probabiliste : de 1 à 3 dans la liste qui précède) ont pour source des intuitions de type mécaniste, alors que les contre-exemples que l'on peut opposer aux conceptions mécanistes (de 5 à 8 dans la liste qui précède) se nourrissent d'intuitions de type explicatif. Nous serons donc amenés à proposer une théorie œcuménique ou « hybride » de la causalité qui distingue

¹⁰ P.F. Strawson, *Individuals*, London, Methuen, 1959, Introduction. Trad. A. Shalom et P. Drong, *Les individus*, Seuil 1973 ; P.F. Strawson, *Analysis and Metaphysics. An Introduction to Philosophy*, Oxford, Oxford University Press, 1992, chap. 3.

deux composantes, à la fois dans le concept de causalité que nous utilisons dans nos jugements et parmi les faits dans le monde qui rendent ces jugements vrais. Il s'agit donc d'une part de métaphysique descriptive puisque le but est de reconstruire l'ensemble de nos jugements de manière à les rendre cohérents, mais l'analyse finale contient aussi une part révisionniste puisqu'elle introduit une distinction qui n'est pas explicite dans le sens commun, celle entre une composante mécaniste et une composante explicative du concept de causalité.

Le modèle « déductif-nomologique » et ses contre-exemples

La doctrine qui a dominé le débat sur la nature de la causalité, au sein de la philosophie de tradition analytique, depuis les années 1930 jusqu'aux années 1970, est l'assimilation, par l'empirisme logique, de la relation causale à la relation d'explication : la relation de cause à effet est équivalente à la relation entre les conditions initiales et l'*explanandum* dans une explication déductive-nomologique¹¹. Cette conception est une héritière lointaine de la conception aristotélicienne : comme cette dernière, elle conçoit la cause et l'effet comme des entités propositionnelles, exprimées par des phrases descriptives. Seules des propositions, porteuses de valeurs de vérité, peuvent jouer les rôles de prémisses et de conclusion dans une explication, ici conçue comme un argument déductif qui fait un appel essentiel aux lois de la nature (d'où l'abréviation D-N, pour « déductif-nomologique »). Mais elle est également héritière de la doctrine empiriste humienne dans la mesure où les lois de la nature qui fondent le lien explicatif entre cause et effet sont elles-mêmes conçues, par l'empirisme logique, comme des régularités universelles. La conception D-N de la causalité partage donc avec Hume la thèse selon laquelle le lien entre une cause et un effet particuliers dépend de l'existence d'une régularité générale. Cependant, à la différence de Hume, la généralité des lois n'est plus conçue sur le plan psychologique de l'expérience : pour qu'un jugement causal entre deux événements particuliers soit vrai, il suffit qu'une régularité nomique *existe* objectivement sans qu'il soit nécessaire qu'un sujet en ait fait l'*expérience*.

De nombreux contre-exemples à la conception D-N de la causalité ont été découverts¹². Les *lois de coexistence* permettent de construire des explications de forme D-N où les prémisses ne font pas référence aux causes de la conclusion. Ce sont des lois exprimant l'existence d'une dépendance nomique entre les propriétés que possède un objet (ou un « système ») à un moment donné. On peut expliquer la température T d'un échantillon donné d'un gaz (suffisamment) « idéal¹³ », à partir de sa pression p et du volume V qu'il occupe si l'on utilise pour cela la loi dite du gaz idéal, ou loi de Boyle-Mariotte : $pV = nRT$

(Dans cette équation, « n » représente la quantité de gaz, en nombres de moles ; R est une constante universelle).

¹¹ L'*explanandum* est ce qu'il s'agit d'expliquer ; l'*explanans* est ce qui explique. Les adjectifs « nomique » et « nomologique » sont dérivés du grec « nomos », la loi : j'utilise « nomique » pour désigner ce qui est en rapport aux lois de la nature, et « nomologique » pour ce qui est en rapport avec le discours sur les lois de la nature, en particulier les énoncés exprimant ces lois. Cf. Anouk Barberousse, Max Kistler, Pascal Ludwig, *La philosophie des sciences au XX^e siècle*, Paris, Flammarion, 2000, chap. 5.

¹² Pour une liste plus riche de tels contre-exemples, cf. P. W. Humphreys, « The Causes, Some of the Causes, and Nothing But the Causes », in *Minnesota Studies in the Philosophy of Science, vol XII: Scientific Explanation*, University of Minnesota Press, Minneapolis 1989, 283-306, 300/1; Wesley Salmon, *Four Decades of Scientific Explanation*, University of Minnesota Press, Minneapolis 1990, 46-50; M. Kistler, « Causation in contemporary analytical philosophy », *Quaestio-Annuaire d'histoire de la métaphysique*, vol. 2, éd. C. Esposito et P. Porro, Turnhout (Belgique), Brepols, 2002, p. 635-668 ; M. Kistler, « Conceptions contemporaines de la causalité », *Intellectica* 38 (2004/1), p. 139-185.

¹³ Un gaz réel s'approche du gaz idéal dans la mesure où ses composantes microscopiques individuelles (atomes ou molécules) occupent un espace petit (relativement à l'espace occupé par le gaz) et ont des interactions faibles entre eux.

Dans ce cas, *l'explanandum* (le fait que le gaz ait une certaine température), ainsi que la condition initiale qui fait partie de *l'explanans* (le fait que le gaz ait une certaine pression et occupe un certain volume), sont bien des phénomènes empiriques entre lesquels il pourrait en principe y avoir une relation causale. Mais l'intuition s'oppose à l'idée qu'il y ait là une relation causale, pour la raison qu'il manque à cette relation *l'antériorité temporelle* et la *non-coïncidence spatiale* de ses termes, ainsi que *l'antisymétrie* caractéristique de la causalité¹⁴. Si *c* est cause de *e*, alors *e* n'est pas cause de *c*. La température et la pression sont des propriétés que le gaz possède au même temps et au même endroit, alors que cause et effet semblent intuitivement devoir être distincts dans l'espace et dans le temps, de telle sorte que la cause précède l'effet et lui soit contiguë (ce qui présuppose que cause et effet ne coïncident pas entièrement dans l'espace, mais ne font que « se toucher »). Or il semble que notre concept de causalité obéisse à cette contrainte : il semble inconcevable que, étant donné deux événements distincts *A* et *B*, *A* soit à la fois cause de *B* et *B* cause de *A*. Par ailleurs, dans la mesure où la relation causale est transitive¹⁵, si l'on voulait admettre que la relation causale puisse ne pas être asymétrique, on serait confronté à la conséquence qu'il y a des « causes de soi » parmi les phénomènes naturels, ce qui est difficilement acceptable : si *A* est cause de *B* et *B* cause de *A*, alors, en vertu de la transitivité, *A* est cause de *A* et *B* cause de *B*.

Les *explications réductrices* constituent un autre type important d'explications scientifiques non-causales. La température et la pression peuvent être réduites à des fonctions des valeurs de certaines propriétés des constituants microscopiques des gaz auxquelles la loi s'applique.¹⁶ Pourtant, l'intuition rejette l'idée que les termes de cette explication réductrice soient les termes d'une relation causale, et ce pour les mêmes raisons que dans le cas de l'explication de la température par la pression (et le volume), en vertu de la loi de Boyle-Mariotte : les molécules avec leurs vitesses se trouvent au même endroit au même temps que le gaz dont ces vitesses permettent d'expliquer la température et la pression ; elles ne peuvent donc pas en être la cause, étant donné qu'une cause doit être spatialement et temporellement distincte de son effet.

Ces contre-exemples montrent la présence, dans le concept intuitif de causalité, d'un élément irréductible à la composante explicative : il existe des explications que l'intuition ne considère pas comme causales alors qu'elles se plient au modèle D-N. Ces cas autorisent à continuer de considérer le modèle D-N comme ce qu'il devait, selon ses concepteurs Hempel et Oppenheim, à l'origine être avant tout : un modèle adéquat de *l'explication*. Cependant, ils donnent des raisons de contester qu'il s'agisse en même temps d'un modèle adéquat de la causalité. On pourrait présenter la situation aussi en ces termes : il y a un sens faible du mot « explication » (et du concept qu'il exprime) où la relation d'explication a la même extension que la relation de *prédiction*. Dans la mesure où la chute du baromètre permet de prédire la tempête, il est légitime de dire que la chute du baromètre explique la tempête. Mais il existe un autre sens plus fort du mot « explication » qui donne lieu à l'intuition contraire : on a l'impression que la chute du baromètre n'explique pas vraiment la tempête, au sens où la chute de la pression de l'air atmosphérique permet de la faire. Dans cette intuition s'exprime le concept mécaniste de la causalité : selon cette deuxième intuition, seules les explications *causales* sont vraiment satisfaisantes ; et une explication doit être plus qu'une déduction D-

¹⁴ Salmon remarque que cette loi « ne dit absolument rien des relations causales entre » la pression, le volume et la température d'un gaz. Lorsque ces quantités subissent des changements, cela est dû à des causes, mais « la loi elle-même est entièrement muette à l'égard de ces considérations causales » (Salmon, *Scientific Explanation*, p. 136).

¹⁵ Ce qui est controversé : de nombreux auteurs considèrent qu'il va de soi que si *A* est cause de *B* et si *B* est cause de *C* alors, *A* est cause de *C*. Ainsi, Lewis (« Causation », p. 167) prend soin d'introduire dans sa théorie une clause qui tient compte de l'intuition selon laquelle la causalité est une relation transitive. D'autres contestent que nos intuitions obéissent à cette contrainte. Cf. l'article de Keil, dans ce numéro.

¹⁶ Cf. Lawrence Sklar, *Physics and Chance*, Cambridge, Cambridge University Press, 1993.

N : elle doit faire référence à une relation causale au sens mécaniste, ce qui est le cas de la relation entre chute de la pression de l'air et tempête, mais non entre chute du baromètre et tempête.

L'analyse contrefactuelle et ses contre-exemples

La théorie contrefactuelle esquissée par Hume¹⁷ et développée récemment par David Lewis¹⁸, conçoit la relation causale en termes de dépendance contrefactuelle entre propositions. Disons que le fait que Jean fume soit la cause du fait qu'il développe un cancer du poumon. Selon la conception contrefactuelle, la vérité de ce jugement causal est fondée sur l'existence d'une relation de dépendance contrefactuelle entre les propositions¹⁹ qui nient ces faits : si Jean n'avait pas fumé, il n'aurait pas développé de cancer du poumon²⁰. L'idée fondamentale de la conception contrefactuelle est donc que les effets d'un événement-cause *C* donné sont les événements *E* dont l'existence *dépend* de cette cause. La dépendance est interprétée en termes de « mondes possibles » : on conçoit une situation possible, différente de la situation actuelle qui contient *C*, où *C ne s'est pas* produite, mais qui est par ailleurs aussi semblable que possible à la situation actuelle (dans laquelle la cause s'est produite). On dira alors de tous les événements *E* qui ne se produisent pas dans ce monde possible, alors qu'ils se produisent dans le monde actuel, qu'ils dépendent contrefactuellement de *C*. « *E* dépend contrefactuellement de *C* signifie » : « si *C* ne s'était pas produit, alors *E* ne se serait pas produit non plus ». Ensuite, Lewis construit la relation causale comme relation transitive fondée sur un enchaînement de dépendances contrefactuelles. *C* est cause de *E*, si et seulement s'il existe un ensemble d'événements intermédiaires entre *C* et *E*, *C*₁, *C*₂, *C*₃ ..., *C*_{*n*},

¹⁷ Selon la seconde définition de la cause par Hume, « un objet est la cause d'un autre » signifie que « si le premier objet n'avait pas existé, le second n'aurait jamais existé » (David Hume, *An Enquiry Concerning Human Understanding*, in *Enquiries Concerning Human Understanding and Concerning the Principles of Morals* (1748), L.A. Selby-Bigge and P.H. Nidditch (eds.), troisième éd., Clarendon Press, Oxford 1975, p. 76 ; trad. par A. Leroy, *Enquête sur l'entendement humain*, Paris, Aubier, 1947, p. 124/5).

¹⁸ Cf. David Lewis, *Counterfactuals*, Oxford, Basil Blackwell, 1973.

¹⁹ Lewis conçoit les événements comme des « propriétés de régions spatio-temporelles » (David Lewis, « Events », in : *Philosophical Papers, Vol. II*, p. 241-269 ; p. 245), ce qui entraîne une correspondance entre les événements et les propositions. Si *c* est un événement, alors la proposition qui fait l'objet de dépendance contrefactuelle est simplement « la proposition que *c* existe (ou a lieu) » (Lewis, « Causation », p. 159). Selon la conception « fine » des événements adoptée par Lewis, le fait que Jean dise « Bonjour ! » à haute voix et le fait que Jean dise « Bonjour ! » sont des faits différents correspondant à deux propositions. Dans la conception de Lewis, il leur correspond également deux événements. La raison de choisir cette conception est qu'elle permet de considérer que ce sont les causes et les effets d'un événement qui déterminent son identité. « La vraie raison pour laquelle nous avons besoin des deux événements [...] est qu'ils diffèrent causalement » (Lewis, « Events », p. 255). Il y a deux événements concernant Jean puisque le fait qu'il dise « Bonjour ! » à haute voix a des causes et des effets que n'a pas le fait qu'il dise « Bonjour ! » simplement. C'est le fait que Jean ait été tendu qui a causé le premier mais non le second événement ; et le premier mais non le second a pour effet que Fred le salue à son tour. La conception des événements de Lewis est semblable à celle de Kim (Jaegwon Kim, « Events as Property Exemplifications », in : M. Brand and D. Walton (eds.), *Action Theory*, Dordrecht, Reidel, 1976, p. 159-177) qui les conçoit comme « des exemplifications de propriétés par des substances à un temps » (Kim, p. 160); elle s'oppose aux conceptions « grossières » de Davidson, « The Individuation of Events » (1969), in : Donald Davidson, *Essays on Actions and Events*, Oxford, Clarendon Press, 1980; trad. P. Engel : *Actions et événements*, Paris, P.U.F., 1993) et Quine (W.V.O. Quine, « Events and Reification », in : Ernest LePore et Brian McLaughlin (eds.), *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, Oxford, Basil Blackwell, 1985) pour lesquels un événement est un particulier possédant de nombreuses propriétés et dont l'identité est déterminée par la zone spatio-temporelle qu'il occupe. La diversité des conceptions des événements peut parfois entraîner des confusions, surtout étant donné que ce que Kim et Lewis appellent les « événements » correspond à ce que d'autres appellent les « faits », pour les distinguer des événements (grossièrement conçus). Cf. Max Kistler, « Causes as events and facts », *Dialectica* 53 (1999), p. 25-46.

²⁰ Cela peut être exprimé par la formule : $\neg Fj \square \rightarrow \neg Pj$, où « $\square \rightarrow$ » représente la dépendance contrefactuelle.

tels que chaque membre de la chaîne dépend contrefactuellement de son prédécesseur : C_1 dépend contrefactuellement de C , C_2 de C_1 , etc., et E dépend contrefactuellement de C_n .

Pour reprendre notre exemple, la théorie contrefactuelle analyse le jugement selon lequel le fait que Jean fume cause le fait qu'il développe un cancer du poumon, comme fondé sur la dépendance contrefactuelle (éventuellement indirecte) entre ces faits : il existe un certain nombre d'événements intermédiaires C_1, C_2, \dots, C_n , tels que : si Jean n'avait pas fumé, C_1 ne se serait pas produit, etc., et si C_n ne s'était pas produit, Jean n'aurait pas développé de cancer du poumon.

Il existe un certain nombre de contre-exemples à cette analyse. De nombreux mécanismes dans les êtres vivants sont « redondants » : admettons que, dans une espèce S , le mécanisme A produit dans les circonstances normales un certain effet E , mais que les organismes S disposent aussi d'un second mécanisme B « de sécurité » (ou « *back-up* ») qui produit le même effet dans les circonstances où le premier mécanisme A est défaillant²¹. Pensons à une situation concrète où le premier mécanisme A produit E . La présence du second mécanisme de sécurité détruit toute dépendance contrefactuelle entre la cause et l'effet : il n'est pas vrai que, si A n'avait pas fonctionné, E ne se serait pas produit, car dans ce cas, E aurait néanmoins eu lieu, grâce au mécanisme B . Cela montre que l'exigence de dépendance contrefactuelle est trop forte. L'intuition juge qu'il y a bien eu production causale de E par A , sans que l'effet dépende contrefactuellement de la cause. On dit que le processus par lequel A cause E empêche (ou « préempte ») B de causer E .

Pour un contre-exemple d'un genre différent, considérons une situation concrète où la chute du baromètre C précède une tempête T . C ne cause pas T , mais dans la mesure où le baromètre est fiable, il est vrai que : si le baromètre n'avait pas chuté il n'y aurait pas eu de tempête²². Les dépendances nomiques exprimées dans les lois de coexistence, telles que la loi de Boyle-Mariotte sont à l'origine d'autres contre-exemples du même genre : si le gaz n'avait pas été sous la pression p , il n'aurait pas été à la température $T = pV/nR$. Ces cas montrent que la dépendance contrefactuelle est en général une exigence trop faible. Il ne suffit pas qu'un fait soit contrefactuellement dépendant d'un autre fait pour que le second soit la cause du premier.

Nous jugeons intuitivement qu'il ne s'agit pas de cas de causalité. Cela semble indiquer que nous n'admettons l'existence d'une relation causale que là où il y a une influence mécanique, ou un processus matériel qui se déroule entre deux événements distincts dans l'espace et le temps, qui commence dans l'événement cause et se termine dans l'événement effet. Dans le cas de « préemption », le mécanisme A constitue un tel processus qui aboutit à E , alors que le processus B n'aboutit pas. Lorsqu'au contraire, dans un cas de « surdétermination », deux processus aboutissent à un même événement effet E , alors l'intuition est en accord avec la conception mécaniste, pour juger que les deux processus sont causes de E . De manière analogue, la théorie mécaniste donne une explication du jugement intuitif selon lequel il n'y a pas de relation causale entre la chute du baromètre et la tempête : il n'y a aucun processus, ou mécanisme, qui prendrait son origine dans la première pour se terminer dans la seconde. Enfin, il ne peut pas y avoir de processus entre la pression et la température d'un gaz, ni entre les vitesses des molécules qui le composent et l'une de ses

²¹ Michel Morange, *La part des gènes*, Paris, O. Jacob, 1998, montre que c'est le cas de nombreux mécanismes génétiques : si le mécanisme A exploitant un gène (ou des gènes) de type 1 est défaillant, un autre mécanisme exploitant un gène (ou des gènes) de type 2 prend la relève. L'avantage évolutif d'une telle redondance pour la stabilité des fonctions vitales est évident.

²² Voici une autre situation du même genre : soit M une mutation dans un gène. Seule M cause les symptômes A et B dans le phénotype de l'organisme. Dans ce cas, il est vrai de dire d'un organisme o qui n'est pas porteur de M : si o avait eu A , alors il aurait eu B ; ainsi que d'un organisme o^* porteur de M : si o^* n'avait pas eu A , il n'aurait pas eu B . Dans cette situation, il est pourtant intuitivement clair que A ne cause pas B et que B ne cause pas A .

propriétés macroscopiques, pour la simple raison qu'un processus, ou mécanisme, doit relier des événements spatialement et temporellement distincts entre lesquels il constitue un « pont ».

La réaction de Lewis à ces contre-exemples révèle sa conception révisionniste de la tâche philosophique de construire une conception satisfaisante de la causalité. Selon Lewis, la logique qui guide notre pratique ordinaire d'évaluation de contrefactuels est en contradiction avec nos jugements intuitifs dans les cas que nous avons mentionnés. « La solution appropriée » du conflit entre l'intuition et l'analyse contrefactuelle « consiste, il me semble, dit Lewis, à nier tout simplement les contrefactuels qui causent les ennuis »²³. Pour montrer que le jugement intuitif est en contradiction avec la logique profonde de notre propre stratégie (intuitive elle aussi) d'évaluation des contrefactuels²⁴, Lewis se concentre surtout sur un cas que nous n'avons pas encore mentionné : nous jugeons intuitivement vrais certains énoncés exprimant une dépendance contrefactuelle *contre la direction du temps*.²⁵ De tels contrefactuels à rebours (« *backtracking* ») sont intuitivement vrais dans chaque situation où, étant données les circonstances, un seul événement *C* a pu causer un événement *E*. Par exemple, il semble plausible de juger que s'il n'y avait pas de fossiles dans les roches d'une certaine montagne, il y aurait pas eu d'écoulement de boue de type *A* à une époque préhistorique que l'on peut déterminer par le taux de carbone-14 dans ces fossiles²⁶. La question de la vérité de tels contrefactuels à rebours est d'une importance stratégique pour l'approche de Lewis : si le contrefactuel sur les fossiles est vrai, alors la fossilisation dépend contrefactuellement de la présence des fossiles dans les roches aujourd'hui, alors qu'elle n'en dépend certainement pas causalement.

Comme le montre clairement l'affirmation citée plus haut, Lewis adopte ici une stratégie révisionniste : il juge qu'en ce qui concerne la vérité des contrefactuels à rebours, l'intuition se trompe. Sur quel genre de raisons peut-il étayer cette thèse si ce n'est sur d'autres jugements intuitifs ? En effet, son argumentation est la suivante : en général, nous évaluons la valeur de vérité d'un contrefactuel $A \square \rightarrow B$ en contemplant le monde le plus proche du nôtre dans lequel l'antécédent *A* du contrefactuel est vrai. Ensuite, nous constatons éventuellement que le conséquent *B* est vrai dans ce monde-là, auquel cas le contrefactuel est vrai. Si, en revanche, *B* y est faux, alors le contrefactuel est jugé faux. Or, dit Lewis, le monde possible qui contient l'événement *A* (que le monde actuel ne contient pas, par définition de ce qu'est un « contrefactuel »), mais qui diffère par ailleurs aussi peu que possible du monde actuel, est un monde où l'événement *A* arrive par « miracle ». Pour évaluer le contrefactuel, on conçoit un monde possible qui contient une violation miraculeuse mais très ponctuelle²⁷

²³ Lewis, « Causation », p. 170.

²⁴ Dans tous les cas, c'est l'intuition qui détermine la valeur de vérité des contrefactuels. La question est de savoir, au cas où différentes intuitions entrent en conflit, auxquelles il convient d'accorder la plus grande importance. Cf. Lewis, « Causation », p. 211 ; et Keil, ce numéro.

²⁵ Une raison de leur accorder une si grande importance est la possibilité que la vérité (selon Lewis seulement apparente) de *tous* les contrefactuels où la dépendance contrefactuelle ne correspond à aucune dépendance causale est fondée sur la vérité (apparente, selon Lewis) d'un contrefactuel à rebours.

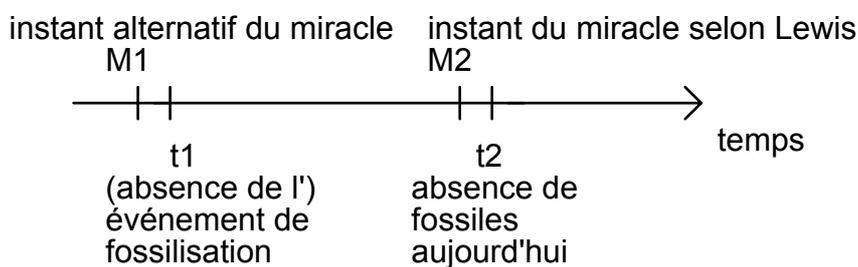
²⁶ Seules des conditions d'un genre très particulier permettent la fossilisation. Cf. Stephen Jay Gould, *Wonderful Life*, W.W. Norton, 1989. Trad. par M. Blanc, *La vie est belle : Les surprises de l'évolution*, Paris, Seuil (Points-Sciences), 1991, p. 82. Notons que la plausibilité de ce contrefactuel à rebours dépend de la plausibilité que l'on accorde à l'hypothèse créationniste : notre certitude qu'il est vrai diminue à mesure que l'on considère comme envisageable que les fossiles aient été créés par Dieu il y a 4000 ans, de sorte que la fossilisation d'animaux vivants il y a des millions d'années ne soit pas la seule cause possible de la présence des fossiles dans la roche, dans les circonstances qui prévalent sur Terre.

²⁷ Lewis parle d'un « tiny little miracle » (David Lewis, « Counterfactual Dependence and Time's Arrow », in : *Philosophical Papers, Vol. II*, p. 32-66, p. 44).

des lois du monde actuel. Cela permet de préserver notamment l'identité qualitative²⁸ de l'ensemble du passé (avant *A*) de ce monde avec le passé correspondant du monde actuel, ainsi que l'identité qualitative de tous les événements qui ne sont pas influencés par *A*. Par ailleurs, Lewis montre que l'introduction de ce petit miracle à un instant *avant A* requiert une plus petite déviation par rapport au monde actuel que son introduction *après A* : *A* ayant de nombreux effets qui se dispersent, il faut un plus grand miracle situé après *A* pour « effacer » toutes les traces que laisse *A* que pour faire surgir *A* suite à un petit miracle ayant lieu *avant A*, ce qui ne requiert qu'une intervention ponctuelle.

Lewis cherche alors à montrer que la nécessité d'évaluer les contrefactuels sur la base de la contemplation d'un monde possible où un petit miracle a lieu avant l'antécédent *A*, entraîne automatiquement la fausseté de tous les contrefactuels à rebours. Il y a donc ici un conflit entre l'intuition selon laquelle un contrefactuel à rebours comme celui sur les fossiles est vrai (ou du moins peut être vrai) avec ce que Lewis juge être un présupposé de toute évaluation intuitive de contrefactuels. En effet, Lewis soutient que l'introduction d'un « petit miracle » *avant* l'antécédent est une condition *générale* de possibilité de notre évaluation des contrefactuels. Pour restaurer la cohérence, il faut donc réviser notre système conceptuel en allant à l'encontre de certaines intuitions. Comme le respect de cette condition de possibilité est une contrainte plus fondamentale, car sous-jacente à l'ensemble de nos jugements sur la valeur de vérité des contrefactuels, que le respect de nos intuitions sur les conditionnels à rebours, il est rationnel de déclarer fausses, ou illusoire, nos intuitions sur la vérité de certains contrefactuels à rebours.

Le point faible du raisonnement que je viens d'exposer est la présence d'une prémisse tacite contestable. Lewis cherche à établir que les contrefactuels à rebours sont systématiquement faux. Appliquée à notre exemple, la raison qu'il donne est la suivante.



Le monde le plus proche du nôtre dans lequel l'antécédent est vrai au temps t_2 , dit Lewis, contient un miracle M_2 qui a lieu à l'instant immédiatement précédent t_2 . En revanche, le cours des choses qui *précèdent* M_2 est le même que dans le monde actuel. Donc, s'il y n'avait pas de fossiles aujourd'hui (t_2), ce n'est pas parce qu'il n'y aurait pas eu d'événement de fossilisation à t_1 : plutôt, il y aurait bien eu un événement de fossilisation à t_1 , et l'absence de fossiles à t_2 est due au miracle qui a lieu à un instant M_2 intermédiaire entre les temps t_1 (du conséquent) et t_2 (de l'antécédent).

²⁸ Ce n'est pas une identité au sens strict, car selon Lewis (David Lewis, *On the Plurality of Worlds*, Oxford, Basil Blackwell, 1986), aucune entité particulière ne peut exister dans différents mondes possibles à la fois. Supposer le contraire soulève le paradoxe des propriétés intrinsèques accidentelles : la propriété d'avoir cinq doigts à sa main droite est une propriété intrinsèque mais accidentelle de Lewis. Il aurait pu en avoir six. Autrement dit, Lewis a six doigts à sa main gauche dans d'autres mondes possibles. Or, si l'actuel Lewis était strictement identique avec les Lewis qui peuplent ces mondes possibles, il aurait à la fois cinq et six doigts à sa main gauche, ce qui est impossible. Donc, conclut Lewis, un Lewis existant dans un monde possible qui a six doigts à sa main gauche n'est que la « contrepartie » (counterpart) du Lewis du monde actuel. Cf. Max Kistler, « Necessary Laws », in : Jan Faye, Paul Needham, Uwe Scheffler, Max Urchs (éds.), *Nature's Principles*, Dordrecht, Springer, 2005, p. 201-227.

La thèse que Lewis omet de justifier est que notre raisonnement sur le monde possible où il n'y a pas de fossiles aujourd'hui contient un miracle à l'instant *immédiatement précédent* t_2 , et donc dans l'intervalle entre l'antécédent et le conséquent. Or cette supposition contredit, me semble-t-il, un autre présupposé (ou une autre condition de possibilité) du raisonnement sur les mondes possibles contrefactuels : pour juger si l'événement décrit dans le conséquent a lieu dans le monde contrefactuel où l'antécédent a lieu, il faut supposer que dans l'intervalle temporel entre les événements décrits par l'antécédent et le conséquent, les lois de la nature soient les mêmes que dans le monde actuel. Seule la connaissance de ces lois et le postulat de ne considérer comme proches que les mondes où ces lois sont respectées, nous permet de fonder notre jugement sur l'occurrence, dans ce monde, de l'événement décrit dans le conséquent. Après tout, dans un monde dont les lois diffèrent des nôtres, tout et son contraire peut arriver et nos jugements ne peuvent s'ancrer sur aucun fond de stabilité.

Or, dès que nous respectons cette condition de possibilité de nos jugements sur les mondes possibles, à savoir celle de maintenir la validité des nos lois actuelles durant l'intervalle entre l'antécédent et le conséquent, l'instant auquel l'introduction du miracle provoque le moins de divergence possible par rapport au monde actuel, n'est plus M_2 mais M_1 , c'est-à-dire l'instant immédiatement précédent le *conséquent* de notre contrefactuel à rebours. Or dès que le miracle se produit à M_1 , rien ne s'oppose plus à suivre l'intuition qui raisonne à rebours pour juger que si les fossiles n'avaient pas été là, cela aurait été à cause de l'absence de l'événement de fossilisation.

En conclusion, nous avons montré que l'argument lewisien selon lequel la seule manière de restaurer la cohérence entre nos jugements est d'aller à l'encontre de nos intuitions sur la vérité des contrefactuels à rebours, repose sur une prémisse discutable. Nous avons montré qu'il y a une bonne raison de nier la prémisse selon laquelle le miracle doit précéder *immédiatement* l'antécédent. A condition de tenir compte de l'ensemble des contraintes qui rendent nos jugements contrefactuels possibles, rien ne s'oppose finalement à l'acceptation des jugements intuitifs de dépendance contrefactuelle à rebours. Par conséquent, dans la mesure où l'événement de la fossilisation dépend contrefactuellement de la présence des fossiles aujourd'hui, sans en dépendre causalement, on ne peut pas réduire la causalité à la dépendance contrefactuelle.

Nous sommes entrés dans le détail de la conception contrefactuelle, pour deux raisons. D'une part, elle constitue (avec les conceptions probabiliste, de la manipulabilité, et de transfert) l'une des théories le plus souvent considérées comme les candidats les plus sérieux à l'analyse adéquate de la causalité²⁹. D'autre part, notre raisonnement permet de montrer que l'analyse contrefactuelle est plausible dans la mesure où elle rend compte de la composante explicative du concept intuitif de causalité. En effet, la dépendance nomique entre la propriété F et la propriété G (qui permet d'*expliquer* un événement qui a G en faisant référence à un événement qui a F ainsi qu'à la loi qui relie ces propriétés) est, dans des circonstances qui ne sont pas exceptionnelles, accompagnée d'une dépendance contrefactuelle. En revanche, dans les situations où deux effets d'une cause commune dépendent contrefactuellement l'un de l'autre sans que l'un soit cause de l'autre (comme dans le cas de la chute du baromètre et de la tempête) et dans les situations où la cause dépend contrefactuellement de l'effet, on peut expliquer notre intuition d'absence de relation causale, en faisant l'hypothèse que dans ces cas au moins, l'intuition exige la présence d'un élément mécanique de transmission, qui fait défaut dans ce type de situations. Il n'y a de processus qui se propage ni d'un effet à l'autre d'une cause commune ni de l'effet à la cause, contre la direction du temps.

La théorie du transfert et ses contre-exemples

²⁹ Cf. l'article de Keil, ce numéro.

L'espace nous manque ici pour examiner de près deux théories importantes et largement discutées aujourd'hui : la théorie probabiliste et la théorie de la manipulabilité. L'idée fondamentale de la première est que *A* est la cause de *B* si et seulement si l'occurrence de *B* est plus probable en présence d'*A* qu'en son absence³⁰. Selon la seconde, ce qui fait d'une chose la cause d'une autre est la possibilité pour un agent d'utiliser la première pour obtenir la seconde.³¹

Examinons plutôt, sans plus attendre, les perspectives d'un ensemble de théories qui articulent ce que j'ai appelé l'intuition mécaniste : selon ces théories, la cause est reliée à ses effets par un processus de transmission. Le transfert d'un caractère entre un émetteur et un récepteur est certainement une relation causale par excellence, et les théories qui s'inscrivent dans ce programme de recherche essaient de trouver une conception appropriée de ce qui doit être transmis pour que la transmission soit causale : pour certains, c'est l'information³², pour d'autres, c'est la capacité du processus de transmettre un caractère³³, ou encore la manifestation, tout au long d'un processus, d'un caractère permanent ou du moins, d'un caractère changeant de manière continue³⁴ ; enfin, ce qui fait qu'un processus est causal est peut-être, comme je l'ai proposé³⁵, le fait que ce qui est transmis entre la cause et l'effet est quelque chose de « matériel », sans être une substance : une quantité individuelle d'une grandeur conservée, comme par exemple une certaine quantité d'énergie.

Les théories du transfert sont conçues sur mesure pour rendre compte des situations qui constituent des contre-exemples aux théories nomologiques, contrefactuelles et probabilistes : elles élaborent l'intuition selon laquelle la causalité est fondée sur la présence d'un mécanisme de propagation qui s'étend progressivement de la cause à l'effet. Si le mécanisme *A* cause un effet *E* en empêchant le mécanisme *B* (de « sécurité ») d'aboutir, au sens où, en présence de *A*, *B* ne va pas jusqu'à son terme où il produirait *E*, l'existence d'une

³⁰ On trouvera un examen critique de cette conception dans l'article de Schaffer, ce volume ; cf. aussi Kistler, « Causation in contemporary analytical philosophy ». On se contentera ici de noter qu'elle est confrontée à des contre-exemples assez simples et analogues aux contre-exemples à l'analyse contrefactuelle. Il y a d'une part des situations où un type d'événement augmente la probabilité d'un événement d'un second type sans pourtant être sa cause ; et d'autre part des situations où la cause diminue la probabilité de son effet. Sans pouvoir le montrer en détail ici, il me semble impossible d'éviter ces contre-exemples en introduisant des modifications « cosmétiques » dans la théorie probabiliste, tout en conservant sa thèse principale. Ces contre-exemples semblent au contraire révélateurs du fait qu'il s'agit, là encore, d'une analyse qui rend compte de l'aspect explicatif du concept de causalité mais néglige l'intuition selon l'existence d'une relation causale exige un mécanisme matériel de transmission.

³¹ Des variantes de cette théorie ont été défendues par Gasking, « Causation and Recipe » ; von Wright, *Explanation and Understanding*, Cartwright, *How the Laws...*, Price, « Agency and Causal Asymmetry » ; Menzies et Price, « Causation as a Secondary Quality » ; Geert Keil, *Handeln und Verursachen*, Vittorio Klostermann, Frankfurt a. M., 2000, et Keil, ce numéro. Le défaut principal de cette conception est qu'elle est insuffisamment fondamentale. La conception de transfert permet d'expliquer pourquoi elle permet de rendre compte des conséquences des mouvements corporels qui constituent les actions – il s'agit de relations causales par transfert – mais aussi d'expliquer pourquoi il n'y a aucune différence à l'égard de leur caractère causal entre des événements provoqués par une action et des événements causés par des événements radicalement inaccessibles à l'action : des événements à l'échelle microscopique ou au contraire cosmologique, ou des événements ayant lieu en dehors du « cône de lumière » dont l'origine contient la durée de l'existence des humains. Ces événements sont pour des raisons fondamentales (éclaircies par la théorie de la relativité) inaccessibles à toute influence causale due à une action humaine ; mais cela ne les empêche intuitivement pas d'avoir des causes et des effets.

³² Wladislaw Krajewski, « Energetic, informational, and triggering causes », *Erkenntnis* 46 (1997), 193-202 ; Don Ross et David Spurrett, « What to Say to a Sceptical Metaphysician: A Defense Manual for Cognitive and Behavioral Scientists », *Brain and Behavioral Sciences*, à paraître.

³³ Reichenbach, *The Direction of Time* p. 198 sq. ; Salmon, *Scientific Explanation*, p. 139 sq.

³⁴ Russell, *Human Knowledge...* ; Dowe, « Wesley Salmon's Process Theory... » ; Dowe, *Physical Causation*.

³⁵ Kistler, « Reducing Causality... » ; Kistler, *Causalité et lois de la nature*.

relation causale entre A et E est due à l'existence d'un processus de transmission qui commence en A et se termine en E , alors que l'absence d'une relation causale entre B et E est due au fait que, en la présence de A , le processus qui va de B à E n'aboutit pas. Lorsque, en revanche, les deux processus vont jusqu'à leur terme, nous jugeons qu'il s'agit d'une situation de surdétermination où les deux processus causent E , ce qui se justifie par l'existence de deux processus de transfert.

Les théories de transfert évitent aussi les autres contre-exemples mentionnés plus haut : aucun processus ne se propage entre la chute du baromètre et la tempête ; aucun ne se propage entre la présence des fossiles, en amont contre la direction du temps, vers l'événement de fossilisation ; et rien ne se propage entre des états de choses qui concernent le même endroit au même temps, comme le fait que l'échantillon de gaz idéal (contenu dans un volume V) possède une température T et le fait qu'il possède une pression p .

Je ne peux pas entrer ici dans le détail d'une évaluation comparative des différentes théories « processuelles ». Quelques remarques devront suffire : la proposition de Reichenbach et Salmon selon laquelle le lien causal est fondé sur l'existence d'un « processus causal » qui se caractérise par sa capacité de transmettre un caractère, souffre de sa circularité, c'est-à-dire de son incapacité de définir le concept de processus, sans utiliser le concept de causalité, que l'on cherche à caractériser au moyen de ce même concept de processus. Par ailleurs, l'analyse de Salmon est insuffisamment fondamentale dans la mesure où elle n'offre aucune explication de la raison pour laquelle les processus ont la capacité de transmettre un caractère alors que cette capacité fait défaut aux « pseudo-processus ». L'analyse de Russell et Dowe (à laquelle Salmon lui-même s'est rallié en 1994) fonde l'existence d'une relation causale sur un processus de propagation *qualitative* qui ne requiert cependant pas de transmission matérielle. Il suffit qu'il y ait un caractère qui se manifeste de manière continue sur une ligne de monde qui relie la cause à l'effet. Cette condition est trop faible : comme j'ai essayé de le montrer ailleurs³⁶, il y a des pseudo-processus (qui ne sont donc pas causaux) qui respectent cette condition, par exemple la propagation d'une tache lumineuse sur la paroi interne d'un cylindre au centre duquel tourne un phare. Selon l'analyse de Glennan³⁷, l'existence d'une relation causale entre C et E est fondée sur l'existence d'un mécanisme qui s'étend entre C et E et qui consiste en un assemblage de parties en interaction selon des lois de la nature. La faiblesse de l'analyse mécaniste de la causalité de Glennan réside dans la limitation de sa validité aux relations causales *macroscopiques*. En effet, la conception d'un mécanisme adoptée par Glennan est inapplicable aux processus fondamentaux : la « collision » d'un photon avec un électron a pour effet un changement de la longueur d'onde du photon (c'est l'effet Compton), mais il n'est pas possible de rendre compte de cette interaction causale en la décomposant en parties articulées comme les rouages d'une montre.

Sans pouvoir justifier cette thèse dans le détail ici, l'expression suivante de l'intuition mécaniste me semble fondamentale, non-circulaire, et généralement applicable : deux événements sont reliés comme cause et effet si et seulement si il existe une quantité d'énergie, de charge électrique ou de toute autre grandeur conservée en vertu d'une loi fondamentale de la nature, qui est transférée entre ces événements. Pour éviter toute circularité, il faut entendre par « transfert » tout simplement la présence d'une quantité donnée d'une grandeur conservée dans l'un et l'autre des deux événements. Cette conception de la causalité comme fondée sur un transfert matériel permet non seulement d'échapper aux contre-exemples examinés auparavant, mais aussi d'expliquer pourquoi les processus peuvent transmettre des caractères : les caractères ne sont rien d'autre que des formes d'apparence de certaines quantités

³⁶ Kistler, « Reducing Causality ... » ; Kistler, *Causalité et lois de la nature*.

³⁷ Glennan, « Mechanisms... ». Cette analyse est développée dans Peter Machamer, Lindley Darden et Carl F. Craver, « Thinking about Mechanisms », *Philosophy of Science* 67 (2000), p. 1-25 et Carl F. Craver, « Role Functions, Mechanisms and Hierarchy », *Philosophy of Science* 68 (2001), p. 53-74.

individuelles d'énergie ; c'est à ce titre qu'ils sont transférés entre cause et effet. Cette conception évite le problème de la théorie de la « manifestation continue » (mentionnée ci-dessus) de Russell et Dowe puisqu'aucune quantité de grandeur conservée n'est transférée le long d'un pseudo-processus, même si celui-ci constitue une ligne de monde qui manifeste une qualité de manière continue. Finalement, elle rend justice à l'intuition mécaniste tout en s'appliquant aux interactions fondamentales : dans l'effet Compton, la cause transfère à l'effet la charge électrique, l'énergie et la quantité de mouvement.

Bien entendu, la théorie du transfert est confrontée à ses propres difficultés. Je me concentrerai ici sur une seule³⁸ : la théorie du transfert ne permet pas de rendre compte de l'intuition selon laquelle des *absences*, *omissions*, ou autres *faits de forme négative* peuvent être causes ou effets. Quelque chose qui n'existe pas, ou qui est absent, ne peut être transféré sur rien, et ne peut donc jouer le rôle de cause. Or, nos pratiques explicatives, dans la vie de tous les jours, dans le discours juridique, ainsi que dans les sciences, regorgent d'appels à des causes ou effets « négatifs », c'est-à-dire exprimées par des descriptions négatives. De nombreux raisonnements juridiques considèrent les omissions comme des causes. Par exemple, « l'omission de livrer à un fabricant une pièce de machine, peut être considérée comme la cause de la perte des profits qui auraient été obtenus par son usage »³⁹.

Il est tout aussi facile de trouver des explications biologiques qui mentionnent des facteurs négatifs comme des causes. La « théorie de l'opéron », développée par Jacob et Monod en 1961, propose un mécanisme qui permet de comprendre pourquoi certains gènes ne sont exprimés en protéines que dans certaines situations⁴⁰. Voici une esquisse très simplifiée de ce mécanisme. Les gènes « structuraux » sont des morceaux d'ADN dont la séquence « encode » des protéines spécifiques. Dans la séquence de l'ADN, il y a, associé à un gène structural donné, au moins trois gènes régulateurs, dans l'ordre : le régulateur, le promoteur et l'opérateur. Lorsqu'une molécule appelée « ARN-polymérase » se fixe sur le promoteur, cela déclenche la transcription des gènes suivants dans la séquence de l'ADN : l'opérateur, puis les gènes structuraux.

Le déroulement de ce processus peut être empêché par une protéine appelée « répresseur » : le gène régulateur code pour le répresseur. Dès que ce répresseur est produit, il se fixe sur l'opérateur, ce qui empêche l'ARN-polymérase de se fixer sur le site promoteur et donc de commencer la transcription. Par conséquent, à l'état de repos, le système ne produit pas de protéine P. Le « signal » qui déclenche sa synthèse prend la forme d'une molécule appelée « inducteur ». Cette molécule s'attache à un site spécifique sur le répresseur, ce qui fait perdre à celui-ci la capacité de se fixer sur l'opérateur. De cette manière, la voie est libre pour l'ARN-polymérase. Elle commence à transcrire les gènes en aval de son site d'attachement, c'est-à-dire le promoteur. Cette transcription mène à la production de la protéine P.

Selon cette description schématique du mécanisme de production de ces enzymes, elle semble être l'effet d'une cause négative, et même doublement négative : l'inducteur déclenche la transcription *en empêchant le répresseur de l'empêcher*. Jonathan Schaffer

³⁸ J'ai montré ailleurs (Kistler, « Causation in contemporary... ») qu'elle échappe aux objections qui ont été soulevées contre des variantes antérieures, notamment celle de David Fair, « Causation and the Flow of Energy », *Erkenntnis* 14 (1979), p. 219-250.

³⁹ H.L.A. Hart et Tony Honoré, *Causation in the Law*, Oxford, Clarendon Press, 1959, p. 56.

⁴⁰ Cf. Pierre Kamoun, Alain Lavoine, Hubert de Verneuil, *Biochimie et biologie moléculaire*, Flammarion, Médecine-Sciences, Paris, 2003, p. 172 ; la structure de l'explication de la régulation de l'expression des protéines par la théorie de l'opéron est analysée par Kenneth F. Schaffner, *Discovery and Explanation in Biology and Medicine*, Chicago, Chicago University Press, 1993, chap. 3, et Michel Morange, *La part des gènes*, Paris, O. Jacob, 1998. Cf. aussi Michel Morange, *Histoire de la biologie moléculaire* (1994), La Découverte, Paris, 2003, chap. 14, et Claude Debru, *Philosophie moléculaire : Monod, Wyman, Changeux*, Paris, Vrin, 1987.

appelle les relations causales de ce type des « déconnexions »⁴¹ : la cause agit en déconnectant un facteur potentiellement actif d'un facteur qui bloque son action. Schaffer trouve un peu partout des instances de cette forme de causalité « doublement négative » où la cause agit en enlevant un facteur empêchant l'effet. On en trouve deux dans ce cas paradigmatique de processus causal qu'est un assassinat accompli en tirant une balle dans le cœur de la victime. Premièrement, l'acte de tirer la gâchette cause l'éjection de la balle en enlevant un facteur inhibiteur : le cran qui empêche la décompression d'un ressort. Deuxièmement, lorsque la balle traverse le cœur, cela cause l'absence de sang oxygéné qui irrigue le cerveau, absence qui cause une absence d'oxygénation des cellules cérébrales, ce qui cause la mort. Il est clair, dit Schaffer, que les théories qui conçoivent la causalité en termes de « relations intrinsèques [...], telles que des connexions physiques, n'ont pas de place pour la causalité par déconnexion, car la causalité par déconnexion dépend de la circonstance externe où une connexion préexistante est rompue »⁴². Dans le mécanisme qui déclenche l'expression des gènes structuraux selon le modèle de l'opéron esquissé plus haut, l'inducteur est une cause qui agit par « déconnexion » au sens de Schaffer : l'inducteur empêche le répresseur d'empêcher la transcription des gènes structuraux.

L'observation de l'omniprésence de la référence aux facteurs négatifs, et même doublement négatifs, dans les explications juridiques et scientifiques, semble nous donner une raison simple mais massive d'écarter la théorie du transfert des propositions en lice pour rendre compte de la nature de la relation causale.

Nous nous trouvons donc apparemment devant un dilemme. D'un côté, la théorie du transfert est la seule qui permet de surmonter les problèmes que posent les problèmes de surdétermination, de préemption et de l'existence de dépendances nomiques, contrefactuelles et probabilistes sans lien causal à toutes les conceptions rivales. D'un autre côté, le cadre conceptuel du transfert semble incapable de permettre d'interpréter une classe importante de jugements causaux faisant appel à des causes négatives ou à des effets négatifs. Il n'y a, me semble-t-il, qu'une manière d'échapper à cette ornière qui semble nous condamner au scepticisme ou au relativisme : il faut distinguer entre deux aspects différents et complémentaires du concept de causalité auxquels on fait appel pour répondre à des interrogations différentes.

Les deux volets du concept de causalité

On peut regrouper les exemples de relations causales impliquant des faits négatifs en trois catégories : 1) un fait négatif cause un fait positif : l'absence du répresseur cause la transcription de la protéine. Les omissions appartiennent à cette catégorie : un père distrait n'intervient pas pour empêcher son enfant de traverser la rue, où une voiture le heurte. L'omission du père de retenir l'enfant cause l'accident. 2) Un fait positif cause un fait négatif : la présence du répresseur cause l'absence d'expression de la protéine. Les préventions appartiennent à cette catégorie : le geste du père de retenir l'enfant prévient l'accident (c'est-à-dire « cause » le fait qu'il n'y ait pas d'accident). 3) un fait négatif cause un autre fait négatif : l'absence de l'inducteur cause l'absence d'expression de la protéine. En l'absence d'un certain gène (appelé « fosB »), les souris négligent leur progéniture : l'absence du gène cause l'absence de soins maternels, ce qui conduit à leur mort. Nous avons un cas de

⁴¹ Jonathan Schaffer, « Causation by Disconnection », *Philosophy of Science* 67 (2000), p. 285-300; Jonathan Schaffer, « The Metaphysics of Causation », *Stanford Encyclopedia of Philosophy*, 2003 (<http://plato.stanford.edu/entries/causation-metaphysics/>), p. 14, et Schaffer, ce volume.

⁴² Schaffer, « Causation by Disconnection », p. 292.

prévention due à un fait négatif lorsque nous jugeons que le fait que la mère n'ait pas distraité le père empêche l'accident de l'enfant.⁴³

Nos jugements causaux s'appuient sur deux genres d'intuitions, correspondant respectivement à la conception explicative et mécaniste. Dans certaines situations, l'intuition semble suggérer que des faits négatifs sont liés comme cause et effet, alors qu'il n'y a semblait-il aucun sens à parler de transmission entre faits négatifs. Ces intuitions se nourrissent de la possibilité d'*expliquer* l'un des faits par l'autre. La qualité d'une telle explication sera jugée en partie en se fondant sur considérations pragmatiques⁴⁴, et en partie en se fondant sur des lois de la nature. Or, il y a des lois portant sur des conditions négatives. Par exemple, toutes les cellules cérébrales humaines meurent en l'absence d'oxygène. Analyser la causalité sur le plan des explications causales justifiées à leur tour par les lois, revient à adopter une « conception épistémique »⁴⁵ de la causalité. Notre conception de la causalité doit tenir compte de ce lien intuitif entre causalité et explication. Mais dans un certain nombre de contextes, on juge qu'une relation pourtant explicative n'est pas causale parce qu'il n'y a aucun processus mécaniste de transmission sous-jacent. Dans ces cas, c'est l'aspect mécaniste du concept de causalité qui informe notre jugement. Le fait de reconnaître la légitimité des analyses qui se situent au niveau épistémique ne s'oppose pas à la légitimité d'une « conception ontique » de la causalité : celle-ci vise la découverte du type de processus, ou mécanisme, fondamental sur lequel repose la vérité des explications causales. Notre tâche consiste à montrer que ces deux types d'intuitions sur la causalité n'existent pas l'une à côté de l'autre, quitte à rendre parfois des verdicts contradictoires. Il s'agit de construire une analyse qui permet de rendre compte des deux tout en les intégrant dans un schéma cohérent.

Au niveau épistémique, l'importance des « causes négatives » est incontestable. Il n'est donc pas surprenant que les conceptions de la causalité rivales de la conception de transfert, qui se situent sur le plan de la relation épistémique entre une condition qui permet d'expliquer une autre condition, s'accrochent de l'existence de relations causales entre faits négatifs.

Je propose d'esquisser en premier lieu la manière dont les aspects mécaniste et explicatif s'articulent dans le cas général, pour revenir ensuite aux causes et effets négatifs. Lorsqu'on donne une explication causale correcte de l'état de choses (ou du fait) que *b* est *G*, en faisant appel à l'état de choses (ou au fait) que *a* est *F*, il existe une relation entre le fait que *a* est *F* et le fait que *b* est *G* qui rend cette explication vraie. On appelle cette relation parfois le « vérifacteur » de l'explication. Le fait que *b* est *G* porte sur l'événement *b* (cela peut être un objet à un instant), et le fait que *a* est *F* porte sur l'événement *a*.

On explique par exemple causalement le fait que la protéine (*b*) est exprimée (*G*) par le fait que l'inducteur (*a*) s'est fixée sur le répresseur (*F*) : *b* est *G* parce que *a* est *F*. Cette explication est vraie parce qu'il existe d'une part une relation de dépendance nomique entre la propriété *F* d'être une fixation de l'inducteur sur le répresseur et la propriété *G* d'être une expression de protéine⁴⁶, et d'autre part parce que ces propriétés appartiennent à des

⁴³ Les exemples d'omissions et de préventions mettant en scène l'accident sont proposés par Phil Dowe, « A Counterfactual Theory of Prevention and 'Causation' by Omission », *Australasian Journal of Philosophy* 79 (2001), p. 216-226.

⁴⁴ Cf. Sylvain Bromberger, *On What We Know We Don't Know*, Chicago, University of Chicago Press, et Stanford, CSLI, 1992.

⁴⁵ Je transpose la distinction que fait Wesley Salmon (*Scientific Explanation*, p. 276) entre deux conceptions de l'explication, aux conceptions de la causalité.

⁴⁶ Cette dépendance n'existe que dans des circonstances bien particulières d'une cellule qui fonctionne normalement. Si l'on veut l'appeler une « loi de la nature », il est important de préciser qu'il s'agit d'une loi dont la validité se limite à ce type de système bien particulier. Dans les termes de Schurz (G. Schurz, « *Ceteris paribus* Laws : Classification and Destruction », *Erkenntnis* 57 (2002) p. 351-372), c'est une « loi de système », alors que, dans les termes de Cartwright (N. Cartwright, *The Dappled World*, Cambridge, Cambridge University

événements (*a* et *b*, respectivement) entre lesquels un processus de transmission d'énergie a lieu : la fixation de l'inducteur sur le répresseur (*a*) cause la production de la protéine (*b*) au sens d'une transmission mécaniste.

Il ne s'agit pas d'opposer simplement la *causalité réelle* qui prend la forme d'un transfert, à l'*explication* entre propositions qui prend la forme d'un argument, comme le fait Davidson lorsqu'il oppose la *relation causale* entre événements à la relation d'*explication causale* entre des *descriptions* de ces événements⁴⁷. Il est nécessaire de postuler des faits en tant qu'entités plus structurées que les événements, pour pouvoir rendre compte de la différence entre des explications causales vraies et fausses. Admettons que l'explication : « *b* est *G* parce que *a* est *F* » soit vraie, mais que l'explication « *b* est *G* parce que *a* est *H* » soit fausse. Les deux portent sur des événements *a* et *b* reliés comme cause et effet (puisque nous supposons que la première explication est vraie). Je propose d'expliquer l'origine de la différence de leurs valeurs de vérité en supposant que ce qui rend vraie une explication causale, son vérifacteur, est une relation complexe entre *faits*. Au niveau des événements, *a* est cause de *b*. En revanche, la première explication est vraie alors que la seconde est fausse parce que la première fait référence à une relation entre des faits qui font intervenir les propriétés correctes. Ce qui rend vraie l'explication du fait *Gb* par le fait *Fa*, c'est l'existence d'une relation de *responsabilité causale* entre ces deux faits. La responsabilité causale entre les faits *Fa* et *Gb* implique deux choses : (1) la relation de transfert entre les événements *a* et *b* et (2) la dépendance nomique entre les propriétés *F* et *G*. Il ne peut y avoir de responsabilité causale entre le fait que *a* est *F* et le fait que *b* est *G* qu'à condition qu'il y ait transfert entre les événements, *a* et *b*, sur lesquels portent ces faits et qu'à condition que les instances de *F* entraînent nomologiquement des instances de *G*.

Nous sommes donc amenés à postuler trois niveaux d'analyse : il y a d'abord le niveau des *propositions* figurant dans une *explication*. S'il s'agit d'une explication causale, elle est rendue vraie par *des faits* qui se trouvent dans une relation de *responsabilité causale*. La relation de responsabilité causale peut être analysée en deux composantes : une relation de transfert entre les événements *a* et *b* sur lesquels portent les faits *Fa* et *Gb*, et une relation de dépendance nomique entre les propriétés *F* et *G* que les faits *Fa* et *Gb* attribuent à *a* et *b*.

Le postulat de ces trois niveaux, au lieu des deux niveaux retenus par Davidson, se révèle être fructueux dans l'analyse des relations causales mettant en jeu des causes et effets négatifs. Lorsqu'on explique la transcription de la protéine par l'*absence* du répresseur, où est l'événement qui transmet l'énergie sur l'effet, c'est-à-dire la transcription de la protéine ? Lorsque l'omission du père de retenir l'enfant cause l'accident, que transmet le père sur l'effet, c'est-à-dire l'accident ?

Il y a semble-t-il deux réponses possibles : dans la mesure où la relation de transmission est le noyau de la relation causale, l'on peut faire valoir qu'il ne s'agit pas de cas de causalité véritable, mais d'une forme d'ersatz-causalité. Phil Dowe argumente en ce sens que les omissions sont des cas de « quasi-causalité »⁴⁸, où l'on fait référence à l'effet *contrefactuel* qu'aurait eu une cause *contrefactuelle*. De manière analogue, nous pourrions défendre la théorie du transfert en faisant valoir que le père n'a rien fait et il n'a pas non plus causé l'accident, au sens strict de « causer » qui requiert une transmission. Lorsqu'on dit néanmoins couramment qu'il l'a causé, on le dit en un sens dérivé, par analogie : un tel

Press, 1999), la cellule est une « machine nomologique » : elle fonde la dépendance nomique en question qui n'existerait pas sans elle.

⁴⁷ D. Davidson, « Causal Relations » (1967), dans D. Davidson, *Essays on Actions and Events*, Oxford, Oxford University Press, 1980. Trad. P. Engel, *Actions et événements*, Paris, Gallimard, 1993.

⁴⁸ Phil Dowe, « A Counterfactual Theory of Prevention and 'Causation' by Omission », *Australasian Journal of Philosophy* 79 (2001), p. 216-226, p. 217; *Id.*, *Physical Causation*, Cambridge, Cambridge University Press, 2000, chap. 6.

énoncé fait référence non pas à une relation causale réelle mais à une *relation causale contrefactuelle*. En disant que le père est responsable de l'accident, et l'a donc « causé » en un sens non littéral, on dit que *si* le père *avait* retenu l'enfant, l'accident ne se *serait* pas produit. Or, dans la mesure où nous n'avons pas trouvé la théorie contrefactuelle acceptable, cela n'est pas équivalent à un jugement causal (entre la non-retention par le père et l'accident). Cette solution préserve l'intuition mécaniste mais au prix d'aller à l'encontre de l'intuition du sens commun selon laquelle certaines omissions causent leurs effets.

Dans la mesure où le respect des intuitions constitue une contrainte essentielle aux yeux de la métaphysique descriptive, on sera tenté de suivre la voie empruntée par Thomson⁴⁹, pour reconnaître que les omissions sont réellement des causes. Cet auteur propose de tenir compte du fait qu'une cause négative, par exemple une omission, n'est néanmoins pas une cause au sens plein du terme, en disant qu'elle n'est qu'une cause *indirecte*, en ce sens : lorsqu'on dit qu'une omission cause *x*, cela signifie qu'*il y a* un événement *E* qui cause *x*. Cette affirmation s'accorde en effet avec une implication de la conception esquissée plus haut. L'omission est d'abord un fait rapporté dans une explication. Dans la mesure où l'explication par une omission est causale, il doit exister un événement cause qui est source de transfert. Quel est cependant le lien entre l'omission et cet événement ? C'est une question laissée entièrement ouverte par Thomson.

Les situations où l'on désigne une omission ou un autre fait négatif *Fa* comme la cause d'un fait *Gb* tombent dans deux catégories : dans la première, le fait *Fa* constitutif de l'omission porte explicitement sur un événement (ou un objet à un instant) qui a causé *Gb* au sens de la transmission physique, mais sans toutefois mentionner la propriété⁵⁰ qui est nomiquement responsable de la propriété pertinente de *Gb*. L'existence des propriétés *F* et *G* liées par une dépendance nomique et instanciées par *a* et *b* respectivement (*a* est *F* et *b* est *G*) est le second ingrédient de la relation de responsabilité causale, vérificateur de toute explication causale vraie. L'événement-cause *a* possède donc toujours une propriété *F* qui fait en sorte que (ou : qui est responsable du fait que) l'effet *b* ait la propriété *G*. Seulement, dans le cas des énoncés mentionnant un fait négatif comme cause, cette propriété n'est pas explicitement nommée. Souvent, le fait négatif permet de faire une inférence plausible, sur l'arrière-plan de la connaissance de la situation et de connaissances générales, à la propriété *F* qui figure dans le fait *A* causalement responsable.

Lorsqu'on dit par exemple que le fait que le répresseur ne soit pas fixé sur l'opérateur provoque la transcription, on attribue la responsabilité causale à un fait négatif. Cependant, la chaîne causale qui mène à la transcription passe bien par un événement qui implique l'opérateur. Le fait mentionné comme cause porte donc sur un événement cause au sens de la théorie de transfert. Seulement, le jugement ne désigne pas directement la propriété de l'opérateur qui est responsable de la transcription : il s'agit d'une certaine conformation que la molécule adopte lorsqu'elle n'est pas liée au répresseur. La connaissance du mécanisme permet d'inférer cette propriété à partir de la description négative donnée.

Dans le deuxième type d'énoncés où une omission ou un autre fait négatif joue le rôle de la cause, l'expression désignant l'omission ne donne aucune indication de l'événement qui cause *Gb* au sens de la transmission. Lorsqu'on dit que le fait que le père n'a pas retenu l'enfant a causé l'accident, ou que « le fait que Jean n'avait pas fermé le portail a provoqué un désastre »⁵¹, ou que l'absence du gène *fosB* chez la mère cause la mort des souriceaux, le contenu littéral des expressions désignant la cause ne contient aucune information sur la

⁴⁹ Judith Jarvis Thomson, « Causation : Omissions », *Philosophy and Phenomenological Research* 66 (2003), p. 81-103.

⁵⁰ On pourrait appeler cette propriété « positive » ou « réelle » si ce n'était pas redondant : toutes les propriétés qui existent sont réelles et positives.

⁵¹ Jonathan Bennett, *Events and Their Names*, 1988, p. 140.

nature de l'événement qui a causé l'événement en question (au sens de la transmission) : le choc entre la voiture et l'enfant au milieu de la route, la collision entre un train et un camion qui a traversé le portail ouvert, le processus de déshydratation des souriceaux. Le sens de ces énoncés a néanmoins un aspect causal, pour deux raisons :

1) l'expression désignant l'omission, absence ou autre « cause négative » nous permet, grâce à notre connaissance du contexte, de reconstruire la nature de l'événement cause.⁵²

2) L'intérêt d'utiliser de telles expressions négatives est qu'elles peuvent véhiculer une information contrefactuelle : à chaque fois, elles placent le fait négatif en contraste avec un fait ou une situation « normale » dans laquelle l'effet ne se serait pas produit⁵³. La « normalité » en question peut relever du devoir moral comme dans les cas du père et du gardien du portail, ou de la normalité biologique au sens de l'adéquation du fonctionnement d'un organisme (ou d'une partie d'un organisme) à sa fonction propre⁵⁴, comme dans le cas des souris. Les énoncés négatifs ne se contentent pas de désigner un fait qui ne porte pas sur la cause au sens de la transmission. Ils désignent un fait négatif dont le pendant positif aurait causé l'effet « normal ». Le fait que le père retienne l'enfant fait que l'enfant n'a pas d'accident. Le fait que le gardien ferme le portail fait qu'il n'y a pas d'accident. La présence du gène fosB fait que la souris nourrit ses petits.

Dans les explications causales contenant des causes ou effets négatifs de ce second type, l'expression négative occupant le rôle du désignateur de la cause dans l'explication ne fait pas référence au fait causalement responsable de l'effet *Gb*, ni même à l'événement sur lequel il porte (comme c'est le cas dans les énoncés du premier type). Cependant, même dans ce cas, le contexte et les connaissances d'arrière-plan permettent de faire des inférences plausibles à la fois sur l'événement *a* et sur la propriété efficace *F*.

On peut faire une distinction analogue à l'égard des expressions négatives désignant les effets : dans un type de cas, ces expressions désignent bien l'événement-effet mais non la propriété « positive » causalement affectée ; dans l'autre type de cas, l'expression désignant l'effet ne contient explicitement ni l'information sur l'événement-effet ni sur la propriété positive concernée : c'est alors, là aussi, le contexte et les connaissances d'arrière-plan qui permettent de faire des inférences probables sur l'identité du fait *B*.

Considérons le cas où l'absence du gène fosB de la souris cause la mort de ses souriceaux. Ici, l'expression désignant la cause appartient au second type, celle désignant l'effet au premier. Il est assez naturel de dire que l'absence du gène fosB, obtenue par la technique dite du « knock-out », dans une souris mère, cause indirectement la mort de ses souriceaux, en causant directement la négligence par la mère de sa progéniture⁵⁵. En même temps, on comprend mal en quel sens il pourrait y avoir un lien causal quelconque entre la

⁵² Thomson a tort de penser que la connaissance de cet événement peut toujours être extrait de l'énoncé négatif lui-même. Cf. Kistler, « Causes as Events and Facts ».

⁵³ Sur le contraste, cf. P. Achinstein, « Causation, Transparency, and Emphasis », *Canadian Journal of Philosophy* 5 (1975), p. 1-23; F. Dretske, « Referring to Events », *Midwest Studies in Philosophy* 2 (1977), p. 90-99; C.R. Hitchcock, « The Role of Contrast in Causal and Explanatory Claims », *Synthese* 107 (1996), p. 395-419.

⁵⁴ Pour simplifier, la fonction propre d'un trait est celui parmi les effets de sa présence dans un animal qui a été la cause de sa sélection naturelle et de sa persistance dans le génome de l'espèce. Selon R. Millikan, « pour qu'une entité A ait la fonction F comme 'fonction propre', il est nécessaire (et quasiment suffisant) que [...] A a son origine dans la 'reproduction' [...] d'une entité ou d'entités antérieures qui, en partie *en vertu* de leur possession des propriétés reproduites, ont réellement effectué F dans le passé, et A existe à cause de (ou historiquement à cause de) cette ou ces performances » (Ruth Garrett Millikan, « In Defense of Proper Functions », in R. Millikan, *White Queen Psychology and Other Essays for Alice*, Cambridge, Mass., MIT Press, 1993, p. 13-29).

⁵⁵ La technique du « knock-out » permet de remplacer un morceau d'ADN correspondant à un gène, par un fragment inactif d'ADN, pour ensuite implanter l'ADN modifiée dans des embryons. Dans ces organismes, le gène ainsi éliminé ne s'exprime pas. Cf. Morange (1998), p. 82, 165.

mère et les petits après l'accouchement, dans la mesure où la mère n'interagit plus du tout avec les petits. Il s'agit d'un cas où les intuitions correspondant aux conceptions « explicative » et « mécaniste » entrent en conflit. L'articulation esquissée plus haut de l'explication avec son vérificateur permet de résoudre ce conflit : dans le cas d'une explication causale où l'*explanans* et l'*explanandum* ont une forme ou un contenu négatif, ni l'*explanans* ni l'*explanandum* ne contient de référence explicite aux faits qui constituent les termes de la relation de responsabilité causale qui rend l'explication vraie. Nos connaissances d'arrière-plan suggèrent parfois une description explicite, ici par exemple : la déshydratation des souriceaux qui provoque un arrêt cardiaque.

En revanche, la mère des souriceaux n'est pas causalement impliquée dans leur mort, et l'absence du gène *fosB* ne fait à plus forte raison pas partie du fait causalement responsable. En disant cela, l'on tire une conséquence de l'intuition mécaniste selon laquelle il n'y a pas relation causale entre la mère et ses petits dans la mesure où elle n'interagit plus physiquement avec eux. La théorie de la transmission voit dans l'absence d'interaction l'absence de transmission. Rien n'est transmis, et il n'y a donc aucune influence causale.

Lorsqu'on fait référence à l'absence du gène chez la mère comme facteur responsable de la mort des souriceaux, on produit donc, au sens strict, une explication non-causale. Cependant, cette explication contient des informations sur des processus causaux, ce qui peut justifier de la considérer comme causale en un sens plus large : il peut être pertinent de faire référence à ce facteur négatif dans la mesure où cela permet de véhiculer une information contrefactuelle : faire référence à l'*absence* du gène *fosB* comme fait responsable de la mort, implique qu'il s'agit d'une situation anormale que l'on contraste avec la situation normale où le gène est présent et où sa présence est causalement responsable de la survie des souriceaux.

Ce qui rend vrai le jugement selon lequel l'absence de soins « est due à » l'absence du gène (quoique non au sens strictement causal), c'est le fait que le gène ainsi que les soins soient absents mais que, dans la situation contrefactuelle qui correspond à la situation normale pour les souris, le gène (ou la présence du gène) *aurait* causé les soins maternels. Ce qui rend vrai le jugement, ce n'est donc pas l'existence, dans la situation réelle, d'une relation causale entre absences, mais plutôt l'absence d'une relation causale à laquelle on fait implicitement référence de manière contrefactuelle.

Conclusion

La théorie de la causalité en termes de transfert suscite souvent une réaction d'incompréhension : quelle pertinence la transmission de grandeurs physiques peut-elle avoir pour les relations causales auxquelles on s'intéresse dans la vie courante et dans les sciences autres que la physique ? Quel rapport entre un accident de voiture causé par un taux d'alcoolémie trop élevé du conducteur, et des transmissions d'énergie ? Quel rapport entre les souriceaux dont la mort a été causée par l'inanition, elle-même due à une modification génétique de l'ADN de leur mère, et des transmissions de grandeurs conservées ? La réponse passe par l'analyse du rapport complexe entre une explication causale et son vérificateur. L'erreur des théories de la causalité que nous avons rejetées, notamment la théorie nomologique et la théorie contrefactuelle, est de réduire ce qui rend vraie une explication causale à son seul aspect explicatif. Or, les contre-exemples que nous avons examinés montrent que beaucoup d'explications obéissent aux contraintes imposées par ces théories sans pour autant être causales. Nous avons fait l'hypothèse que ce qui distingue les explications causales des autres explications est la présence d'une relation de transfert entre les événements sur lesquels porte l'explication. Cette hypothèse a été confirmée lorsque nous avons pu constater que les explications qui semblent intuitivement non-causales sont précisément celles qui portent sur des événements non reliés par une transmission.

Néanmoins, il faut éviter l'erreur inverse qui consisterait à réduire le vérificateur des explications causales à la seule transmission. Il y a bien transmission d'énergie entre le cerveau de l'automobiliste, son pied, la pédale d'accélération, la voiture et l'arbre contre lequel elle finit par se heurter. Mais cette transmission ne rend pas à elle seule vraie l'explication, et elle ne contient surtout normalement pas l'information la plus pertinente pour les intérêts pratique que l'on poursuit : pour déterminer si le taux d'alcool a été un facteur déterminant, on cherche plutôt à savoir s'il existe une relation de dépendance nomique entre le taux d'alcool et le contrôle des mouvements volontaires. On veut aussi savoir contrefactuellement, si le conducteur aurait, s'il avait moins bu mais dans des circonstances par ailleurs identiques, gardé le contrôle de la voiture. Notre analyse tient compte du fait que ces relations sont les plus directement visées dans des explications causales. Nous rendons compte également du fait qu'une bonne explication causale doit donner la réponse aux questions de dépendance nomique et de dépendance contrefactuelle : le vérificateur d'une explication causale vraie est d'abord une relation de responsabilité causale entre deux faits qui ne contiennent en général aucune référence aux grandeurs conservées physiques. Puis, l'analyse de la relation de responsabilité causale montre que la dépendance nomique en est un ingrédient essentiel. Seulement, le seul fait qu'une explication fasse référence à une relation de dépendance nomique ne suffit pas pour en faire une explication *causale* ; ce qui la rend causale, c'est que la relation à laquelle elle fait référence contient un élément de transmission. C'est cet élément qui rend justice à l'intuition mécaniste qui accompagne tous nos jugements proprement causaux, alors que l'élément de dépendance nomique rend compte de l'intuition explicative.

Est-il possible qu'une explication qui mentionne un fait négatif en guise d'*explanans* ou d'*explanandum* soit néanmoins causale ? Elle ne peut certainement pas l'être au sens plein d'une explication qui nomme à la fois l'événement qui est la cause mécanique et la propriété causalement responsable. Nous avons distingué deux formes d'explications de forme négative. Il y a des explications par des faits négatifs qui sont néanmoins causales dans la mesure où l'*explanans* désigne l'événement qui est la cause mécanique (c'est-à-dire la source de la transmission) même s'il ne désigne pas explicitement la propriété responsable de l'*explanandum*. D'autres explications par des faits négatifs ne sont causales que dans un sens dérivé : l'*explanans* ne fait pas référence à l'événement qui est la cause mécanique de l'*explanandum* ; au lieu de cela, il contient l'information que la relation entre *explanans* et *explanandum* n'est pas « normale », où les connaissances d'arrière-plan permettent d'inférer que le pendant positif de l'*explanans* négatif est normalement la cause (au sens littéral) du pendant positif de l'*explanandum* négatif. Lorsque l'absence du gène est responsable de la mort des souriceaux, cette absence ne cause pas la mort au sens littéral et mécanique. Cependant, elle l'explique par référence à une dépendance nomique, et la formulation permet d'inférer qu'il s'agit d'une situation anormale, en ce sens que le fait mentionné dans l'*explanans* est en contraste avec la présence normale du gène qui cause alors, dans les circonstances normales, la survie des souriceaux⁵⁶.

⁵⁶ Je remercie Jérôme Pelletier, Véronique Antomarchi et Bruno Gnassounou pour leur lecture critique de cet article.