

La causalité dans la philosophie contemporaine

Max Kistler

► **To cite this version:**

Max Kistler. La causalité dans la philosophie contemporaine. *Intellectica - La revue de l'Association pour la Recherche sur les sciences de la Cognition (ARCo), Association pour la Recherche sur la Cognition*, 2004, 38 (1), pp.139-185. ijn_00000549

HAL Id: ijn_00000549

https://jeannicod.ccsd.cnrs.fr/ijn_00000549

Submitted on 5 Nov 2004

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

La causalité dans la philosophie contemporaine

Max Kistler*

Résumé : La réflexion philosophique moderne sur la notion de causalité s'articule autour de son rapport au concept de loi de la nature. Cet article présente et soumet à un examen critique les conceptions majeures de la causalité développées au XXe siècle : la conception « déductive-nomologique » assimile la relation causale à une explication. Selon l'approche « contrefactuelle », le rapport entre cause et effet se réduit au rapport de dépendance contrefactuelle. L'approche interventionniste ou « agentive » analyse le rapport causal à partir du rapport entre un agent et son action. L'analyse probabiliste réduit la relation causale à celle d'augmentation de probabilité. A la lumière de l'analyse critique de ces théories, il apparaît que la réduction de la causalité au transfert d'une quantité individuelle d'une grandeur conservée, en particulier d'énergie, est en mesure d'éviter les difficultés auxquelles elles se heurtent. Cependant, la relation objective qui correspond à la notion commune de causalité contient, outre le transfert entre événements, un aspect nomologique : un fait est causalement responsable d'un autre fait s'il existe une loi de la nature en vertu de laquelle le premier détermine le second.

Mots-clés : Causalité, loi, explication, dépendance contrefactuelle, probabilité conditionnelle, action, transmission, responsabilité causale.

Abstract: **Causality in contemporary philosophy.** Modern philosophical thinking about the notion of causation revolves around its relation to the concept of law of nature. This article presents and critically examines the major conceptions of causation developed in the 20th century: the deductive-nomological approach assimilates causation to explanation. The counterfactual analysis reduces the relation between cause and effect in terms of the relation of counterfactual dependence. The interventionist or agentive approach analyses causation in terms of the relation between an agent and its actions. According to the probabilistic analysis, causation is grounded on the augmentation of probability. In light of our analysis of these accounts, it appears that the reduction of causation to the transference of an individual amount of a conserved quantity, in particular energy, can avoid the difficulties encountered by those theories. However, the common notion of causation expresses an objective relation containing,

* Département de philosophie, Université Paris X – Nanterre, 200 av de la République, 92001 Nanterre Cedex. Institut Jean Nicod – CNRS, 1 bis, avenue de Lowendal, 75007 Paris, e-mail: kistler@ehess.fr

beyond the relation of transference, an aspect of causal responsibility, by which one fact determines another in virtue of a law of nature.

Keywords: Causation, causality, law of nature, explanation, counterfactual dependence, conditional probability, action, transference, causal responsibility.

La causalité continue aujourd'hui de faire l'objet d'intenses débats en philosophie des sciences. Son destin de concept philosophique respectable semblait pourtant scellé au début du XXe siècle lorsque Russell le déclarait définitivement obsolète : dans son célèbre article « On the Concept of Cause », Russell défend premièrement la thèse historique selon laquelle la disparition de la terminologie causale dans une branche scientifique est un indice du fait que cette branche est entrée dans sa maturité. Cette thèse est surtout fondée sur l'examen de la physique théorique où selon Russell, le concept de dépendance fonctionnelle remplace progressivement celui de causalité. Deuxièmement, il juge qu'il s'agit là d'une évolution salutaire que la philosophie ferait bien de suivre : « la complète extrusion [du mot « cause »] du vocabulaire philosophique [est] désirable » (Russell 1912, p. 173).

Russell donne trois arguments pour cette dernière thèse normative selon laquelle « la raison pour laquelle la physique a cessé de chercher des causes est qu'il n'existe pas de telles choses » et selon laquelle tout l'appareillage conceptuel de la causalité est « une relique d'une époque révolue qui ne survit, comme la monarchie, que parce qu'on suppose à tort qu'elle ne fait pas de dégâts » (Russell 1912, p. 173)¹. Son premier argument est celui de l'anthropomorphisme : c'est l'expérience de notre propre volonté qui donne son sens au concept de cause, et ce n'est que le fruit d'un préjugé non fondé que d'appliquer ce concept aux processus naturels dépourvus de volonté. « Rien ne peut opérer en dehors d'un acte de volonté. La croyance selon laquelle les causes opèrent résulte de leur assimilation, consciente ou inconsciente, aux actes de volonté. » (Russell 1912, p. 183)² Cette critique ne remet pas en cause la réalité de la causalité en elle-même, mais seulement le bien-fondé des attributions de causalité à la nature inanimée. Ce qui montre l'inadéquation du concept de volonté pour la description de processus naturels, c'est qu'il contient les idées d'une différence objective entre un facteur actif et des fac-

¹ Pour l'analyse de l'argumentation de Russell, je suis redevable à Keil (2000, chap. II, 4).

² Russell maintient sa condamnation de cet aspect anthropomorphique de la notion courante de causalité lors qu'il s'apprête, deux ans plus tard, à réhabiliter le concept de causalité et de loi causale : « Historically, the notion of cause has been bound up with that of human volition. [...] The cause is supposed to be "active", the effect "passive". [...] But all such ideas, as applied to physics, are mere anthropomorphic superstitions » (Russell 1914, p. 227).

teurs passifs et surtout celle d'une asymétrie de détermination et de situation dans le temps. Ce qu'on trouve dans la nature, à en croire les équations différentielles de la physique théorique, ce sont des relations de dépendance et de détermination qui sont symétriques par rapport au temps et qui ne présentent aucun point d'ancrage pour la distinction entre activité et passivité. Les régularités découvertes dans les processus naturels déterminent l'état d'un système dans les deux directions temporelles à la fois : étant donné l'état d'un système (les « conditions initiales »), comme le système planétaire, et les lois qui gouvernent son changement dans le temps (il vaut mieux éviter le terme « évolution » puisqu'il suggère une détermination asymétrique du futur par le passé), les états du système dans le passé sont déterminés exactement de la même manière que les états du système dans l'avenir.

Le deuxième argument de Russell attaque la causalité par le biais de son association avec l'une des variantes du « principe de causalité » selon lequel la causalité engendre des régularités, ou même si l'on suit Hume, consiste en régularités. Selon ce principe, la même cause engendre toujours les mêmes effets. Or il est essentiel que le concept de causalité puisse s'appliquer aux événements macroscopiques à partir desquels nous l'apprenons et qui constituent des instances d'application paradigmatiques. L'avalanche est la cause de la destruction de la maison. Le choc de la voiture contre l'arbre est la cause des blessures des passagers. Ces événements macroscopiques sont d'une formidable complexité, et leur complexité contraste fortement avec les conditions simples qui sont décrites par les lois formelles de la physique. Russell soutient que les exigences (pourtant toutes les deux caractéristiques de la causalité) de s'appliquer à des événements complexes macroscopiques et de suivre la régularité stricte d'une loi de la nature sont irréconciliables. Plus la description d'un événement macroscopique complexe s'approche d'une complétude et d'un détail suffisant pour permettre d'y appliquer les lois de la physique, moins il semble possible que l'événement se répète. « Dès lors que les antécédents ont été indiqués avec une complétude suffisante pour permettre le calcul du conséquent avec quelque précision, les antécédents seront devenus si complexes qu'il est très improbable qu'ils se répéteront. » (Russell 1912, p. 181)³ Hume avait tenu compte du fait que les événements macroscopiques qui sont les causes et effets paradigmatiques de notre schème conceptuel ne se

3 Cette observation Russellienne a été récemment reprise par N. Cartwright qui en tire pourtant la conclusion opposée : « Je maintiens exactement la position inverse par rapport à celle de Russell. Je suis favorable aux causes et opposée aux lois. » (Cartwright 1983, p. 74). L'observation commune de Russell et de Cartwright est qu'il n'existe pas de régularités strictes au niveau des entités qui jouent le rôle de termes des relations causales : les événements concrets.

répètent jamais exactement, en n'exigeant des régularités constitutives de la causalité que la *similarité* et non l'identité qualitative. Selon la fameuse définition de Hume, une cause est « un objet antérieur et contigu à un autre, tel que tous les objets semblables au premier soient placés dans une relation semblable de priorité et de contiguïté par rapport à des objets semblables au second » (Hume 1739, p. 171, tr. p. 259; cf. aussi Hume 1748, p. 76, tr. p. 124/5). Mais nous sommes aujourd'hui plus sensibles que ne l'était Hume au fait que le retrait vers l'exigence vague de similarité risque de rendre la thèse de la régularité triviale : comme Quine l'a bien expliqué, d'un point de vue logique tout est semblable et dissemblable de tout, puisque n'importe quelle paire d'objets partage un nombre infini de propriétés et diffère par un nombre infini de propriétés.

Le retrait vers la notion vague de similarité menace de ruiner le projet de sauver le concept de causalité dans le cadre d'une conception scientifique du monde, de ses résonances anthropomorphiques ou animistes, en le réduisant au concept de régularité nomique⁴. Depuis Galilée⁵, les philosophes ont interprété le projet de la science moderne comme celui de substituer à la recherche des causes la recherche des lois de la nature⁶. Or cette substitution ne peut conduire

⁴ « Nomique » - du grec « nomos », la loi – est l'adjectif relatif au nom « loi » que l'on utilise lorsqu'il s'agit de lois de la nature, « légal » étant réservé pour la loi au sens juridique. Dans cet article, j'utiliserai le terme « nomique » pour désigner ce qui est relatif aux lois elles-mêmes, alors que « nomologique » désigne ce qui est relatif au discours et aux théories sur les lois de la nature.

⁵ Galilée exprime l'idée d'une telle purification de la conception scientifique d'un phénomène naturel, en substituant à la terminologie obscure de la causalité la terminologie plus claire de la régularité des processus observables, sans utiliser le terme de « loi » : « L'occasion ne me semble pas favorable pour rechercher la cause de l'accélération du mouvement naturel, problème sur lequel différents philosophes ont formulé différentes opinions. [...] Pour le moment le but de notre auteur est seulement de nous faire comprendre qu'il a voulu découvrir et démontrer quelques propriétés d'un mouvement accéléré (quelle que soit la cause de son accélération). » (Galilée 1638, p. 202, trad. p. 135) Sur la causalité chez Galilée, cf. Yakira (1994).

⁶ Deux passages qui témoignent de cette conviction largement partagée, par Berkeley : « la physique ou la mécanique se proposent de fournir, non les causes efficientes, mais seulement les règles des impulsions et des attractions et, pour tout dire, les lois des mouvements, puis, sur ces bases, de donner une explication des phénomènes particuliers, et non leur cause efficiente. » (Berkeley 1721, § 35, p. 40 ; trad. p. 166), et Comte : « La révolution fondamentale qui caractérise la virilité de notre intelligence consiste essentiellement à substituer partout, à l'inaccessible détermination des causes proprement dites, la simple recherche des lois, c'est-à-dire des relations constantes qui existent entre les phénomènes observés. Qu'il s'agisse des moindres ou des plus sublimes effets, de choc et de pesanteur comme de pensée et de moralité, nous n'y pouvons vraiment connaître que les diverses *Intellectica*, 2004/1, 38

à une conception plus scientifique des processus naturels que si le concept de régularité ou de loi, censé prendre la relève du concept de causalité, permet une formulation quantitativement précise de la détermination de l'effet par la cause. La substitution de la variante nomologique du principe de causalité qui dit que les mêmes causes ont toujours les mêmes effets, au principe plus abstrait selon lequel tout événement a une cause, ne constitue un progrès que s'il existe des régularités qui le sont au sens strict et quantitatif et qui correspondent aux causes et effets qui sont les termes de la relation causale du sens commun. L'argument de Russell consiste à dire que cette substitution ne peut se faire qu'au prix d'un abandon du concept de causalité puisque les termes des lois sont incommensurables avec les termes de la relation causale : les lois fonctionnelles de la physique, exprimées en équations différentielles, relient un nombre restreint de quantités de grandeurs mesurables possédées par un système, alors que la causalité relie des événements concrets d'une complexité infinie. On ne peut pas, comme le propose Hume, réduire la causalité à une instance de régularité puisqu'il n'existe aucune régularité sur le plan des événements qui sont les termes de la relation causale. Les régularités qui font l'objet des lois strictes de la physique ne concernent pas les événements concrets, mais certaines de leurs propriétés mesurables. Les lois de la physique ne méritent cette appellation que parce qu'elles sont strictes ; elles expriment des rapports de détermination quantitativement précise. Dans le cadre d'une telle conception des lois, ce serait un leurre de vouloir avoir recours aux régularités approximatives qui correspondent aux régularités entre événements « similaires » de Hume⁷. Une régularité où les termes de la succession ne sont pas identiques mais seulement semblables n'est pas une régularité scientifique du tout. Par conséquent, la substitution de la recherche des lois à la recherche des causes qui caractérise la science moderne, et qu'accomplit la physique théorique, ne permet pas de réduire le concept de causalité - ce qui signifierait qu'il serait conservé et même légitimé - mais mène à son élimination.

Le troisième argument de Russell renforce la conclusion du second. Non seulement il n'existe pas de régularités au niveau des événements concrets (c'était la conclusion du second argument), mais la physique du 20^{ème} siècle ne connaît même plus de lois de succession du tout. Le progrès de la physique vers sa maturité s'accompagne d'un remplacement progressif des lois de succession par des lois de coexistence ou d'équivalence qui forment le contenu des équations exprimant des dépendances fonctionnelles.

liaisons mutuelles propres à leur accomplissement, sans jamais pénétrer le mystère de leur production. » (Comte 1844, p. 19/20)

⁷ En ce sens, "Hume missed the difficulty that lay hidden in the concept of similarity" (Davidson 1995, p. 268).

« Ces lois de succession probable tendent, malgré leur utilité dans la vie de tous les jours et pendant l'enfance d'une science, à être remplacées par des lois assez différentes dès qu'une science est couronnée de succès. La loi de gravitation illustrera ce qui se passe dans toutes les sciences avancées. Dans les mouvements des corps qui gravitent mutuellement, il n'existe rien qui puisse être appelé une cause et rien qui puisse être appelé un effet ; il y a seulement une formule. On peut trouver certaines équations différentielles qui valent à tout instant pour chaque particule du système, et qui, étant donné la configuration et les vitesses à un instant, ou la configuration à deux instants, rendent théoriquement calculable la configuration à n'importe quel instant antérieur ou postérieur. En d'autres termes, la configuration à n'importe quel instant est une fonction de cet instant et des configurations à deux instants donnés. Cette affirmation est vraie à travers toute la physique, et non seulement dans le cas particulier de la gravitation. Mais dans un tel système il n'existe rien qui puisse correctement être appelé « cause » et rien qui puisse correctement être appelé « effet » ». (Russell 1912, p. 186)

Ce qui explique selon Russell l'échec définitif du programme moderne de sauvetage du concept de causalité en le réduisant à celui de régularité nomique, c'est sa thèse sur le type des lois qui caractérisent la science mûre. Selon la conception nomologique de la causalité caractéristique de l'époque moderne⁸, si deux événements *c* et *e* sont reliés comme la cause à l'effet, alors les événements *c* et *e* sont des instances de types d'événements de telle sorte qu'il existe une loi de la nature selon laquelle tous les événements du même type que *c* sont suivis d'un événement du type de *e*. Or, selon Russell, les lois fonctionnelles de la physique moderne sont incapables de jouer ce rôle. On peut reconstruire l'argument dans la forme d'une réduction à l'absurde : les deux propriétés suivantes sont essentielles au concept de causalité. Premièrement, les termes de la relation causale sont séparés dans le temps, la cause étant antérieure à l'effet. Deuxièmement, il y a une asymétrie entre cause et effet qui est intrinsèque, c'est-à-dire non pas relative à la description qu'on en fait : la cause produit, ou détermine l'effet, alors que l'inverse n'est pas vrai. Or si l'on essaye de réduire la causalité à une succession régulière conforme à une loi de la nature, il faut avoir recours aux seules lois dont la physique mûre reconnaît l'existence, les lois fonctionnelles de coexistence. Or, cela fait disparaître les deux caractéristiques de la causalité car les termes des régularités fonctionnelles ne sont pas séparés dans le temps, et il n'existe entre eux aucune asymétrie objective. Donc, il faut rejeter le principe nomologique de

⁸ Kant propose une justification transcendantale du fondement de la causalité sur la régularité nomique. Davidson (1995) a tenté une reformulation contemporaine d'une telle justification transcendantale.

la causalité qui constituait l'hypothèse de départ de ce raisonnement. A partir de cette conclusion, il apparaît que le rôle du second argument de Russell est de donner une raison conceptuelle pour la non-existence de lois de succession entre événements complexes. Il est important pour la démonstration russellienne de la déchéance du concept de causalité, de montrer que l'absence de telles lois n'est pas seulement un fait historique contingent mais découle d'une nécessité conceptuelle, à savoir de l'incommensurabilité des termes des lois quantitatives (un nombre restreint de grandeurs mesurables) et des termes des relations causales (les événements concrets qui possèdent un nombre infini de propriétés).

L'ABSORPTION DE LA CAUSALITÉ DANS LE MODÈLE DE L'EXPLICATION DÉDUCTIVE-NOMOLOGIQUE

Les arguments de Russell n'ont pas suffi à convaincre les philosophes de la nécessité d'éliminer tout l'appareil conceptuel de la causalité, ainsi que la terminologie causale. Pour différentes raisons, le concept de causalité est indispensable non seulement pour comprendre nos propres interactions avec notre environnement dans la vie de tous les jours, mais aussi pour comprendre la science. Nancy Cartwright se déclare en accord avec la thèse russellienne selon laquelle « il est impossible de dériver des principes causaux à partir des lois d'association qui sont causalement symétriques », mais conteste sa thèse selon laquelle « les lois d'association sont toutes les lois qui existent » (Cartwright 1983, p. 21). La raison qu'elle avance est que nous sommes capables d'intervenir dans le monde, ce qui présuppose que ces interventions soient soumises à des régularités que nous sommes en outre capables de connaître, au moins approximativement. De manière plus pertinente pour la science, les sciences permettent de développer, d'affiner et surtout d'expliquer l'efficacité de nos « stratégies efficaces » (Cartwright, *ibid.*) d'intervention, et cette capacité présuppose qu'il existe des lois causales. Putnam a également attiré l'attention sur le fait que la notion de causalité est incontournable, non seulement dans les sciences moins complètement parvenues au stade de la formulation de lois quantitatives, comme la biologie, mais aussi dès qu'on conçoit d'un point de vue physique l'interaction de systèmes physiques concrets : « Quand nous en venons aux sciences moins fondamentales que la physique fondamentale des particules, par exemple à la sociologie ou à l'histoire, ou même à la biologie évolutionniste, ou même à la chimie, ou à la physique des solides, alors, bien sûr, nous nous apercevons que l'on invoque constamment la causalité en son sens d'occasion. [...] Même en physique fondamentale des particules, un tel langage devient indispensable lorsque nous *appliquons* la physique à des systèmes réels qui ne sont *pas* (bien entendu) tout

l'univers-considéré-comme-système clos » (Putnam 1984, p. 8, italiques de Putnam; repr. 1990, p. 86; trad. p. 225/6)⁹. Russell lui-même, seulement deux ans après sa critique dévastatrice du concept de cause, entreprend sa réhabilitation partielle¹⁰, toutefois conditionnée par la détermination de laisser en dehors du concept purifié de causalité tous ses aspects anthropomorphiques : la science ne justifie pas l'existence dans la nature d'une nécessité (*compelling*) avec laquelle la cause produirait l'effet ; l'idée d'une telle nécessité provient entièrement de l'analogie avec l'expérience de nos propres actes volontaires¹¹. Il faut également cesser d'assimiler la cause à un facteur « actif » et l'effet à un facteur « passif » ; il faut éliminer du concept de causalité « tout ce qui différencie entre passé et futur » (Russell 1914, p. 227). Cette réhabilitation partielle du concept de cause s'avère d'après Russell nécessaire dès lors qu'on cherche à comprendre l'évolution de systèmes physiques réels qui sont « moins grands que l'état total de l'univers » (Russell 1914, p. 230), ce qui est bien entendu pratiquement indispensable. Il entend désormais justifier l'idée qu'il existe des « relations constantes entre événements de types spécifiés, séparés par des intervalles temporels donnés » (*ibid.*) et qui méritent d'être considérées comme des « lois causales ». De telles lois ne sont pas strictes, mais « susceptibles d'avoir des exceptions » dont l'explication requiert la prise en considération plus complète des événements étudiés.

La réintégration du concept de causalité dans l'appareil conceptuel de la philosophie des sciences trouve son expression canonique dans l'assimilation de la causalité à la relation d'explication, conçue selon le modèle dit « nomologique-déductif ». Cette conception de la causalité faisait partie du cœur doctrinal de la philosophie de l'empirisme logique, et constituait à ce titre la doctrine majoritaire quant à la nature de la causalité, depuis les années 1930 jusqu'aux années 1970 lorsque les critiques et propositions alternatives commencèrent

⁹ Depuis, Putnam a changé d'avis sur la réalité objective de la causalité. Il la considère désormais comme un effet de perspective qui n'apparaît que relativement à notre intérêt dans la recherche d'explications, recherche qui est gouvernée par une logique pragmatique. Cf. Putnam (1992, p. 47sq.).

¹⁰ Russell maintient sa thèse selon laquelle « dans une science suffisamment avancée, le mot « cause » n'apparaîtra dans aucun énoncé de lois invariables » (Russell 1914, p. 223). Mais il admet désormais qu'« il existe néanmoins un usage quelque peu grossier et peu rigoureux du mot « cause » qui peut être préservé. Les uniformités approchées qui conduisent à son emploi préscientifique peuvent s'avérer dans toutes circonstances en dehors de quelques-unes très rares et exceptionnelles, et peut-être même dans toutes les circonstances qui se produisent dans la réalité. Dans de tels cas, il est utile de pouvoir parler de l'événement antécédent comme de la « cause » et de l'événement successif comme de « l'effet » ». (Russell 1914, p. 223)

¹¹ Cf. Russell (1914, p. 229).

à proliférer. Dans la mesure où ces conceptions alternatives ont toujours pour but de surmonter les difficultés que rencontre la doctrine de l'absorption de la causalité dans le modèle déductif-nomologique de l'explication, il faut présenter cette conception en quelque détail¹².

Selon le modèle de l'explication scientifique proposé par Hempel (1942) et développé par Hempel et Oppenheim (1948), une explication est un argument déductif dont la conclusion constitue l'*explanandum* - ce qu'il s'agit d'expliquer - et les prémisses l'*explanans* - ce qui explique. Parmi les prémisses, une partie consiste en « énoncés de conditions antécédentes » et une partie en « lois générales » (Hempel et Oppenheim 1948, p. 249), alors que la conclusion contient une « description du phénomène empirique qu'il s'agit d'expliquer » (*ibid.*). Une explication adéquate doit notamment satisfaire les exigences suivantes : il faut que l'*explanandum* soit une conséquence logique de l'*explanans*, car seule la nécessité logique de la conclusion d'une déduction valide par ses prémisses peut suffisamment « fonder » (notre croyance en la vérité de) l'*explanandum*. Il faut que l'*explanans* contienne des lois générales, de manière essentielle. Autrement dit, une explication n'est pas adéquate s'il est possible de déduire l'*explanandum* des seules prémisses particulières (les conditions initiales). En revanche, Hempel et Oppenheim reconnaissent explicitement la légitimité d'explications purement générales où l'*explanans* ne contient que des lois, et où l'*explanandum* est également une loi. Ainsi on peut intégrer dans le modèle la dérivation de lois à partir de lois de plus grande généralité, comme par exemple la dérivation de la forme générale¹³ des lois de Kepler à partir des lois de Newton. Il faut en outre exiger que l'*explanans* ait un contenu empirique, autrement dit qu'il soit susceptible d'être testé par l'observation ou l'expérimentation, exigence qui découle de la restriction du domaine d'application du modèle à l'explication des seuls phénomènes empiriques. Finalement, une explication ne peut être correcte que si les énoncés de l'*explanans* sont tous vrais. Cela implique que nous devons renoncer à l'ambition de savoir avec certitude si une explication est correcte : à cause du problème de l'induction¹⁴, nous sommes condamnés pour des raisons de principe à ne disposer que de lois hypothétiques dont la vérité objective reste à jamais incertaine.

Hempel et Oppenheim considèrent que toute explication qui satisfait aux exigences du modèle D-N est *ipso facto* une explication

¹² Sur le modèle déductif-nomologique de l'explication et son rapport avec la causalité, cf. Kistler (1999a), chap. 5, et Barberousse, Kistler, Ludwig (2000), chap. 5.

¹³ Dans les lois de Kepler, le Soleil est mentionné en tant qu'objet individuel. Elles ne peuvent donc être dérivées sans prémisses particulières portant notamment sur la masse du Soleil. Mais la dérivation des lois gouvernant le mouvement relatif de deux masses quelconques l'une autour de l'autre, ne requiert pas de telles prémisses particulières.

¹⁴ Cf. Barberousse, Kistler, Ludwig (2000), chap. 2.

causale. Selon eux, on peut donc utiliser ce modèle comme analyse de la relation causale : cela est possible si l'on interprète la conjonction des conditions initiales comme une description de la cause, et l'événement décrit par l'*explanandum* comme son effet¹⁵. Dans la mesure où les lois qui apparaissent dans l'*explanans* sont explicitement considérées comme des énoncés de régularités empiriques, l'assimilation de la relation causale à la relation explicative constitue une revendication de la théorie régulariste de la causalité. Il faut donc s'attendre à ce qu'elle prête toujours le flanc au second et troisième argument de Russell. Il est intéressant d'observer que plus tard, Hempel a dans une certaine mesure reconnu l'inadéquation de principe du modèle D-N comme reconstruction logique de la relation entre cause et effet, et cela pour plusieurs raisons.

Il reconnaît notamment (en accord avec la conclusion du second argument de Russell) que tant que la cause ne comprend pas l'état de l'univers entier à un instant, il n'est à strictement parler pas possible de déduire avec nécessité logique, la description de l'état d'un système limité, à partir de la description d'un autre état du système ainsi que des lois. « Si le système n'est pas isolé, c'est-à-dire si des influences externes pertinentes agissent sur le système, durant l'intervalle temporel entre l'état initial mentionné et l'état à expliquer, alors les circonstances particulières qu'il faut mentionner dans l'*explanans* incluent aussi ces influences extérieures » (Hempel 1965a, p. 352). Plus tard, Hempel (1988) reconnaît plus explicitement que la déduction d'événements particuliers (plus restreints que l'état de l'univers dans sa totalité) n'est jamais strictement possible car « si l'on veut éviter que les prémisses nomologiques des explications soient ouvertement fausses, il faut qu'on reconnaisse explicitement le fait que l'explication repose sur des clauses auxiliaires en général tacites » (Grünbaum 1988, p. x), clauses que Hempel appelle « *provisos* », et que d'autres appellent les clauses « *ceteris paribus* »¹⁶.

La deuxième raison que donne Hempel pour expliquer l'inadéquation du modèle D-N en tant que moyen d'analyser la relation de cause à effet (et qui reprend le troisième argument de Russell) concerne l'existence des « lois de coexistence » (Hempel 1965a, p.

¹⁵ "The type of explanation which has been considered here so far is often referred to as causal explanation. If E describes a particular event, then the antecedent circumstances described in the sentences C₁, C₂, ..., C_k may be said jointly to 'cause' that event, in the sense that there are certain empirical regularities, expressed by the laws L₁, L₂, ..., L_r, which imply that whenever conditions of the kind indicated by C₁, C₂, ... C_k occur, an event of the kind described in E will take place" (Hempel and Oppenheim 1948, p. 250). De manière plus succincte, « en vertu du fait qu'elle présuppose des lois générales qui relient la « cause » et « l'effet », l'explication causale est conforme au modèle D-N » (Hempel 1965a, p. 348).

¹⁶ Sur ce problème, cf. Kistler (1999a, chap. 3 ; et 2003).

352). Hempel admet que ce n'est pas la déductibilité en soi de l'*explanandum* à partir des conditions initiales et de lois générales qui fait de ces conditions la cause de l'événement décrit par l'*explanandum*. Comme contre-exemple à cette thèse trop générale, il présente le cas de l'explication de la période d'un pendule (idéal) par sa longueur. Cette explication qui repose sur la loi de coexistence

entre la période T et la longueur L , selon laquelle $T = 2\pi\sqrt{L/g}$,

relie deux propriétés¹⁷ que le pendule possède en même temps et qui ne peuvent donc pas être dans le rapport de cause à effet qui exige une séparation temporelle¹⁸. De manière plus générale, Hempel admet qu'il existe un certain nombre de types d'explications qui satisfont aux critères du modèle D-N, sans qu'il soit possible de les interpréter comme des explications causales : à côté des explications faisant appel à des lois de coexistence, cela est notamment le cas où ce n'est pas un événement particulier qui est expliqué mais une loi. Dans la déduction des lois de Kepler à partir des lois de Newton, aucun événement particulier n'est nommé qui pourrait être interprété comme cause ou effet.

Finalement, Hempel note qu'il existe des explications satisfaisant toutes les exigences du modèle D-N mais où les conditions « initiales » contiennent une référence à un état du système qui est *postérieur* à l'état décrit par l'*explanandum*. L'explication de la réfraction d'un rayon de lumière lors d'un passage à travers une interface entre deux milieux à indice de réfraction différent, constitue un exemple d'une telle explication : l'*explanans* mentionne, outre le principe de Fermat selon lequel le chemin du rayon est celui qui minimise le temps du parcours, les points de départ et d'arrivée du rayon, pour en déduire le point de passage dans l'interface, alors que l'événement d'arrivée est évidemment postérieur à ce passage. Hempel remarque que la « gêne » (« *uneasiness* » Hempel 1965a, p. 353) produite par des explications de ce genre vient de l'apparente impossibilité de les interpréter comme des explications causales puisqu'il apparaît comme *a priori* exclu que la réfraction, et d'autres cas de ce genre,

¹⁷ Afin de pouvoir attribuer la période au pendule à un instant précis, il faut la considérer comme une « caractéristique dispositionnelle » (Hempel 1965a, p. 352).

¹⁸ Cummins (1983, p. 3) utilise cet exemple pour réfuter plus généralement l'identification de l'explication scientifique avec la subsumption sous une loi causale. Hempel (1965a, p. 352 sq.) note, à propos de cet exemple, qu'on accepterait intuitivement plus volontiers une explication de la période du pendule par sa longueur que l'explication inverse de sa longueur par sa période. Or rien dans les exigences logiques posées à l'explication D-N ne permet de justifier cette préférence intuitive, étant donné que la loi reliant les deux grandeurs est une loi symétrique de coexistence qui exprime à la fois la dépendance de la période par rapport à la longueur et la dépendance de la longueur par rapport à la période.

constituent des cas de causalité rétroactive où l'effet précède la cause¹⁹. Mais Hempel se déclare convaincu de l'impossibilité de formuler les conditions générales qu'il faudrait imposer à une explication D-N afin qu'elle puisse être interprétée comme causale, c'est-à-dire comme constituant une analyse de la relation causale.

La plupart des présentations de la doctrine qu'on peut appeler la « théorie nomologique-déductive de la causalité », par les principaux philosophes de la tradition de l'empirisme logique prennent moins de précautions que Hempel et identifient directement la relation de cause à effet avec la relation entre conditions initiales et *explanandum* dans une explication D-N correcte. A titre d'exemple, selon Carnap, « Que signifie alors l'énoncé : l'événement B est causé par l'événement A? Qu'il existe dans la nature certaines lois dont on peut déduire logiquement l'événement B, à condition de les conjuguer avec la description exhaustive de l'événement A » (Carnap 1966, p. 194; trad. p. 189). Popper, pourtant à bien des égards critique de l'empirisme logique, identifie lui aussi l'explication causale avec l'explication D-N tout court : « Donner une *explication causale* d'un événement signifie déduire un énoncé le décrivant en utilisant comme prémisses de la déduction une ou plusieurs *lois universelles* et certains énoncés singuliers. » (Popper 1934, trad. p. 57; italiques de Popper)²⁰

Comme nous l'avons vu, Hempel a indiqué un certain nombre de raisons qui s'opposent à cette doctrine : on ne peut pas réduire le concept de relation causale à la relation entre les conditions initiales et l'*explanandum*, dans une explication qui correspond au modèle D-N. Les contre-exemples à la doctrine déductive-nomologique de la

¹⁹ La question de savoir s'il existe des cas de causalité rétroactive en mécanique quantique est controversée. Cf. Reichenbach (1956) et Dowe (1996).

²⁰ Il existe de nombreuses variantes de cette doctrine. Stegmüller exprime la même idée dans un registre sémantique, sur le plan des conditions de vérité des énoncés causaux. « Un énoncé causal singulier implique un énoncé exprimant une régularité générale. [...] Là où l'on constate une relation causale individuelle, on affirme qu'une séquence particulière d'événements peut être subsumée sous une loi générale de la nature (admise à titre hypothétique) ». (Stegmüller 1969, p. 512/3) Davidson défend la doctrine de l'équivalence entre une relation causale particulière et l'existence d'une loi générale qui permettrait de déduire une description de l'effet à partir d'une description appropriée de la cause, dans le cadre de notre justification de la croyance en la vérité des énoncés causaux particuliers : « Il ne s'ensuit pas que si nous savons que tel énoncé causal est vrai nous devons être capables d'en tirer une loi ; tout ce qui s'ensuit est que nous savons qu'il doit y avoir une loi de couverture (*covering law*). Et très souvent, je pense, les raisons que nous pouvons avoir d'accepter un énoncé causal singulier tiennent au fait que nous avons de bonnes raisons de croire qu'une loi causale appropriée existe, bien que nous ne sachions pas quelle elle est. » (Davidson 1967, p. 160 ; trad. p. 213)

causalité peuvent être regroupés en deux catégories²¹ : la première consiste en explications D-N où le rapport entre *explanans* et *explanandum* se situe sur un plan totalement étranger à celui des relations causales entre événements particuliers. La deuxième catégorie contient des explications qui concernent des processus causaux mais où les conditions initiales ne peuvent pas être identifiées avec la cause, ni l'*explanandum* avec l'effet. A la première catégorie appartiennent :

a) l'explication des lois générales : par exemple celle des lois de Kepler à partir des lois de Newton;

b) l'explication logique et mathématique, par déduction à partir d'axiomes qui ne sont pas interprétés ou dont l'interprétation consiste en modèles abstraits.

c) l'explication par des lois de coexistence : par exemple celle de la période par la longueur du pendule ;

d) l'explication analytique ou compositionnelle²², en particulier l'explication fonctionnelle : un type important d'explications scientifiques concerne non pas l'occurrence d'un événement ni l'acquisition d'une propriété par un objet, mais plutôt la nature d'une propriété. Par exemple, la physique statistique parvient à expliquer en quoi consiste la propriété d'être à une certaine température, ou à indiquer la nature sous-jacente de cette propriété, en indiquant sa base de réduction, l'énergie cinétique moyenne des composantes de l'objet qui possède la température. Il n'existe entre les propriétés des composantes et leur relations mutuelles et la propriété ainsi réduite de l'objet complexe aucune relation causale, ne serait-ce que parce qu'il s'agit de la possession de différentes propriétés par des objets qui se trouvent au même endroit en même temps, alors que la relation causale requiert la séparation de la cause et de l'effet à la fois dans le temps et dans l'espace. Un cas particulier de ce type d'explication non-causale concerne les propriétés dispositionnelles : on explique la solubilité du sucre dans l'eau par les propriétés des composantes moléculaires du sucre et de l'eau.

e) Il existe des explications triviales qui satisfont les exigences purement formelles du modèle D-N, mais où la loi utilisée est triviale au lieu d'être une véritable loi de la nature. On peut « expliquer » le fait que Jean (un homme) ne soit pas tombé enceint (sic), en le dé-

²¹ Un certain nombre de ces contre-exemples ont été énumérés par Humphreys (1989, p. 300/1) et Salmon (1990, p. 46-50).

²² Cummins (1983) distingue les explications de ces types de l'explication causale par subsomption sous une loi causale. Par ailleurs, Cummins distingue l'analyse compositionnelle d'un *système* de l'analyse d'une *propriété* qu'il appelle « analyse fonctionnelle » (Cummins 1983, p. 15) lorsque la propriété analysée est dispositionnelle, et « analyse de propriété » lorsqu'elle ne l'est pas.

duisant du fait qu'il a pris la pilule contraceptive pendant toute l'année et de la « loi » selon laquelle tous les hommes qui prennent régulièrement la pilule contraceptive évitent de tomber enceints. Il n'existe aucun lien pertinent et encore moins de lien causal entre la prise de la pilule et l'absence de la grossesse chez les hommes.

La seconde catégorie de contre-exemples concerne les cas où l'explication concerne bien un processus causal mais où les conditions initiales ne désignent pas la (ou une) cause de l'*explanandum*.

a) les explications à partir de principes extrémaux, telle que l'explication de la réfraction de la lumière à partir du principe de Fermat sur le chemin qui minimise le temps de parcours. Dans ces explications, l'identification de la cause avec les conditions « initiales » est impossible parce qu'elles contiennent la description d'un événement qui a lieu *après* l'événement décrit par l'*explanandum*.

b) Le modèle D-N autorise l'explication de la cause par l'effet. Pour reprendre un exemple célèbre²³, une tour projette une ombre. L'explication de la hauteur de la tour par la longueur de l'ombre satisfait aux exigences du modèle tout aussi bien que l'explication de la longueur de l'ombre par la hauteur de la tour. Pourtant, seule la seconde autorise l'interprétation des conditions initiales comme la cause de l'*explanandum*.

c) Le modèle D-N n'exclut pas non plus l'explication d'un événement par un autre si les événements sont reliés comme les effets d'une cause commune, mais non par une relation causale entre eux. Ainsi, l'explication de la tempête par la chute du baromètre est légitime dès lors qu'elle repose sur un lien nomique, c'est-à-dire non seulement accidentel, entre les chutes de baromètre et les tempêtes qui ont lieu peu de temps après. Pourtant, les conditions initiales ne désignent pas la cause de l'*explanandum*, mais plutôt un autre effet de leur cause commune qu'est la chute de la pression de l'air. Il est clair qu'il n'existe pas de chaîne causale qui ferait de la chute du baromètre une cause, même indirecte, de la tempête.

c) Kyburg (1965) a remarqué que le modèle D-N rend légitime l'explication du fait qu'une pincée de sel de cuisine se dissout dans de l'eau tiède, à partir du fait que ce sel a été ensorcelé et que nécessairement, tous les échantillons de sel ensorcelé se dissolvent dans l'eau tiède. Il y a effectivement une relation de cause à effet en jeu ; mais la présence dans les conditions initiales d'une information sans pertinence pour l'effet empêche l'identification de l'ensemble des conditions initiales avec la cause.

d) Scriven (1962) a découvert un contre-exemple qui est semblable au précédent dans sa structure, mais qui pose un problème plus sérieux, aujourd'hui connu comme le « problème de la préemption »

²³ Cf. Bromberger (1966, p. 83).

(Lewis 1986, p. 171) : le cas d'une surdétermination causale potentielle. Un pont s'écroule à cause du poids trop élevé des charges qui pèsent sur lui. Une bombe qui aurait été suffisante pour détruire le pont explose juste après son écroulement. L'explication de l'écroulement à partir de l'explosion satisfait aux exigences du modèle D-N mais ne mentionne pas la cause de cet événement. On dit que l'efficacité de la bombe a été « préemptée » par l'intervention de la surcharge. Selon Hempel (1965a, p. 420), il est possible de sauver le modèle de la réfutation par des cas de ce genre si l'on exige la prise en compte du déroulement du processus gouverné par la loi dans le temps. A y regarder de plus près, la loi qui permet de déduire la destruction du pont à partir de l'explosion concerne la propagation de l'onde de choc à partir de l'endroit spatio-temporel de l'explosion, à une vitesse finie. Or dans le cas envisagé par Scriven, l'explication D-N n'aboutira pas parce qu'au moment où l'onde de choc atteint l'endroit où se trouvait le pont, celui-ci n'existe déjà plus et elle ne peut donc pas le détruire.

Cet ensemble de contre-exemples n'a pas pour seul intérêt de faire apparaître le caractère inadéquat de la théorie nomologique-déductive de la causalité, forme contemporaine de la théorie régulariste de Hume. C'est le nombre et surtout la diversité des types de contre-exemples qui constitue la motivation principale dans la recherche d'une théorie de rechange qui analyse la causalité sur des bases nouvelles²⁴. Les quatre théories majeures dont on débat actuellement devront donc se laisser juger à l'aune de leur capacité d'éviter ces contre-exemples. Nous considérerons tour à tour la théorie contrefactuelle, interventionniste, probabiliste et de transfert.

CONCEPTION CONTREFACTUELLE

David Lewis (1986) a proposé de retourner à la seconde définition de la cause par Hume selon laquelle « un objet est la cause d'un autre » signifie que « si le premier objet n'avait pas existé, le second n'aurait jamais existé » (Hume 1748, p. 76, tr. p. 124/5). La proposition de Lewis consiste à réduire la causalité entre deux événements à l'existence d'une chaîne causale intermédiaire qui est une séquence finie d'événements dont chaque membre dépend causalement du précédent²⁵. La *dépendance causale* entre événements est analysée

²⁴ Ainsi Lewis motive sa proposition d'une théorie contrefactuelle de la causalité, par l'énumération des problèmes qui se posent à la théorie déductive-nomologique : "I have no proof that regularity theories are beyond repair, nor any space to review the repairs that have been tried. Suffice it to say that the prospects look dark. I think it is time to give up and try something else." (Lewis 1986, p. 160)

²⁵ La précédence est ici entendue au sens logique et est déterminée par l'ordre des membres de la chaîne, sans préjuger de la question de la précédence temporelle de la cause par rapport à l'effet. Lewis donne en effet un argument indépendant contre la possibilité de

par l'intermédiaire de la dépendance *contre-factuelle* entre les propositions qui affirment que ces événements ont lieu. Selon la théorie contre-factuelle de la causalité, l'événement *e* dépend causalement de l'événement *c* si les deux contre-factuels suivants sont vrais : si *c* avait eu lieu, *e* aurait eu lieu ; si *c* n'avait pas eu lieu, *e* n'aurait pas eu lieu²⁶. La distinction entre la dépendance causale et la relation causale elle-même est introduite afin de garantir la transitivité de la relation causale, car la dépendance contre-factuelle ne l'est pas²⁷. On peut par exemple concevoir une situation où les prémisses de l'argument suivant sont vraies, alors que sa conclusion est fautive²⁸ :

Si Otto était allé à la fête, alors Anna y serait allée.

Si Anna était allée, alors Waldo y serait allé.

Si Otto était allé, alors Waldo y serait allé.

La conclusion peut par exemple être fautive parce que Waldo est le rival malheureux d'Otto pour les faveurs d'Anna, et fait tout pour éviter son concurrent. On peut expliquer la non-transitivité des contre-factuels par le fait que leur évaluation présuppose implicitement une situation concrète et que la situation présupposée dans la conclusion diffère de celles présupposées par les prémisses²⁹.

La proposition de Lewis est confrontée à certains des contre-exemples à la théorie déductive-nomologique.

a) Il semble exister des énoncés contre-factuels qui affirment la dépendance de la cause par rapport à l'effet. Selon un tel contre-factuel « à rebours » (« *backtracking* »), s'il n'y avait pas eu d'éclipse solaire le 11 août 1999, le Soleil, la Lune et la Terre auraient eu des

relations causales « rétroactives » où l'effet précède la cause dans le temps. Je présenterai cet argument plus loin pour le contester.

²⁶ Le premier contre-factuel porte le poids de l'affirmation de causalité dans le cas où les événements *c* et *e* n'ont pas eu lieu, tandis que le second est décisif dans le cas où il s'agit d'événements qui ont réellement eu lieu. Cf. Lewis (1986, p. 167).

²⁷ "Causation must always be transitive; causal dependence may not be so." (Lewis 1986, p. 167) Parmi les défenseurs de l'approche contre-factuelle, certains remettent en doute le bien-fondé de la décision de Lewis de garantir la transitivité de la causalité par stipulation. Cf. Keil (2000, p. 297). Dans le cadre d'une théorie qui cherche à réduire la causalité à la dépendance contre-factuelle, il semble en effet plausible de se laisser guider par le fait que la dépendance contre-factuelle n'est pas transitive, et d'interpréter ce fait comme une raison de penser que la causalité elle-même n'est pas transitive non plus. Hall (2000) montre que si l'on admet les « événements négatifs », tels que les absences et les omissions parmi les causes et effets, alors la thèse selon laquelle la dépendance contre-factuelle est suffisante pour l'existence d'une relation causale entre en conflit avec la thèse de la transitivité de la causalité.

²⁸ L'exemple est dû à Lewis (1973, p. 32).

²⁹ Cf. Keil (2000, p. 295/6).

positions ou des vitesses différentes le 10 août 1999. Lewis appelle cela « le problème des effets » (Lewis 1986, p. 170).

b) De manière semblable, deux effets d'une cause commune semblent dépendre contrefactuellement l'un de l'autre : étant donné que M. Dupont et M. Dupond regardent tous les matches de football transmis par la chaîne de télévision XY (et seulement ceux-là), il semble vrai de dire que si M. Dupont avait vu le match le jour J, M. Dupond l'aurait vu (parce qu'il a été transmis par XY), et si M. Dupont ne l'avait pas vu, M. Dupond ne l'aurait pas vu non plus (parce qu'il n'a pas été transmis par XY). C'est ce que Lewis appelle le « problème des épiphénomènes » (Lewis 1986, p. 170).

c) Enfin, cette analyse semble fournir un résultat incorrect dans les cas de préemption et de surdétermination causale. Si un prisonnier est exécuté par deux balles *A* et *B* tirées par deux bourreaux en même temps, où chacune aurait été suffisante pour sa mort, il s'agit de surdétermination causale, ou, dans les termes de Lewis, de « causalité redondante » (Lewis 1986, p. 193). Chaque balle semble intuitivement être une cause de la mort, mais il n'est pas vrai que si par exemple la balle *B* n'avait pas été tirée, le prisonnier ne serait pas mort. Pour reprendre l'exemple de préemption déjà mentionné plus haut, si le pont n'avait pas été surchargé, le pont se serait écroulé quand même ; l'analyse contrefactuelle doit donc nier, de manière erronée, que la surcharge est, dans la situation esquissée, la cause de l'écroulement du pont³⁰.

³⁰ Lewis (1986, p. 205-7) propose de justifier l'intuition que la surcharge est la cause de l'écroulement quoiqu'il n'en dépende pas contrefactuellement, en faisant valoir que l'écroulement est néanmoins *quasi-dépendant* de la surcharge : la relation entre la surcharge et l'écroulement est causale parce qu'elle est intrinsèquement semblable à d'autres relations entre surcharge et écroulement où aucune cause préemptée n'est présente. Dans ces situations-là, il y a dépendance contrefactuelle, et il y a donc quasi-dépendance dans la situation où la cause préemptée est présente. Comme la causalité d'une relation est une affaire intrinsèque qui ne dépend que des propriétés du processus reliant la cause et l'effet, la relation est causale même dans la situation où la cause préemptée est présente. Donc, la quasi-dépendance suffit pour la causalité, même s'il n'y a pas de dépendance. Mais il existe des cas de préemption où le processus préempté possède lui aussi la quasi-dépendance. Schaffer (2000) montre que la théorie de la quasi-dépendance échoue systématiquement dans des cas de préemption « coupante » (« *trumping preemption* ») où le processus préempté n'est jamais interrompu mais va jusqu'au bout, tout comme le processus préemptant. La théorie de la quasi-dépendance donne le verdict correct dans le cas du processus qui va de l'explosion de la bombe à l'écroulement du pont puisque ce processus n'aboutit pas : quand l'onde de choc arrive, il n'y a plus de pont. (Voir le raisonnement de Hempel rapporté plus haut). Ce processus n'est donc pas intrinsèquement semblable à d'autres processus où aucune surcharge n'est présente et où il y a dépendance contrefactuelle et causalité. Schaffer (2000, p.175) rapporte l'exemple suivant qu'il attribue à Bas van Fraassen et Ned Hall : Le major et le sergent, tous les deux devant le caporal, crient en

Il me semble que la théorie contrefactuelle hérite d'une partie des difficultés de la théorie nomologique parce qu'elle repose tacitement sur les mêmes fondements³¹. (Cela est bien entendu une thèse controversée que j'essaie de justifier par la suite). Suivons Lewis (1973) (et Stalnaker 1968) dans leur conception des conditions de vérité des conditionnels contrefactuels, en termes de mondes possibles. Admettons que *A* est faux dans le monde actuel (ou que l'événement *a* ne s'est pas produit dans le monde actuel : *A* est la proposition que *a* se produit. Cf. Lewis 1986, p. 166). A quelle condition le contrefactuel *A* → *C* (« si *A* était le cas, alors *C* serait le cas » ; équivalent à : « si *a* s'était produit alors *c* se serait produit ») est-il vrai ? Selon Lewis, ce contrefactuel est vrai dans le monde actuel « si et seulement si un monde (accessible³²) où *A* et *C* sont vrais tous les deux est globalement plus semblable au monde actuel que tous les mondes où *A* est vrai mais *C* est faux » (Lewis 1986, p. 41). Il n'y a pas de critères généraux qui permettraient de construire un algorithme universel qui jugerait la similarité de deux mondes possibles en toute généralité. Toutefois il est possible d'établir une hiérarchie de critères, dans l'ordre décroissant de leur importance : ce qui est, selon Lewis, à l'origine de la plus grande dissimilitude entre un monde *w*₀ et un monde *w*₁, c'est le fait que *w*₁ contient des violations des lois de *w*₀ qui sont « larges, répandues et diverses » (Lewis 1986, p. 47). Autrement dit, pour trouver le monde le plus proche parmi ceux où l'antécédent *A* du contrefactuel est vrai, il faut d'abord écarter les mondes qui contiennent des violations majeures des lois du monde

même temps « chargez ! » et le caporal décide de charger. L'ordre de l'officier supérieur « coupe » celui de l'officier inférieur ; la cause de la décision du caporal est donc l'ordre du major, non celle du sergent. Pourtant il est au moins concevable que le processus concurrent de l'ordre du sergent n'est jamais interrompu. Il possède donc autant la quasi-dépendance que le processus causal de l'ordre du major. Lewis (2000, p. 185) reconnaît que ce genre de situations réfute la théorie de la quasi-dépendance.

³¹ Ces problèmes ne se posent toutefois pas exactement de la même manière. Dans un cas de préemption, par exemple, le problème de la théorie nomologique est qu'elle mène au jugement erroné que l'événement préempté est une cause, alors que la théorie contrefactuelle mène au contraire au jugement erroné que la cause réelle n'en est pas une - puisque l'effet aurait eu lieu même en son absence. Cf. Lewis (1986, p. 200).

³² Le concept d'accessibilité d'un monde possible à partir d'un autre permet de construire différents concepts de possibilité et de nécessité. On peut par exemple distinguer la nécessité physique de la nécessité logique, en se fondant sur la distinction entre les mondes possibles où valent les lois de la nature de notre monde actuel et ceux où elles ne valent pas. Les premiers seront considérés comme les mondes accessibles à l'égard de la modalité physique : une proposition est physiquement nécessaire si et seulement si elle est vraie dans tous les mondes possibles accessibles ; une proposition est physiquement possible si et seulement s'il existe un monde accessible dans lequel elle est vraie. Cf. Lewis (1973, p. 4/5).

selon lequel on juge la valeur de vérité du contrefactuel. Le deuxième critère dans l'ordre décroissant de leur importance dit que les mondes les plus semblables sont ceux qui sont en correspondance parfaite (« *perfect match* ») par rapport au monde de référence, à l'égard de tous les faits existant dans la plus grande région spatio-temporelle. Seulement ensuite vient le critère selon lequel la proximité est plus grande s'il n'y a même pas de violations de lois qui sont « petites, localisées, simples » (Lewis 1986, p. 48), c'est-à-dire s'il n'y a pas dans w_1 des événements qui apparaissent comme des « miracles » du point de vue de w_0 . Le poids le moins important revient à la similarité à l'égard de faits particuliers.

C'est à partir de l'application de ces critères au jugement de similarité entre mondes que Lewis espère fonder le jugement selon lequel les conditionnels à rebours (c'est-à-dire des contrefactuels dont l'antécédent porte sur un événement *ultérieur* à l'événement décrit dans le conséquent) ne sont jamais vrais³³. Ce jugement s'avère être la clé de voûte de la défense de la théorie contrefactuelle, dans la mesure où un certain nombre de contre-exemples (et peut-être tous) tirent leur plausibilité (apparente, selon Lewis) d'un emploi au moins tacite de dépendance contrefactuelle à rebours. Le premier contre-exemple qui pose le « problème des effets » ((a) plus haut) s'appuie explicitement sur un contrefactuel à rebours. Dans le contre-exemple de la cause commune ((b) plus haut), le jugement selon lequel si M. Dupont n'avait pas vu le match, alors M. Dupond ne l'aurait pas vu non plus, se fonde sur le contrefactuel à rebours selon lequel si M. Dupont ne l'avait pas vu, la station de télévision ne l'aurait pas transmis³⁴.

Considérons l'exemple suivant³⁵. Dans le monde w_0 , à l'époque de la guerre froide et de la présidence de Nixon aux États-Unis, il existe un bouton rouge qui permet de déclencher de manière absolument

³³ La théorie contrefactuelle pourrait être tentée, comme d'ailleurs la théorie nomologique, d'éviter le problème de devoir prédire à tort que les effets sont aussi des causes de leurs propres effets (le « problème des effets »), en stipulant simplement que la cause précède toujours l'effet dans le temps. Cette solution facile est souvent rejetée pour trois raisons (Cf. Lewis 1986, p. 170) : 1) une telle stipulation ne permettrait pas de résoudre le problème des épiphénomènes ni celui de la préemption. 2) Elle préjuge *a priori* que la causalité « contre la direction du temps » est impossible pour des raisons conceptuelles. L'existence de tels cas de causalité « à rebours » est pourtant une question controversée qui devrait être laissée au jugement scientifique, en dernière instance fondé sur des critères empiriques. Cf. note 19 plus haut. 3) Elle exclut *a priori* la possibilité d'une théorie causale du temps, qui semble pourtant prometteuse.

³⁴ Lewis (1986, p. 200/1) montre qu'au moins certain cas de préemption ne constituent des contre-exemples à la théorie contrefactuelle qu'à condition d'admettre des contrefactuels à rebours.

³⁵ Cf. Lewis (1986, p. 43sq.).

fiable une catastrophe nucléaire. On peut imaginer que w_0 est notre monde actuel. Dans w_0 , Nixon n'appuie pas sur le bouton. Selon Lewis, le contrefactuel « Si N. avait appuyé sur le bouton, alors il y aurait eu une catastrophe nucléaire » est vrai, alors que le contrefactuel à rebours « s'il y avait eu une catastrophe nucléaire, alors N. aurait appuyé sur le bouton » est faux. La raison de la vérité du premier contrefactuel est que le monde le plus proche w_1 où N. appuie sur le bouton est, selon Lewis, un monde exactement comme le nôtre (w_0) jusqu'à l'instant immédiatement précédent l'acte fatal de N. mais où se produit à cet instant-là un « minuscule miracle », c'est-à-dire que « les lois déterministes de w_0 sont violées dans w_1 d'une manière simple, localisé, discrète » (Lewis 1986, p. 44)³⁶. w_1 ne diffère de w_0 que par « quelques neurones supplémentaires [qui] déchargent dans un coin du cerveau de Nixon » (Lewis 1986, p. 44). Parmi les mondes qui satisfont à ces exigences, le plus proche est celui où aucun autre miracle ni violation des lois de la nature n'a lieu. Le conséquent du contrefactuel « en avant » y est donc vrai puisque nos lois déterminent inévitablement la catastrophe nucléaire, une fois appuyé sur le bouton.

Qu'en est-il de la valeur de vérité du contrefactuel à rebours, selon lequel, « s'il y avait eu une catastrophe, N. aurait auparavant appuyé sur le bouton rouge » ? Selon Lewis, ce contrefactuel est faux parce que monde le plus proche de w_0 où l'antécédent est vrai, où il y a donc une catastrophe nucléaire, est un monde où cet holocauste est causé par un petit miracle immédiatement précédent. Lewis montre en effet que les mondes où le miracle se situe après l'holocauste plutôt qu'avant, sont nécessairement plus éloignés du nôtre que ceux où le miracle se situe avant : un miracle qui aurait lieu juste après la catastrophe permettrait de maximiser la similarité de ce monde par rapport au monde actuel, pour ce qui a trait à sa phase finale qui commence après la catastrophe. Mais un tel miracle ne serait pas « petit, localisé et discret », car il faut un très grand miracle pour « effacer » le grand nombre d'effets qui se propagent à partir d'une seule cause³⁷. Cette asymétrie provient de l'existence de processus irréversibles, asymétrie qui est selon une hypothèse plausible également à l'origine de l'asymétrie de la causalité et du temps.³⁸

Contrairement à ce qu'affirme Lewis, l'irréversibilité des processus physiques n'empêche pas la possibilité de raisonner à rebours. L'observation de Lewis selon laquelle les miracles sont toujours plus petits lorsqu'ils se situent dans le passé plutôt que dans l'avenir par rapport à l'antécédent du contrefactuel nous contraint seulement de

³⁶ Sur les miracles, cf. également Lewis (1973, p. 75).

³⁷ "The big miracle required for perfect reconvergence consists of a multitude of little miracles, spread out and diverse" (Lewis 1986, p. 47).

³⁸ Cf. Reichenbach (1956).

chercher le monde le plus proche où un miracle a eu lieu *avant* la catastrophe décrit dans l'antécédent. Mais cela ne détermine pas encore la situation dans le temps de ce petit miracle : Lewis propose de le situer *immédiatement avant* l'événement décrit par l'antécédent, la catastrophe nucléaire, ce qui rendrait effectivement le contrefactuel faux. « La solution correcte aux deux problèmes [celui des effets et celui des épiphénomènes ; M.K.] est, je pense, tout simplement de nier les contrefactuels qui causent les ennuis. Si *e* avait été absent, ce n'est pas parce que *c* aurait été absent. [...] Plutôt, *c* se serait produit tout comme il l'a fait actuellement mais il n'aurait pas causé *e*. » (Lewis 1986, p. 170) Dans notre cas, le verdict lewisien analogue serait : si *e* (la catastrophe) avait été présent, ce n'est pas que *c* (l'événement où N. appuie sur le bouton) aurait été présent, mais plutôt que *c* aurait toujours été absent, et *e* se serait produit grâce à un miracle immédiatement précédent *e*. Mais cette proposition repose sur l'idée d'évaluer le contrefactuel sur la base d'un monde où un miracle survient *dans le laps de temps qui sépare l'événement décrit par l'antécédent de l'événement décrit par le conséquent* du contrefactuel. Or cela détruit la capacité de ce monde-là de fonder notre jugement contrefactuel. La possibilité de raisonner de manière contrefactuelle présuppose que nous considérons que les lois du monde actuel valent sans exception dans le monde où nous évaluons le conditionnel et où l'antécédent est vrai, *pendant le laps de temps qui sépare l'antécédent du conséquent*.

En effet, sur quoi pouvons-nous fonder nos jugements intuitifs de la valeur de vérité des conditionnels contrefactuels, si ce n'est sur la supposition que dans le monde alternatif au nôtre où l'antécédent est vrai, valent les mêmes lois que dans le monde actuel ? Traditionnellement³⁹, on considère que l'un des principaux rôles de la connaissance des lois de la nature est de nous permettre de juger des cours alternatifs que les choses auraient pris dans des circonstances contrefactuelles. Notre capacité de faire ces jugements est notamment indispensable lorsque nous prenons des décisions : nous évaluons les conséquences de chacune de nos actions possibles. Nous ne pouvons estimer ces conséquences qu'à condition de supposer que les lois de la nature ne changent pas après l'acte possible en question. Il me semble donc incontournable de fonder, avec Jackson (1977), l'évaluation des contrefactuels sur l'exigence que le monde possible qui est considéré doit être le plus proche possible du nôtre à l'égard des faits qui existent au temps de l'antécédent et qu'il doit obéir aux mêmes lois que le nôtre pendant le laps de temps qui sépare l'antécédent du conséquent⁴⁰. L'exigence de maintenir les lois de la nature de w_0

³⁹ Cf. Kneale (1949), Goodman (1955), Nagel (1961), Jackson (1977), Armstrong (1983), Zimmermann (2000).

⁴⁰ Cf. Jackson (1977, p. 9).

valides au moins pendant le laps de temps qui sépare l'antécédent du conséquent nous oblige donc de situer le « miracle » avant le conséquent (l'événement où N. appuie sur le bouton). Dans ce cas, il n'y a plus d'obstacle principal au raisonnement contrefactuel à rebours. Le contrefactuel est vrai si le conséquent est vrai dans le monde qui est le plus proche du nôtre à l'égard des faits au temps de l'antécédent (ultérieur dans le temps) et dans lequel les lois qui valent dans le laps de temps qui sépare l'antécédent du conséquent sont les mêmes que dans le monde actuel⁴¹. Même si toutes les lois s'appliquant à une situation ne sont pas réversibles, il suffit, pour la vérité du contrefactuel à rebours, que dans les circonstances, l'événement conséquent *c* soit une condition nécessaire de l'événement antécédent *e*. Le conditionnel à rebours sur N. est vrai dans la mesure où, dans les circonstances, d'autres causes possibles qui auraient pu déclencher la catastrophe sans l'intervention de miracles se situent dans des mondes possibles plus éloignés du nôtre que le monde où N. déclenche la catastrophe en appuyant sur le bouton. Dans la mesure où notre raisonnement réfute la thèse de Lewis selon laquelle aucun contrefactuel à rebours n'est vrai, la défense de la théorie contrefactuelle de la causalité menacée par les contre-exemples mentionnés plus haut, s'écroule. La possibilité que les contrefactuels à rebours soient vrais laisse la théorie contrefactuelle sans défense face aux problèmes des effets et des épiphénomènes.

Jusqu'ici, j'ai accepté de raisonner en termes de mondes « miraculeux ». Comme Lewis tient à le préciser (Lewis 1986, p. 44/5), les mondes miraculeux ne sont pas inconsistants au sens où ils contiendraient des violations de leurs propres lois. Ils apparaissent seulement comme miraculeux de notre point de vue : leurs lois (celles de w_1 , obéies strictement dans w_1) diffèrent de manière à permettre que le cours des choses diffère du cours actuel seulement dans une minuscule portion de l'espace-temps, mais nulle part ailleurs⁴². Or, la conception d'un tel monde semble renfermer une contradiction. Une loi est par définition une contrainte qui détermine une régularité globale⁴³ ; par conséquent, la moindre divergence entre deux mondes possibles à l'égard d'une loi de la nature entraîne une vaste divergence à l'égard des faits et événements particuliers, à travers tout le passé et l'avenir. Par ailleurs, un changement par rapport à un fait particulier, si minuscule soit-il, comme les quelques neurones qui déchargent dans le cerveau de N., requerrait la violation de

⁴¹ Cf. Jackson (1977 p. 11/12).

⁴² Dans l'exemple du contrefactuel où N. appuie sur le bouton, le monde possible qui sert de fondement à l'évaluation de ce contrefactuel est un monde qui ne diffère du nôtre qu'à l'égard du fait que « quelques neurones supplémentaires déchargent dans un coin du cerveau de Nixon » (Lewis 1986, p. 44).

⁴³ Selon les théories régularistes des lois, une loi *est* simplement une régularité globale.

plusieurs lois, à commencer par les lois de conservation. Lewis se trouve devant le dilemme suivant. Ou bien les mondes qui sont candidats pour être la base d'évaluation d'un contrefactuel parce qu'ils sont proches du nôtre, sont des mondes cohérents au sens où leurs lois valent de manière stricte. Alors ces lois doivent être les mêmes que les nôtres, car autrement la divergence seraient nécessairement vaste à l'égard des faits, et nous n'aurions aucun moyen d'y porter des jugements, car nous ne pouvons pas raisonner à partir de lois différentes. Ou bien ce sont des mondes vraiment miraculeux, mais un monde contradictoire dont ses propres lois ne sont pas des lois (puisqu'elles sont violées), ne peut pas servir de base pour l'évaluation des contrefactuels.

Afin d'éviter la considération de mondes possibles miraculeux, j'ai suggéré⁴⁴ que l'évaluation des contrefactuels se fasse à partir de la contemplation de mondes possibles qui partagent strictement nos lois de la nature, mais qui ne contiennent aucune contrepartie d'une partie de l'histoire du monde actuel. Dans l'évaluation d'un contrefactuel « en avant » où l'antécédent a lieu à t , on considère un monde possible qui contient une contrepartie de la partie temporelle du monde actuel qui s'étend de t à l'infini dans le futur, mais qui ne contient aucune contrepartie du passé du monde actuel.

CONCEPTION INTERVENTIONNISTE

Les défenseurs de la conception interventionniste⁴⁵ proposent de fonder l'asymétrie de la causalité qui pose des difficultés formidables à la fois à la théorie nomologique et à l'approche contrefactuelle, sur l'asymétrie entre le passé que nous ne pouvons pas influencer et l'avenir qui se présente à nous en tant qu'agents comme un « espace » ouvert de possibilités » (Keil 2000, p. 440). Si A est la chute d'une avalanche, et B et C les écroulements de deux maisons causés par A, qu'est-ce qui permet de justifier le rejet des contrefactuels problématiques suivants ? « Si B n'avait pas eu lieu, alors A n'aurait pas eu lieu » est un contrefactuel explicitement à rebours qui pose le problème des effets. « Si B n'avait pas eu lieu, alors C n'aurait pas eu lieu » qui pose le problème des épiphénomènes, repose lui aussi implicitement sur une dépendance contrefactuelle à rebours. L'interventionniste propose de juger leur valeur de vérité par rapport au monde qui diffère du nôtre seulement par une intervention humaine qui joue le rôle du miracle lewisien. Les mondes que nous comparons dans notre évaluation des contrefactuels partagent nos lois de la nature, mais diffèrent à l'égard d'une action qu'un agent choisit librement. Or l'intervention minimale qui permet d'empêcher B

⁴⁴ Cf. Kistler (2002 ; 2002b ; et à paraître).

⁴⁵ Gasking (1955), von Wright (1971), Cartwright (1983), Price (1992), Menzies et Price (1993), Keil (2000).

consiste à protéger directement la maison B. Mais si cela avait été le cas, A, ainsi que C auraient toujours eu lieu, ce qui fonde le jugement selon lequel les deux contrefactuels litigieux car à rebours (au moins implicitement, au cas du second), sont faux⁴⁶.

Plus généralement, la théorie interventionniste, ou théorie de la manipulabilité⁴⁷, réduit l'existence d'une relation causale entre les événements distincts A et B au fait que « provoquer A serait un moyen efficace par lequel un agent libre pourrait accomplir B » (Menzies et Price 1993, p. 187)⁴⁸. Cette proposition se heurte à des difficultés dont deux sont majeures⁴⁹. En premier lieu, elle est anthropocentrique en postulant une dépendance conceptuelle entre la causalité et la capacité humaine d'agir. Elle ne se contente pas de constater que nos actions sont des causes, mais affirme en outre qu'en toute généralité, un événement *c* est cause de *e* si et seulement si, un agent humain pourrait obtenir *e*, en accomplissant *c*. Cette approche est donc condamnée à considérer toute relation causale qui a lieu hors de l'atteinte possible d'une action, comme étant causale seulement par analogie. Il est en principe impossible pour un agent humain de déclencher la décomposition radioactive d'un noyau atomique. L'analyse interventionniste ne permet donc pas de la considérer comme un processus causal. Ensuite, toute décision de considérer telle relation comme causale ou tel processus comme causal car analogue, ou suffisamment semblable, à une action, et tel autre comme non causal car insuffisamment analogue, souffre d'une indétermination qui prive la théorie d'un fondement permettant de porter des jugements causaux de manière non arbitraire, en dehors du domaine des actions humaines.

La deuxième difficulté vient de sa circularité : l'approche interventionniste considère la capacité d'agir comme primitive, pour analyser le concept de causalité par son intermédiaire. Mais lorsque nous agissons, nous causons des événements. La compréhension du concept d'action présuppose donc une compréhension préalable du concept de causalité. La circularité peut être évitée si l'on pose explicitement le concept d'action ou d'intervention comme primitif, et non susceptible d'être analysé à son tour, et surtout non en termes cau-

⁴⁶ C'est le raisonnement que propose Keil (2000, p. 442).

⁴⁷ Menzies et Price l'appellent "agency theory" (Menzies and Price 1993, p. 187).

⁴⁸ "Bringing about the occurrence of A would be an effective means by which a free agent could bring about the occurrence of B."

⁴⁹ Hausman mentionne trois difficultés : « La théorie de la manipulabilité est étroitement circulaire, anthropomorphique dans une mesure injustifiable ; et elle rend des inférences causales sur la base de l'observation passive problématiques sans que cela soit nécessaire. » (Hausman 1986, p. 145)

saux⁵⁰. Mais alors il semblerait que nous n'ayons pas d'analyse de la causalité en tant que telle du tout, mais seulement l'affirmation que les interactions causales d'une certaine espèce, à savoir les actions humaines, sont paradigmatiques et peuvent servir de modèle pour le genre entier. En l'absence d'une analyse directe de ce qui caractérise le genre, à savoir la causalité elle-même, l'analyse devra procéder dans l'ignorance de ce qui, parmi les propriétés des actions humaines, se laisse généraliser à tous les cas de causalité.

Ces critiques ne remettent pas en cause le fait, surtout mis en évidence par von Wright (1971), que sur le plan épistémique, nous ne pouvons connaître les relations causales objectives que par l'intermédiaire de nos interventions expérimentales⁵¹. Notre capacité d'intervenir dans le cours naturel des choses est le seul levier sur lequel peut s'appuyer notre connaissance des relations causales. Comme le dit von Wright, « il est établi qu'il existe une connexion causale entre p et q si nous nous sommes convaincus que, en manipulant l'un de ces facteurs, nous pouvons faire en sorte que l'autre est ou non présent. D'habitude, nous parvenons à cette conviction en faisant des expériences » (von Wright 1971, p. 72)⁵². Reconnaître ce lien épistémique important ne devrait pourtant pas obscurcir le fait qu'une relation causale naturelle quelconque (qui n'est pas une action humaine) ne dépend pas, pour son existence objective, d'une action humaine, ni même d'une action possible.

CONCEPTION PROBABILISTE

La thèse fondamentale de l'approche probabiliste de la causalité est que A est la cause de B si et seulement si l'occurrence de A augmente la probabilité de l'occurrence de B. L'origine de cette théorie est double. D'une part, elle peut être considérée comme une version affaiblie de la théorie déductive-nomologique, en substituant des lois probabilistes ou statistiques aux lois déterministes de l'*explanans*⁵³. Mais l'on peut aussi y parvenir à partir de la théorie de la décision

⁵⁰ La dépendance du concept d'intervention par rapport à celui de causalité rend la théorie interventionniste de la causalité clairement inutilisable dans le cadre de la théorie causale de l'action, selon laquelle les actions se distinguent d'autres mouvements corporels par leur origine causale dans un épisode mental approprié. Il est ouvertement circulaire d'analyser le concept d'action en partant du concept de causalité, et d'analyser en même temps la causalité à partir de l'action. Cf. Keil (2000, chap. III, 3).

⁵¹ Ces méthodes d'investigation des relations causales font l'objet d'une célèbre classification par Mill (1843).

⁵² "It is established that there is a causal connection between p and q when we have satisfied ourselves that, by manipulating the one factor, we can achieve or bring it about that the other is, or is not, there. We usually satisfy ourselves as to this by making experiments."

⁵³ C'est de cette manière que la théorie est introduite par Suppes (1970).

rationnelle⁵⁴. Il est rationnel de faire A dans le but d'obtenir B si et seulement si l'on croit que A cause B. Selon la théorie normative de la décision, il est rationnel de faire A pour obtenir B si et seulement si on croit que $P(B|A) > P(B|\neg A)$. Si ces équivalences sont correctes, on peut conclure que A cause B si et seulement si $P(B|A) > P(B|\neg A)$ ⁵⁵. Le problème majeur de cette approche concerne son application à l'analyse de la causalité singulière, c'est-à-dire aux relations causales entre des paires particulières d'événements. Pour reprendre l'un des exemples préférés des débats sur la causalité probabiliste, supposons que la probabilité de développer un cancer des poumons est plus grande parmi les fumeurs que parmi les non-fumeurs, au sens où la fréquence des cancers est plus grande dans la population des fumeurs que dans celle des non-fumeurs. La théorie probabiliste de la causalité en conclut que fumer est une cause du cancer. Or il n'est pas clair si la dépendance statistique entre les deux populations permet de fonder un jugement dans un cas singulier⁵⁶ : il semble parfaitement possible que l'acte de fumer de Jean ne cause pas chez lui le développement d'un cancer - soit parce qu'il n'en développe pas du tout, soit parce que son cancer est (causalement) dû à un autre facteur, par exemple à l'inhalation de particules d'amiante. L'augmentation de la probabilité de l'occurrence d'un événement de type B, en conséquence de l'occurrence d'un événement de type A, n'est donc pas suffisant pour que l'occurrence de l'événement a (de type A) soit la cause d'un événement b (de type B) qui a pourtant bien eu lieu. L'augmentation de la probabilité de B par A n'est pas non plus nécessaire, car l'événement a (de type A) peut très bien causer b (de type B) alors que, en général, l'occurrence d'un événement de type A *baisse* la probabilité de l'occurrence d'un événement de type B. C'est

⁵⁴ A ce titre, on peut la considérer comme une variante de la théorie de la manipulabilité, considérée dans la section précédente.

⁵⁵ Cet argument est présenté comme une motivation majeure de la théorie probabiliste par Papineau (1989, p. 321 et 328) et Mellor (1995, chap. 7). Il n'est certainement pas concluant dans la mesure où il ne permet de conclure de manière valide qu'à l'équivalence des *croiances* : il montre qu'il est équivalent de *croire* que A cause B et de *croire* que $P(B|A) > P(B|\neg A)$. Il est évidemment illicite d'en conclure, sans autre argument, à l'équivalence des *objets* de ces croiances.

⁵⁶ Ellery Eells, l'un des majeurs défenseurs de la théorie probabiliste, le reconnaît très clairement: "Given the conceptual independence of token-level causal facts from type-level causal facts, it should not be surprising that what is true at the type level is [...] conceptually independent of what is true of the token level, and that token-level causation cannot be straightforwardly understood in terms of type-level causal relations." (Eells 1991, p. 16)

le cas dans la situation suivante⁵⁷ : Moriarty, Watson, et un rocher instable se trouvent au bord d'une falaise. Holmes se trouve au pied de la falaise, sans se douter de rien. Watson doit partir, et il sait que s'il ne fait rien avant de partir, Moriarty poussera le rocher par-dessus le bord de la falaise, avec l'intention d'écraser Holmes. La seule chose que Watson peut faire pour tenter de sauver Holmes, c'est d'essayer de pousser le rocher lui-même, mais si fort qu'il passe par-dessus Holmes. Watson réussit à atteindre le rocher et il le pousse aussi fort qu'il peut, afin d'essayer de sauver Holmes. Malheureusement l'improbable se produit, la tentative échoue et Holmes est écrasé. La cause de la mort de Holmes est un événement d'un type qui diminue la probabilité de sa mort. L'événement de pousser le rocher « avait une tendance à empêcher » la mort de Holmes, mais l'a néanmoins causé.⁵⁸

Deux autres considérations montrent que la théorie probabiliste est mal appropriée à l'analyse de la relation causale entre événements concrets. Les termes des relations de probabilité sont des facteurs, des types d'événements ou propriétés d'événements, comme le fait de fumer ou le fait de développer un cancer. Ce sont des types d'entités qui ne sont situés ni dans l'espace ni dans le temps. Ces entités ne peuvent donc pas être causalement reliées au sens strict, dans la mesure où la relation causale a lieu entre des termes situés dans l'espace et le temps, généralement conçus comme événements. Cela suggère que la théorie probabiliste soit en réalité une théorie non pas de la causalité, mais de l'explication ou du lien nomique entre propriétés, autrement dit des lois de la nature. Dans le cadre du modèle DN de l'explication, elle peut jouer les deux rôles à la fois.

La deuxième raison de penser que la théorie probabiliste est inadéquate pour rendre compte de la causalité entre événements spatio-temporels est la suivante. Les corrélations statistiques qui constituent le point de départ de l'analyse probabiliste sont souvent factices (*spurious*), dans le sens où la corrélation a lieu entre deux (types d') effets d'une cause (d'un type de cause) commune, sans que ces effets soient liés comme cause à effet entre eux. La théorie doit donc exiger que l'augmentation de la probabilité de B par A ne permet de conclure à la présence d'une relation causale entre (les événements de type) A et B que s'il n'existe pas de facteur tiers qui s'interpose en tant que cause réelle des deux. On appelle un facteur de ce genre un « facteur écran » (*screening factor*). La corrélation entre les chutes de baromètre et les tempêtes est factice ; la théorie probabiliste, enrichie de l'exigence de l'absence de facteurs écran, prédit qu'elle n'est

⁵⁷ Cet exemple est dû à Good (1961-2) ; On peut trouver d'autres exemples dans Eells and Sober (1983) ; Eells (1991, p. 281sq.) ; Heidelberger (1992, p. 151) ; Edgington (1997, p. 420).

⁵⁸ Cf. Good (1983, p. 216-7).

pas causale car la chute de la pression de l'air constitue un facteur écran : à l'intérieur d'un ensemble de situations où la pression de l'air est constante, la corrélation entre les chutes de baromètre et les tempêtes disparaît. Mais cette stratégie ne peut pas être généralisée. Tant que les ensembles à l'intérieur desquels on évalue les probabilités conditionnelles contiennent plus d'un élément, il est toujours possible de partitionner cet ensemble de manière à inverser les probabilités conditionnelles, alors que bien entendu ces changements dans la manière de calculer la probabilité n'affectent pas la réalité des relations causales. Ce fait est connu comme le « paradoxe de Simpson »⁵⁹. Le célèbre cas de la prétendue discrimination sexiste de l'université de Berkeley dans sa politique des admissions peut servir d'illustration. Cartwright (1983, p. 37) cite une étude qui montre que la fréquence d'admission à Berkeley parmi l'ensemble des candidats masculins est supérieure à la fréquence d'admission parmi les candidats féminins. Mais l'existence de cette corrélation statistique ne permet pas de conclure à l'existence d'une relation « causale »⁶⁰ entre le fait d'être un homme et le fait d'être favorisé dans la sélection pour l'admission à Berkeley. Car il existe une partition de l'ensemble des candidats qui, lorsqu'on évalue les probabilités conditionnelles pertinentes séparément à l'intérieur des ensembles de cette partition, inverse la corrélation. Il s'agit de la partition par rapport aux différents départements qui effectuent la sélection des candidats. Il s'est avéré que la corrélation entre le sexe et l'admission des candidats n'est pas fondée sur l'existence d'une relation « causale » entre ces deux facteurs, mais provient d'un facteur « causal » commun : les hommes ont tendance à présenter leur candidature dans les départements où il est plus facile d'être admis, alors que les femmes se présentent plus souvent dans les départements plus sélectifs. Pour simplifier, considérons un cas où il n'y a que deux départements et où les fréquences pertinentes sont les suivantes⁶¹ :

	Département 1	Département 2	Total
Masculin	81/90	2/10	83/100
Féminin	9/10	18/90	27/100

Dans ce cas précis, lorsque les probabilités d'être admis sont évaluées globalement - sans distinguer entre les deux départements - la probabilité d'être admis est plus que trois fois plus grande pour les hommes que pour les femmes. En revanche, il n'y a aucune corréla-

⁵⁹ D'après Simpson (1951). Cf. Eells (1991, chap. 2).

⁶⁰ Comme je l'ai déjà expliqué, il me semble que l'on commette une erreur de catégorie lors qu'on appelle « causes » des entités abstraites, comme des propriétés ou facteurs. Il serait plus correct de parler de dépendance nomique. Celle-ci semble être le domaine où la théorie probabiliste peut s'appliquer fructueusement.

⁶¹ Le détail numérique a été élaboré par Eells (1991, p. 63).

tion entre le sexe des candidats et leur admission lorsque les probabilités sont évaluées à l'intérieur des sous-ensembles de candidatures que l'on obtient par la partition de l'ensemble en deux départements. Le premier département accepte 90% des hommes mais aussi 90% des femmes, alors que le second département n'accepte que 20% des hommes mais aussi 20% des femmes. Eells (1991, p. 66sq.) montre que l'existence d'une corrélation globalement positive entre le fait d'être masculin et le fait d'être admis est non seulement compatible avec l'absence d'une telle corrélation, lorsque les probabilités pertinentes sont évaluées à partir d'une partition appropriée, mais aussi avec l'existence d'une corrélation *négative* (c'est-à-dire le cas où les départements admettent les femmes plus facilement que les hommes) dans le cas d'une évaluation par département. Il est aussi possible que l'évaluation à partir des départements montre bien qu'il existe une discrimination contre les femmes mais qu'elle est moindre que celle qui apparaît à l'évaluation globale⁶². Bref, on ne peut tirer aucune information définitive de l'analyse des probabilités effectuées à partir d'un ensemble inhomogène. Tant qu'il est possible d'introduire des partitions à partir de nouveaux facteurs, il est toujours possible que les probabilités conditionnelles s'inversent. Or le fait que les jugements de corrélation statistique restent toujours relatifs à la partition choisie, et que le choix d'une nouvelle partition risque toujours de les inverser, les rend inappropriés à fonder les jugements causaux.

La seule solution semble être d'exiger, comme le fait Cartwright (1983), que les ensembles à partir desquels l'évaluation des probabilités est effectuée soient absolument homogènes, c'est-à-dire qu'*aucun* facteur causalement pertinent ne soit laissé en dehors de l'analyse⁶³. Or, cette solution tend à éliminer l'élément probabiliste de la théorie, pour la rendre équivalente à la théorie nomologique. Pour le cas des relations causales déterministes, la relation de probabilité conditionnelle de B, étant donné A, se réduit à la relation nomique selon laquelle B dépend nomiquement de A. Comme nous l'avons vu, pour évaluer correctement la dépendance de B par rapport à A, il faut exclure toutes les situations où il existe une corrélation entre A et B sans que A soit causalement relié à B, et les situations où la corrélation entre A et B est renforcée ou diminuée par d'autres facteurs qui influent sur B, sans être causés par A. Pour juger si A cause B, on évalue donc la probabilité de B, étant donné A, dans toutes les

⁶² Cf. Eells (1991, p. 68).

⁶³ Selon Cartwright, il faut évaluer les probabilités conditionnelles « en tenant invariables les autres causes » (Cartwright 1983, p. 28), et il n'est possible de conclure à la présence d'une relation causale, à partir de l'augmentation de probabilité, que lorsque l'ensemble des situations est « homogène à l'égard de tous les autres facteurs causaux » (Cartwright 1983, p. 25). Mais lorsqu'il s'agit d'événements concrets, il y a toujours d'autres facteurs causaux, et aucune paire de situations n'est homogène à l'égard de tous les facteurs causaux.

populations qui sont de type A et B mais parfaitement homogènes à l'égard de tous les facteurs potentiellement interférents. A ce moment-là, l'élément probabiliste disparaît : dans un monde déterministe, si tous les facteurs extérieurs potentiellement interférents sont tenus constants (supposons qu'ils soient systématiquement absents), alors la probabilité de Y, étant donné X est 1 si X cause Y, et 0 si X ne cause pas Y. Les valeurs intermédiaires ont leur source dans la variabilité des facteurs extérieurs interférents, variabilité qui a ici été exclue. Ce raisonnement remet en cause l'idée fondamentale de la théorie probabiliste de la causalité. Les probabilités intermédiaires ne mesurent pas l'impact causal de X sur Y (qui ne peut être qu'existant ou non existant) mais la probabilité des facteurs interférents. Mais celle-ci est indépendante de la relation entre X et Y elle-même. En dehors des processus irréductiblement indéterministes, le contenu des attributions de probabilités intermédiaires ne porte donc pas sur la relation causale elle-même, mais sur notre ignorance des facteurs extérieurs. Par conséquent, la théorie probabiliste ne fournit pas une analyse originale de la relation causale elle-même : en ce qui concerne la causalité objective, la théorie probabiliste se réduit à la théorie nomologique.

La théorie probabiliste tire sa plausibilité de l'analogie entre l'évaluation de probabilités conditionnelles et l'évaluation d'une suite d'expériences, suivant les méthodes de Mill. Pratiquement, on ne peut pas exclure tous les facteurs interférents, et la probabilité de B étant donné une expérience où A est réalisée et où les facteurs interférents ont été éliminés selon les possibilités pratiques, sera intermédiaire. Mais cela n'a pas sa source dans une caractéristique de la relation entre A et B elle-même.

Eells (1991, chap. 6) propose une théorie probabiliste des relations causales particulières qui cherche à réduire ce que j'appellerai plus loin la « responsabilité causale » d'un fait à l'égard d'un autre, à une certaine évolution de la probabilité du fait « effet » au cours du temps, à partir du temps correspondant au fait « cause », en fonction de ce fait « cause » ainsi que de facteurs extérieurs interférents. Il me semble cependant que la théorie probabiliste des relations causales particulières se heurte aux mêmes difficultés que la théorie probabiliste de la dépendance causale entre facteurs.

Dans le cas des relations causales particulières, la probabilité de l'effet étant donné la cause est la limite des fréquences de l'effet dans une suite infinie de populations où la cause est présente. Cette suite contient un nombre infini de répliques de la situation, dans son ensemble, avec tous les facteurs interférents.⁶⁴ Mais dans un univers déterministe, dans la mesure où les situations dans cette suite ne diffèrent pas du tout en ce qui concerne tous les facteurs qui affectent

⁶⁴ Cf. Eells (1991, chap. 6).

directement ou indirectement l'effet, si le fait que x est X est causalement responsable du fait qu' y est Y , alors la probabilité qu' y est Y , étant donné que x est X , est 1. Dans ce cas, il n'y a plus d'analyse probabiliste de la relation de responsabilité causale. Il me semble que la théorie d'Eells bute là sur le problème qui fait l'objet du deuxième argument russellien contre la théorie de la causalité en termes de régularité (examiné plus haut). L'attribution de nombres inférieurs à 1 ne se justifie que parce qu'on ignore une partie du contexte.

THÉORIES DE LA CAUSALITÉ EN TERMES DE TRANSMISSION

La conclusion générale qui s'impose après l'examen des alternatives à la théorie déductive-nomologique examinées jusqu'ici, est que deux d'entre elles, la théorie contrefactuelle et la théorie probabiliste sont mieux appropriées à l'analyse de la dépendance nomique qu'à celle de la causalité. Cela semble particulièrement clair dans le cas de la théorie probabiliste : son objet est l'étude systématique des corrélations entre différentes propriétés (ou entre différents facteurs, ou entre les ensembles qui constituent leurs extensions). Ces corrélations peuvent être révélatrices de l'existence de dépendances nomiques qui existent entre ces propriétés⁶⁵. En revanche, elles ne donnent pas d'information définitive sur l'existence d'une relation causale entre les événements qui exemplifient ces propriétés. Comme nous l'avons vu lors de l'examen critique de la conception déductive-nomologique, la dépendance nomique n'est pas équivalente à la causalité. Deux propriétés peuvent être nomiquement liées sans que deux événements qui exemplifient ces propriétés soient liés comme cause et effet. Mais j'ai aussi essayé de montrer (ce qui est probablement encore plus controversé) que la théorie de la causalité en termes de dépendance contrefactuelle se laisse également plus facilement interpréter comme une théorie de la dépendance nomique que comme une théorie de la causalité : l'antécédent d'un contrefactuel est une proposition qui affirme l'exemplification d'une propriété (exprimé par le prédicat de cette proposition). Afin de déterminer si la proposition exprimée par le conséquent est vraie dans le monde le plus proche dans lequel l'antécédent est vrai, ce que l'on fait est d'examiner s'il y a une dépendance nomique entre la propriété désignée par le prédicat de l'antécédent et la propriété désignée par le prédicat du conséquent. On ne peut répondre à cette question qu'en contemplant des mondes possibles qui partagent nos lois de la nature.

Le fait que l'objet propre de ces théories soit la dépendance nomique entre propriétés plutôt que la relation causale entre événements concrets est la raison profonde de leur échec face aux contre-exemples où A est nomiquement lié à B , mais où un événement a

⁶⁵ Sur les conceptions contemporaines les plus importantes à l'égard des lois de la nature, cf. Kistler (1999a, chap. 2) et Barberousse, Kistler, Ludwig (2000, chap. 4).

exemplifiant A n'est pourtant pas la cause de b exemplifiant B : soit parce que b est cause de a (problème des effets), soit parce que a et b sont les effets d'une cause commune, soit parce que b a été causé par une cause préemptante c. La clé de la solution à tous ces problèmes réside donc dans la recherche d'une théorie qui ne se situe plus sur le plan des propriétés mais sur celui des événements spatio-temporels qui sont les termes des relations causales particulières⁶⁶.

L'approche qui me semble être la plus prometteuse, fait l'hypothèse suivante : ce qui fait qu'un événement *a* est la (ou une) cause de l'événement *b*, c'est le fait que quelque chose soit transmis, ou transféré, entre ces événements. Il a d'abord été proposé⁶⁷ que le propre des processus causaux est d'avoir la capacité de transmettre un caractère, où un caractère est « le résultat d'une intervention par l'intermédiaire d'un processus irréversible » (Reichenbach 1956, p. 198)⁶⁸. Un cas exemplaire de l'introduction d'un caractère qui se propage ensuite tout le long d'un processus causal est le passage d'un rayon de lumière à travers un filtre rouge. La couleur est transmise par le rayon aussi longtemps que ce processus ne fait pas l'objet d'une nouvelle interaction. En cela les processus causaux s'opposent aux pseudo-processus dont Salmon (1984, p. 141/2) décrit un exemple très clair : un phare tourne au centre d'un cylindre creux et projette un point lumineux blanc sur sa paroi interne. La ligne de monde constituée par la séquence de portions illuminées de la paroi constitue un pseudo-processus, et non un processus causal, dans la mesure où, si l'on pose un filtre rouge à l'un des endroits parcourus par le point lumineux, il sera marqué de rouge, mais cette marque n'est pas transmise à la suite du pseudo-processus qui consiste en lumière blanche comme avant.

Cette version d'une théorie de la causalité en termes de transfert rencontre une objection à mon avis décisive⁶⁹. Le problème est que les définitions de la « capacité à transmettre un caractère » et celle de « processus causal » forment un cercle vicieux⁷⁰ : selon la thèse principale de la théorie, un processus causal est un processus capable de transmettre un caractère. Mais la notion de caractère ne peut à son tour être définie qu'à l'aide du concept de processus causal. Car il faut exiger que seules les interactions *causales* soient capables d'introduire un caractère dans un processus. Un caractère est une altération de propriété due à une interaction locale singulière. Il faut

⁶⁶ C'est aussi le constat que dresse Ehring (1997).

⁶⁷ Cf. Reichenbach (1956, p. 198sq.), Salmon (1984).

⁶⁸ La similarité avec la théorie de la manipulabilité n'est qu'apparente. Les interactions de la théorie du transfert ne sont pas nécessairement le fruit d'interventions humaines.

⁶⁹ C'est aussi l'avis du principal défenseur de la théorie, Wesley Salmon, qui l'a abandonnée pour cette raison. Cf. Salmon (1994).

⁷⁰ Cf. Dowe (1992, p. 200/1), Kistler (1999a, p. 89).

exiger que cette interaction soit causale car autrement une intersection de deux pseudo-processus pourrait introduire un caractère. (Cela contredirait l'idée d'un pseudo-processus.)

Pour sortir de ce cercle, Dowe (1992) a proposé que deux événements soient liés comme la cause et l'effet si et seulement s'ils manifestent la même quantité d'une grandeur conservée (désormais abrégée « GC »). La théorie de Dowe est explicitement opposée à l'idée selon laquelle ces quantités de GC peuvent être transmises, et ne peut donc pas être considérée comme une variante de la théorie de transfert. Mais il semble naturel d'utiliser le concept d'une quantité de grandeur conservée pour généraliser une variante traditionnelle de la théorie du transfert selon laquelle la relation causale consiste en une transmission d'énergie⁷¹. Dans le contexte de la conception contemporaine des grandeurs conservées fondées sur des symétries, il semble arbitraire de fonder la causalité sur la seule transmission d'énergie, ou sur toute autre liste de grandeurs, qui inclut la quantité de mouvement (Fair 1979), la force (Aronson 1971b), ou l'information (Krajewski 1982). L'idée de généraliser aux quantités de n'importe quelle grandeur conservée comme support « matériel » des relations causales s'impose aussi à partir d'une réflexion sur le concept de transfert lui-même. Ne peut être transféré que ce qui se conserve au cours du transfert. J'ai donc proposé de réduire la relation causale entre deux événements particuliers à l'existence d'une transmission entre eux d'une quantité de GC⁷². C'est une tâche scientifique de dresser la liste explicite de toutes les GC qui existent dans la nature et qui peuvent donc être porteurs de relations causales.

Cette conception peut être élaborée de façon à rendre compte de tous les cas qui posent des difficultés insurmontables aux théories examinées auparavant. La première qualification concerne l'asymétrie de la causalité. Le concept de transfert est symétrique : la quantité individuelle de GC X est transférée de a à b si elle est présente à la fois dans l'événement a et dans l'événement b . Dans un monde où tous les processus sont réversibles, il y aurait de la transmission et donc de la causalité, mais les relations causales (et avec eux, le temps, si l'on réduit la direction du temps à celle de la causalité) y seraient symétriques. Il n'y aurait pas de différence intrinsèque quant à son « efficacité » entre les deux termes d'une relation causale. En revanche, dans le monde actuel, la plupart des processus réels sont irréversibles. Cette irréversibilité est à la fois à l'origine de l'asymé-

⁷¹ Cette théorie a été défendue depuis la découverte de la loi de conservation de l'énergie au 19^e siècle, à commencer par Julius Robert Mayer, puis par les défenseurs de l'énergétisme, en particulier W. Ostwald. Cf. Krajewski (1997). Elle a été reprise plus récemment par Konrad Lorenz (1941), puis dans la littérature anglophone par Aronson (1971a; 1971b) et Fair (1979).

⁷² Cf. Kistler (1998), Kistler (1999a, chap. 1).

trie de la causalité et de celle du temps. En empruntant à Reichenbach (1956) le concept d'un réseau causal où tous les événements sont causalement liés à d'autres, de sorte qu'aucun événement reste absolument isolé, il est possible de considérer que dès lors qu'une écrasante majorité des processus dans ce réseau possède la même direction intrinsèque, qui est la direction de l'augmentation de l'entropie, cette direction est celle de l'ensemble du réseau. Ainsi il est possible d'attribuer une direction même aux processus intrinsèquement symétriques, c'est-à-dire aux processus réversibles : cette attribution se fait indirectement, par l'intermédiaire du réseau dans son ensemble.

C'est donc en ayant recours à la propriété empirique de la majorité des processus causaux réels d'avoir une direction intrinsèque que la théorie du transfert peut résoudre le problème des effets : *a* est la cause de *b* parce que premièrement, il y a transmission d'une quantité individuelle d'une GC entre eux, et deuxièmement, ou bien le processus de transmission entre *a* et *b* est physiquement asymétrique et irréversible, ou bien ce processus est réversible, mais fait partie d'un réseau de processus causaux qui possède une direction globale et impose cette direction au processus de transmission entre *a* et *b*.

Le problème des épiphénomènes, le problème de la préemption et de la surdétermination causale ne sont pas des problèmes pour la théorie du transfert. Il n'y a pas de transmission de quoi que ce soit entre M. Dupond et M. Dupont qui regardent toujours les mêmes émissions de télévision et qui sont donc en corrélation parfaite sur ce plan-là. L'exemple est particulièrement clair à cet égard parce que dans cette situation, la théorie de la relativité restreinte interdit la possibilité de transmission⁷³. On a donc une explication simple de l'absence de relation causale. Dans le cas de préemption considéré plus haut, c'est la surcharge qui a transmis de l'énergie sur le pont, énergie transformée en énergie de déformation au cours de l'écrasement. L'explosion de la bombe a bien transmis de l'énergie, mais non sur le pont puisque celui-ci n'existe plus lorsque l'onde de choc qui transporte cette énergie, arrive sur les lieux.

Mais la théorie du transfert rencontre ses propres problèmes. Sosa et Tooley ne sont pas les seuls à juger que « les perspectives pour une réduction physicaliste de la causalité ne paraissent pas très bonnes » (Sosa et Tooley, 1993, p. 4), pour trois raisons dont les deux premières ne concernent à strictement parler que la version de la théorie proposée par Fair (1979), mais ne posent pas de difficultés à la présente version. Cependant, le troisième argument de Sosa et Tooley exige une prise de position quant au statut conceptuel de la

⁷³ On dit que la distance entre les événements de réception est « spatiale » (*space-like*), alors que seuls les événements séparés par des distances « temporelles » (*time-like*) peuvent être reliés par des processus causaux, en particulier par des signaux.

théorie. Leur première objection à la théorie du transfert est que la relation causale ne peut être réduite à un type de relation physique que si celle-ci possède une direction qui n'a pas à son tour un fondement causal. En effet, le concept de transfert, dans son acception commune, ne possède une telle direction que parce qu'il a un sens causal, et sans doute même anthropomorphique, son application paradigmatique étant aux *actions* de transfert (l'action de transférer un objet d'un endroit à un autre). Afin d'éviter à la fois la circularité et l'anthropomorphisme, Fair lui-même propose de fonder l'asymétrie sur l'asymétrie temporelle : de deux événements liés par transfert, la cause est simplement celui qui précède. Mais nous avons déjà remarqué premièrement que cette stratégie rend impossible la réduction de la direction du temps elle-même, et deuxièmement qu'elle stipule pour des raisons *a priori* qu'il n'existe pas de causalité rétroactive, alors que cela semble être une question empirique. Nous avons déjà indiqué la solution à cette difficulté : notre proposition consiste à concevoir le concept de transfert d'une quantité d'une GC comme la présence d'une telle quantité individuelle dans deux événements différents. Cette conception n'est ni implicitement causale ni anthropomorphique ; en revanche, elle est symétrique. Notre thèse est alors que l'asymétrie de la causalité et l'asymétrie du temps (l'existence d'une direction privilégiée du temps) ne sont pas des composantes *a priori* du concept de causalité, mais que cette asymétrie est une propriété contingente des relations causales du monde réel, qui a son origine dans l'existence d'une large majorité de processus causaux intrinsèquement asymétriques et irréversibles : ce sont les processus d'évolution des systèmes dont l'entropie augmente⁷⁴.

La deuxième objection mentionnée par Sosa et Tooley concerne la théorie causale de la persistance. À l'égard de la persistance des objets macroscopiques dans le temps, deux conceptions s'affrontent : selon la conception de « l'endurance », pour utiliser le terme

⁷⁴ Il y a également des processus microphysiques intrinsèquement asymétriques, en particulier la désintégration des mésons-K, ou « kaons ». On dit de ces processus de décomposition qu'ils « violent » la symétrie à l'égard de l'inversion temporelle (« T »). L'argument selon lequel les décompositions en question ne sont pas réversibles, autrement dit symétriques par rapport au temps, est indirect. Ce qu'on observe (bien entendu, encore de façon indirecte) est une violation des symétries à l'égard des opérations de conjugaison de charge (« C ») et d'inversion spatiale (parité : « P »). Sur la base de la validité du théorème dit de « CPT » selon lequel toutes les interactions sont conjointement symétriques à l'égard de la conjugaison de la charge (« C »), à l'égard de l'inversion spatiale (parité : « P »), et à l'égard de l'inversion temporelle (« T »), on infère que la symétrie à l'égard du temps est, elle aussi, violée au cours des interactions de décomposition des mésons-K. Cf. Lee et al. (1957); Sachs (1987); Dowe (1992a, p. 189).

employé dans la littérature anglophone⁷⁵, l'identité des objets n'est déterminée que dans les trois dimensions spatiales, mais non dans la dimension temporelle, d'où l'appellation de « tridimensionnalisme » parfois utilisée pour caractériser cette doctrine. Un tel objet tridimensionnel qui correspond à la « substance » de la tradition philosophique, persiste à travers le temps sans pour autant posséder des « parties temporelles ». L'objet lui-même, dans sa totalité, est présent à différents instants temporels. Le concept d'un objet qui est entièrement présent à plusieurs instants est analogue à la conception traditionnelle d'une propriété universelle. On considère qu'une propriété comme le rouge est entièrement présente dans toutes ses instances dans les objets rouges. La capacité d'être entièrement présent à plusieurs endroits et plusieurs instants est donc commune aux substances et aux universaux, alors qu'une telle capacité est *a priori* exclue pour les entités quadridimensionnelles : un événement qui a lieu à un endroit et un instant bien délimité ne peut pas être *entièrement* présent à un *autre* endroit spatio-temporel. Seules ses *parties* occupent différentes zones spatio-temporelles. Selon la conception alternative de la persistance des objets dans le temps, en termes de « perdurance », l'identité des objets est déterminée dans quatre dimensions : selon cette doctrine souvent appelée « quadridimensionnalisme », un objet a des limites à la fois dans l'espace et dans le temps, et il a non seulement des parties spatiales, mais aussi des parties temporelles.

Dans le cadre du tridimensionnalisme, l'identité d'un objet dans le temps, ou « génidentité »⁷⁶, est une vraie identité qui est primitive et irréductible ; en revanche, pour le quadridimensionnalisme, il ne s'agit pas d'une identité au sens logique strict, mais plutôt d'une relation qui relie les différentes parties temporelles d'un même objet. Dans le cadre de cette dernière doctrine, il semble plausible de penser que cette relation, fondatrice de l'identité des objets dans le temps, est la relation causale⁷⁷. Or – et cela est l'objection de Sosa et Tooley contre la version de la théorie de transfert défendue par Fair (1979) – tant que la relation causale est conçue sur le modèle d'un transfert entre différents objets, la relation entre les différentes parties temporelles d'un même objet ne peut pas être conçue comme causale, ce qui prive la conception quadridimensionnaliste de la persistance dans le temps de son fondement le plus plausible. La version de la théorie du transfert que nous avons proposée plus haut et qui conçoit la causalité en terme d'un transfert entre *événements*, et non plus entre *objets*, évite cette objection : elle peut expliquer la persistance des objets dans le temps, à partir de l'existence d'une relation

⁷⁵ Ce terme, ainsi que celui de « perdurance », a été introduit par Mark Johnston. Cf. Lewis (1986a).

⁷⁶ Sur la notion de génidentité, cf. Kistler (2001b).

⁷⁷ Cela a notamment été suggéré par Armstrong (1980).

causale entre ses parties temporelles qui sont des événements, grâce au transfert de quantités de différentes GC, à commencer par la masse-énergie.

La troisième objection porte sur le statut de la théorie. La théorie de transfert implique un changement radical dans la conception même du problème de l'analyse du concept de causalité, en ce sens qu'elle ne cherche plus à en fournir une analyse purement conceptuelle et *a priori*, comme c'est le cas de toutes les doctrines alternatives. Au contraire, son hypothèse fondamentale est que la relation causale est un type naturel de relation dont il s'agit de découvrir *a posteriori*, par une spéculation en principe empiriquement testable, la nature sous-jacente ou, en jargon philosophique, « l'essence réelle ». En ce sens, la découverte de cas de télépathie ou d'autres phénomènes d'action à distance réfuterait la théorie de transfert de manière empirique, alors que les autres conceptions ne peuvent être réfutées que par des arguments conceptuels *a priori*. Selon notre conception de la causalité comme espèce naturelle de relation, la causalité n'existe pas nécessairement. Comme pour d'autres espèces naturelles comme l'or, rien n'empêche l'existence de mondes possibles qui ne contiennent pas cette espèce. Cela signifie que dans un monde possible dans lequel il n'y a pas de grandeurs conservées, il n'y a pas non plus de causes et d'effets. Par exemple, dans un monde où seuls des esprits immatériels existent ceux-ci n'interagiraient pas causalement. Plus pertinent pour notre propre monde est l'observation que les corrélations à distance prédites par la mécanique quantique et mises en évidence expérimentalement, ne sont pas causales. Ces corrélations instantanées à distance entre les parties A et B (plus précisément, A et B sont des événements simultanés situés à des endroits distants) d'un système enchevêtré (« *entangled* ») quantique ne réfutent pas la conviction selon laquelle il n'existe pas d'actions à distance dans notre monde (plus précisément, d'actions qui agissent *immédiatement* à distance), car ces corrélations ne peuvent pas être employées pour transmettre des signaux, et plus généralement pour obtenir un changement instantané à B, en opérant un changement en A.

Nous pouvons accepter la thèse de Sosa et Tooley selon laquelle « la causalité possède une nature intrinsèque, de manière à ce que la causalité doit être la même relation dans tous les mondes possibles » (Sosa et Tooley 1993, p. 3), en ajoutant qu'elle a cette nature essentielle dans tous les mondes possibles où elle existe. Nous concevons l'essence de la causalité comme une essence réelle découverte empiriquement, *a posteriori*, plutôt que, comme le font ces auteurs en suivant la tradition, comme une essence nominale qui constitue le contenu *a priori* du concept de causalité. Mais cette conception ne justifie pas l'objection de Sosa et Tooley selon laquelle « l'on peut faire appel à la possibilité de mondes qui contiennent de la causalité

mais qui ne contiennent pas les relations physicalistes en question [en l'occurrence, la relation de transfert de GC ; M.K.] – ou de manière plus radicale, qui ne contiennent aucun état physicaliste du tout – afin de tirer la conclusion que la causalité ne peut, même dans ce monde-ci, être identique à aucune relation physicaliste » (Sosa et Tooley, *ibid.*). Si la causalité est une espèce naturelle, de tels mondes qui réfuteraient notre conception, n'existent pas plus que des mondes possibles où existe de l'or qui n'a pas le nombre atomique 79⁷⁸. Dans de tels mondes, il y aurait peut-être une autre source de corrélations, mais ce ne serait pas la causalité : ce ne serait pas une relation qui partage l'essence réelle de la causalité qui existe dans notre monde.

RESPONSABILITÉ CAUSALE

J'ai essayé de montrer que les théories nomologique, contrefactuelle et probabiliste ne parviennent pas à analyser la relation causale singulière entre événements. Il nous reste à montrer que leur utilité peut être reconnue dans le cadre d'un usage différent du concept de causalité qui ne concerne pas la relation causale entre événements, mais une relation entre *faits* concernant ces événements. Si une balle blanche de billard A met en mouvement une balle rouge B (par choc élastique frontal, où B a été au repos avant le choc), il s'agit là d'une relation causale par excellence. Mais nous pouvons caractériser cette relation de manière plus fine, car certaines propriétés de la cause sont responsables de certaines propriétés de l'effet, alors que d'autres ne le sont pas. Le fait que A vienne heurter B avec une quantité de mouvement M est causalement responsable du mouvement de B avec M après le choc. Mais le fait que A soit blanche n'est pas causalement responsable du fait que B soit rouge après le choc. Or, la dépendance de la possession de certaines propriétés par l'effet, par rapport à la possession de certaines propriétés par la cause est une dépendance nomique qui est à l'origine d'une dépendance contrefactuelle, ainsi que d'une relation d'augmentation de probabilité. Si A n'avait pas eu M, B n'aurait pas eu M après le choc non plus. En revanche si A n'avait pas été blanche, B aurait été rouge quand même. La probabilité que B ait M est plus grande si A est sur une trajectoire de choc avec M que si A n'y est pas. En revanche, la probabilité que B soit rouge après le choc est la même si A est présente ou non.

La relation de *responsabilité causale* (désormais abrégée « RC ») nous permet donc d'intégrer la contribution des théories nomologi-

⁷⁸ Il a été établi par Kripke (1972) que les espèces naturelles comme l'or possèdent une essence réelle qui leur appartient nécessairement, tout en étant découverte *a posteriori*. Par conséquent, il n'existe aucun monde possible où l'or n'a pas ses propriétés essentielles, parmi lesquelles on peut supposer que le nombre atomique est central.

que, contrefactuelle et probabiliste à l'analyse de la causalité⁷⁹. Nous écrivons la responsabilité causale du fait que l'événement cause (*c*) a la propriété F pour le fait que l'effet (*e*) a la propriété G, ainsi : CR (Fc, Ge).

Dans une première approche, on peut analyser la relation de RC comme une relation complexe fondée sur l'existence d'une relation causale entre les événements *c* et *e*⁸⁰. En ayant recours à une conception nomologique de la dépendance entre propriétés⁸¹, la relation de responsabilité causale CR (Fc, Ge) d'un fait Fc concernant la cause *c* pour un fait Ge concernant l'effet *e* se laisse analyser en une conjonction C(*c*,*e*) Fc Ge L(F,G). C(*c*, *e*) signifie que les événements *c* et *e* sont liés comme cause et effet en vertu d'un transfert, Fc et Ge signifient que *c* et *e* possèdent respectivement les propriétés F et G, et L(F, G) signifie que l'exemplification de F par *e* dépend nomiquement de l'exemplification de F par *c*.

L'introduction de la responsabilité causale est censée remédier au manque de spécificité de l'analyse de la causalité en termes de transfert. Ce manque de spécificité est le prix à payer pour sa grande généralité. En particulier, lorsqu'il s'agit de comprendre des rapports causaux biologiques ou psychologiques, il ne semble pas très pertinent d'évoquer des transferts d'énergie ou de quantité de mouvement. Le concept de responsabilité causale est censé fournir un cadre qui permet d'intégrer des jugements causaux plus spécifiques. Le transfert en est une partie constitutive essentielle : il n'y a pas de causalité sans transfert. Seulement, lorsqu'il s'agit de comprendre des processus causaux entre objets et événements complexes, les jugements causaux pertinents doivent faire référence à des propriétés bien plus spécifiques que les propriétés physiques qui jouent un rôle direct dans les transferts de quantités de grandeurs conservées.

De nombreuses recherches biologiques visent par exemple à découvrir ce dont différents gènes sont causalement responsables.

⁷⁹ Des relations semblables ont été étudiées par d'autres auteurs (Horgan 1989, McLaughlin 1993, Kim 1993a, Marras 1998) sous le nom de "*qua-causation*" or "*quausation*". D'autres considèrent la relation entre faits comme la relation causale tout court. "Token causal claims are about the causal role of the actual *exemplification of one type* for the actual *exemplification of another type*. [...] 'y's being Y was *because of* (or *token caused by*) x's being X" (Eells 1991, p. 288 ; italiques d'Eells).

⁸⁰ J'ai défendu cette conception dans Kistler (1999a, chap. 5; 2001a). Kistler (1999b) défend l'analyse esquissée ici contre des arguments proposés par Mellor (1995) et autres selon lesquels *seuls* les faits peuvent être termes de relations causales, mais non les événements.

⁸¹ Alternativement, cette dépendance pourrait être conçue sur le modèle de la dépendance contrefactuelle ou d'augmentation de probabilité. Ici n'est pas le lieu pour défendre ma conviction que la dépendance nomique a la priorité métaphysique, et que les autres relations ne sont que différentes manières de la dépendance nomique de se manifester.

L'ouvrage de Michel Morange sur « la part des gènes » (Morange 1998) évoque, entre autres, « des gènes qui causent des maladies », « des gènes qui contrôlent la vie et la mort » et « des gènes qui affectent le comportement ». Un concept de causalité à vocation générale doit être adéquat aux mécanismes causaux de niveau moléculaire qui font l'objet des découvertes sur les effets des gènes. A titre d'exemple, le mécanisme qui mène d'une mutation du gène codant normalement pour la chaîne de l'hémoglobine, à l'anémie falciforme (*sickle cell anemia*) est une suite de processus causaux : la mutation conduit au remplacement d'un acide aminé dans la chaîne de la protéine. Ce changement de la structure primaire est responsable d'un changement de la conformation de la molécule (c'est-à-dire sa forme dans l'espace) qui provoque à son tour l'agrégation de molécules d'hémoglobine lorsque la concentration d'oxygène est basse, ce qui cause la déformation des cellules rouges du sang. A la fin de « cette chaîne causale » (Morange 1998, p. 49), la forme anormale des cellules les empêche de transiter normalement par les capillaires fins. Par conséquent, elles les bloquent, ce qui mène à priver les tissus situés en aval des capillaires d'oxygène, ce qui constitue l'anémie. Dans chacun des chaînons de cette longue chaîne de responsabilités, en majeure partie causales, ce n'est pas tant le transfert omniprésent d'énergie qui est pertinent pour l'explication causale, mais la dépendance nomique d'une propriété complexe de l'effet par rapport à une propriété de la cause. C'est la dépendance de la structure primaire de la chaîne de l'hémoglobine par rapport à la séquence du gène qui justifie le jugement que ce gène est causalement responsable de la structure primaire de la protéine (c'est-à-dire de la séquence de ses acides aminés). Mais ce qui fait de cette dépendance nomique un cas de responsabilité *causale* est le transfert d'énergie qui se déroule dans l'espace et le temps. D'autres instances de dépendance nomique ne relèvent pas de la responsabilité causale : la conformation tridimensionnelle de l'hémoglobine dépend nomiquement de sa structure primaire. Cependant, cette dépendance n'est pas causale parce qu'elle concerne deux propriétés que chaque molécule d'hémoglobine possède à chaque instant. Le décalage spatial et temporel caractéristique de la causalité fait défaut. En revanche, la conformation des molécules est bien causalement responsable de leur agrégation, elle-même causalement responsable de la déformation des globules rouges.

Il ressort de cette analyse que l'application du concept de causalité à l'explication de phénomènes complexes, par exemple biologiques, est conditionnée à la fois par la présence d'un transfert et par l'existence d'une dépendance nomique entre certaines propriétés des événements à expliquer. Il ne s'agit toutefois pas d'une simple juxtaposition de deux conditions sans relations entre elles, le transfert et la

dépendance nômique : plutôt, dans les cas les plus simples⁸², les propriétés dont on dit qu'elles sont liées par la responsabilité causale, sont des formes d'apparence de grandeurs conservées, et en particulier de l'énergie. Faute de place, il n'est pas possible de développer ici la nature de la dépendance des propriétés complexes par rapport aux propriétés physiques élémentaires sous-jacentes, auquel on fait allusion lorsqu'on dit que les premières « surviennent »⁸³ sur les secondes ou que les secondes « émergent » à partir des premières.

Les mêmes remarques s'appliquent aux rapports de causalité auxquels on fait appel en psychologie. Dans les recherches sur les espaces de représentation des qualités sensorielles⁸⁴, les représentations jouent le rôle d'intermédiaires causaux entre les stimuli et les actions. Lorsqu'on montre par exemple à un sujet deux couleurs différentes, pour solliciter ensuite un jugement de la similarité de ces couleurs, la stimulation des organes sensoriels est causalement responsable des représentations des deux couleurs, et celles-ci sont à leur tour causalement responsables du jugement. Comme dans les rapports causaux biologiques, les propriétés directement mises en rapport de responsabilité causale ne sont pas des quantités conservées ; les propriétés psychologiques surviennent sur des (ou émergent à partir de) propriétés plus élémentaires neuronales qui à leur tour surviennent sur des propriétés chimiques et physiques. Dans

⁸² L'une des complications peut être illustrée par l'un des chaînons du mécanisme causal par lequel la mutation génétique provoque l'anémie falciforme et qui a une forme négative : le blocage des capillaires est responsable de l'absence d'oxygène dans les tissus normalement irrigués par ces capillaires. Comment une absence peut-elle causer ou être causée si la causalité est fondée sur un transfert ? Une absence ne peut certainement ni transférer ni recevoir quoi que ce soit par transfert. La réponse passe par le constat que les prédicats négatifs ne font pas directement référence aux propriétés impliquées dans les rapports de responsabilité causale. Ces jugements peuvent être vrais sans que cela justifie l'idée selon laquelle les absences et autres « propriétés négatives » sont réelles au sens de pouvoir causer et être causé, d'une manière analogue à celle dont Russell a montré que le jugement « l'actuel roi de France est chauve » peut être faux sans que cela justifie d'attribuer à l'actuel roi de France une forme quelconque de réalité. J'ai montré que l'analyse en termes de responsabilité causale est compatible avec certains jugements de forme négative, tels que les omissions, les interruptions et les déclenchements, dans Kistler (1999a ; 1999b ; 2001a).

⁸³ La survenance est une relation de dépendance asymétrique qui a récemment suscité beaucoup d'intérêt, notamment parce qu'elle promet de porter un nouvel éclairage sur la dépendance de l'esprit par rapport au corps : un ensemble de propriétés A (par exemple mentales) survient sur un ensemble de propriétés B (par exemple, physiques), si et seulement si il est impossible que deux objets diffèrent à l'égard de leurs propriétés A sans différer à l'égard de leurs propriétés B. Cette idée fondamentale a notamment été développée par Kim (1993b).

⁸⁴ Cf. Shepard (1962), Clark (1993).

tous les cas, il y a un processus de transfert physique sous-jacent au rapport de responsabilité causale qui exige en outre la dépendance nomique des propriétés complexes mises en jeu.

CONCLUSION

Le débat philosophique sur le concept de causalité est loin d'être achevé. Comme j'ai essayé de le montrer, toutes les tentatives pour trouver une analyse relativement simple qui réduit la causalité à la dépendance contrefactuelle, à l'augmentation de probabilité, ou à un transfert sont confrontées à des difficultés importantes. Certains répondront que, tout bien pesé, Hume, Russell et Carnap avaient raison de conclure que le concept de causalité n'a pas sa place dans l'ontologie, c'est-à-dire dans la théorie des types d'entités qui existent objectivement. Plutôt que de s'obstiner à concevoir la causalité comme relation objectivement réelle, d'une manière qui s'accorde à la fois avec nos intuitions et avec les résultats scientifiques, on ferait mieux de se rabattre sur l'étude de l'*apparence* causale : il n'y a pas de sens à accorder à l'idée selon laquelle la réalité est causalement structurée ; tout ce qui existe sont des *représentations* de causalité ; or, l'étude de telles représentations appartient au domaine de la psychologie, non de l'ontologie. Une autre conclusion que l'on pourrait être tenté de tirer de l'échec (provisoire) de ces tentatives d'analyse serait de dire qu'il n'existe aucune relation unique qui s'applique dans tous les contextes : il n'y aurait que des relations causales hétérogènes qui ne partagent aucun principe commun. Selon ce point de vue, il y aurait une causalité propre à la physique, une autre propre à la biologie, et une autre encore qui s'applique en psychologie.

Cependant, il me semble prématuré de se résigner à la conclusion selon laquelle la tentative de trouver une analyse unique et satisfaisante de ce concept a définitivement échoué, dans la mesure où tout un axe de recherche n'a pas encore été suffisamment exploré, à savoir celui de la construction d'analyses hybrides. Ainsi, Dowe (2001) a proposé une conception qui réunit des aspects de la théorie de la causalité comme processus avec la conception contrefactuelle ; Schaffer (2001) a examiné la possibilité d'enrichir la conception en termes de processus avec des éléments de la théorie probabiliste. C'est cette stratégie que j'ai suivie dans la dernière partie de cet article. En effet, à la différence de ses concurrentes contrefactuelle, probabiliste et interventionniste, il me semble que la théorie du transfert ne soit pas confrontée à des contre-exemples aptes à réfuter sa prétention à être compatible avec toutes les relations causales. Plutôt, le défaut de la théorie du transfert qui fait aussi sa force, réside dans sa grande généralité. Cette généralité la fait apparaître peu pertinente à l'analyse de rapports causaux en dehors de la physique. Cependant, j'ai essayé de montrer que ce manque peut être comblé en complétant la notion de transfert par celle de responsabilité causale qui per-

met de reconnaître l'efficacité causale de propriétés non-physiques, notamment biologiques et psychologiques.⁸⁵

Références

- Armstrong, D. M. (1980), Identity through time, in : Peter van Inwagen (ed.), *Time and cause*, Dordrecht, Reidel, p. 67-78.
- Armstrong D. M. (1983), *What is a Law of Nature?*, Cambridge, Cambridge University Press.
- Aronson, J. J. (1971a), On the Grammar of 'Cause', *Synthese* 22, p. 414-430.
- Aronson, J. J. (1971b), The Legacy of Hume's Analysis of Causation, *Stud. Hist. Phil. Sci.* 2, p. 135-165.
- Barberousse, A., Kistler, M., Ludwig, P. (2000), *La philosophie des sciences au XXe siècle*, Flammarion, coll. Champs-Université.
- Berkeley, G. (1721), Of Motion (De Motu), in: *The Works of George Berkeley*, A.A. Luce and T.E. Jessop (eds.), vol. IV. repr. Nendeln 1979, p. 31-52. Trad. Par D. Berlioz-Letellier et M. Beyssade, Du mouvement, in: George Berkeley, *Oeuvres II*, G. Brykman (dir.), p. 151-181.
- 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 Sylvain Bromberger, *On What We Know We Don't Know*, Chicago, University of Chicago Press, et Stanford, CSLI, 1992, p. 75-100.
- Carnap, R. (1966), *Philosophical Foundations of Physics*, New York, Basic Books; trad. par J.M. Luccioni et A. Soulez : *Les fondements philosophiques de la physique*, Paris, A. Colin, 1973.
- Cartwright, N. (1983), *How the Laws of Physics Lie*, Clarendon Press, Oxford.
- Clark, A. (1993), *Sensory Qualities*, Clarendon Press, Oxford.
- Comte, Auguste (1844), *Discours sur l'esprit positif*, Paris, Vrin, 1983.
- Cummins, Robert (1983), *The Nature of Psychological Explanation*, Cambridge MA, MIT Press.
- Davidson, D. (1967), Causal Relations, repr. in : D.D., *Essays on Actions and Events*, Clarendon Press, Oxford 1980 trad. par P. Engel, *Actions et événements*, P.U.F., 1993.
- Davidson, D. (1995), Laws and cause, *Dialectica* 49, p. 263-279.
- Dowe, P. (1992), Wesley Salmon's Process Theory of Causality and the Conserved Quantity Theory, *Phil. of Science* 59, pp. 195-216.
- Dowe, P. (1992a), Process Causality and Asymmetry, *Erkenntnis* 37, p. 179-196.
- Dowe, P. (1996), Backwards Causation and the Direction of Causal Processes. *Mind* 105, p. 1-22.

⁸⁵ Je tiens à remercier les rapporteurs pour *Intellectica* de leurs remarques critiques, et Benoît Hardy-Vallée pour son aide linguistique.

- Dowe, P. (2001), A Counterfactual Theory of Prevention and “causation” by omission, *Australasian Journal of Philosophy* 79, p. 216-226.
- Edgington, D. (1997), Mellor on Chance and Causation, *British Journal for the Philosophy of Science* 48, p. 411-433.
- Eells, E. (1991), *Probabilistic Causality*, Cambridge University Press.
- Eells, E. and Sober, E. (1983), Probabilistic Causality and the Question of Transitivity, *Phil. of Science* 50, p. 35-57.
- Ehring, D. (1997), *Causation and persistence: A Theory of Causation*. New York: Oxford University Press.
- Fair, D. (1979), Causation and the Flow of Energy, *Erkenntnis* 14, p. 219-250.
- Galilei, G. (1638), Discorsi e dimostrazioni matematiche intorno a due nuove scienze, in: *Le Opere di Galileo Galilei*, vol. 8, Edizione Nazionale, Firenze, G. Barbera 1898; trad. par M. Clavelin, *Discours concernant deux sciences nouvelles*, A. Colin 1970.
- Gasking, D. (1955), Causation and Recipes, *Mind* 64, pp. 479-487.
- Good, I.J. (1961-62), A Causal Calculus I-II, *British Journal for the Philosophy of Science* 11, p. 305-318; 12, p. 43-51; Errata and corrigenda : 13, p. 88. Repr. in I.J. Good, *Good Thinking*, Minneapolis, University of Minnesota Press, 1983.
- Goodman, N. (1955), *Fact, Fiction and Forecast*, 3e édition, Indianapolis, Bobbs-Merrill, 1973; trad. revue par P. Jacob : *Faits, fictions, et prédictions*, Paris, Minuit, 1984.
- Grünbaum, A. (1988), Preface, in: Grünbaum A. et Salmon W. (eds.) (1988), *The Limitations of Deductivism*, Los Angeles, University of California Press, p. ix-xv.
- Hall, N. (2000), Causation and the Price of Transitivity, *Journal of Philosophy* 97, p. 198-222.
- Hausman, D. (1986), Causation and Experimentation, *American Phil. Quarterly* 23, No. 2, pp. 143-154
- Heidelberger, M. (1992), Kausalität: Eine Problemübersicht. *Neue Hefte für Philosophie* 32/33, p. 130-153
- Hempel, C. G. (1942), The Function of General Laws in History, repr. in Hempel (1965b).
- Hempel, C. G. et Oppenheim, P. (1948), Studies in the Logic of Explanation, repr. in Hempel (1965b).
- Hempel, Carl G. (1965a), Aspects of Scientific Explanation, in: Hempel (1965b).
- Hempel, C. G. (1965b), *Aspects of Scientific Explanation*, New York: The Free Press.
- Hempel, C. G. (1988), Provisos: A Problem Concerning the Inferential Function of Scientific Theories, in: A. Grünbaum et Wesley C. Salmon, *The Limitations of Deductivism*, University of California Press, Berkeley, p. 19-36
- Horgan, T. (1989), Mental Quausation, in: Tomberlin J.E. (ed.), *Philosophical Perspectives 3: Philosophy of Mind and Action Theory*, Atascadero, CA: Ridgeview, 1989.

- Hume, D. (1739), *A Treatise of Human Nature*, seconde édition, par L.A. Selby-Bigge et P.H. Nidditch, Oxford, Clarendon Press, 1978; trad. fr. par A. Leroy : *Traité de la nature humaine*, Paris, Aubier 1983.
- Hume, D. (1748), *An Enquiry Concerning Human Understanding*, in *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, eds. L.A. Selby-Bigge et P.H. Nidditch, 3e édition, Oxford, Clarendon Press, 1975; trad. fr. par A. Leroy, *Enquête sur l'entendement humain*, Paris, Aubier, 1947.
- Humphreys, P. W. (1989), The Causes, Some of the Causes, and Nothing But the Causes, in: *Minnesota Studies in the Philosophy of Science, vol XII: Scientific Explanation*, Minneapolis, University of Minnesota Press, p. 283-306.
- Jackson, F. (1977), A Causal Theory of Conterfactuals, *Australasian Journal of Philosophy*, 55, p. 3-21.
- Keil, G. (2000), *Handeln und Verursachen*. Vittorio Klostermann, Frankfurt a.M.
- Kim, J. (1993a), The Non-Reductivist's Troubles with Mental Causation, in: J. Heil and A. Mele (eds.), *Mental Causation*, Clarendon Press, Oxford, 1993, p. 189-210.
- Kim, J. (1993b) *Supervenience and mind*, Cambridge, Cambridge University Press.
- Kistler, M. (1998), Reducing Causality to Transmission, *Erkenntnis* 48 (1998), p. 1-24.
- Kistler, M. (1999a), *Causalité et lois de la nature*, Vrin.
- Kistler, M. (1999b), Causes as events and facts, *Dialectica* 53, p. 25-46.
- Kistler, M. (2001a), Causation as transference and responsibility, in W. Spohn, M. Ledwig & M. Esfeld (eds.), *Current Issues in Causation*, Paderborn, Mentis, p. 115-133.
- Kistler, M. (2001b), Le concept de génidentité, in S. Laugier (éd.), *Carnap et la construction logique du monde*, Paris, Vrin, p. 163-188.
- Kistler, M. (2002), The Causal Criterion of Reality and the Necessity of Laws of Nature, *Metaphysica* 3, p. 57-86.
- Kistler, M. (2002b), L'identité des propriétés et la nécessité des lois de la nature, *Cahiers de philosophie de l'Université de Caen, Numéro spécial 38/39 : Le réalisme des universaux*, p. 249-273.
- Kistler, M. (2003), Laws of Nature, Exceptions and Tropes, *Philosophia scientiae* 7 (2), p. 189-219.
- Kistler, M. (à paraître), Necessary Laws, in : J. Faye, M. Urchs and U. Scheffler (eds.), *Principles of Nature*, Dordrecht, Kluwer.
- Kneale, W. (1949), *Probability and Induction*, Oxford, Oxford University Press.
- Krajewski, W. (1982), Four Conceptions of Causation, in: W. Krajewski ed., *Polish Essays in the Philosophy of the Natural Sciences*, Reidel: Dordrecht, 1982.
- Krajewski, W. (1997), Energetic, informational, and triggering causes. *Erkenntnis* 46, p. 193-202.

- Kripke, S. A. (1972), Naming and Necessity. in: Harman and Davidson (eds.), *Semantics of Natural Language*. Dordrecht: Reidel.
- Kyburg, H. E. jr (1965), Comment, *Philosophy of Science* 32, p. 147-151.
- Lee, T.D., Oehme, R., and Yang, C.N. (1957), Remarks on Possible Noninvariance under Time Reversal and Charge Conjugation, *Phys. Rev.* 106, p. 340-345.
- Lewis, D. (1973), *Counterfactuals*, Oxford, Basil Blackwell.
- Lewis, D. (1986), *Philosophical Papers, vol. II*, New York, Oxford University Press.
- Lewis, D. (1986a), *On the Plurality of Worlds*, Oxford, Basil Blackwell.
- Lewis, D. (2000), Causation as Influence, *Journal of Philosophy* 97, p. 182-197.
- Lorenz, K. (1941), Kants Lehre vom Apriorischen im Lichte gegenwärtiger Biologie, *Blätter für deutsche Philosophie* 15.
- Marras, A. (1998), Kim's Principle of Explanatory Exclusion, *Australas. J. Phil.* 76, p. 439-451.
- McLaughlin, B. (1993), On Davidson's Response to the Charge of Epiphenomenalism, in J. Heil and A. Mele (eds.), *Mental Causation*, Clarendon Press, Oxford, 1993, p. 27-40.
- Mellor, D.H. (1995), *The Facts of Causation*, London et New York, Routledge.
- Menzies, P. et Price, H. (1993), Causation as a Secondary Quality, *British J. Phil. of Science* 44, p. 187-203.
- Mill, J. S. (1843), *A System of Logic, Ratiocinative and Inductive*, London, Parker; trad. par L. Peisse, *Système de logique déductive et inductive*, Paris, Ladrangé 1866; repr.: Liège: Pierre Mardaga 1988.
- Morange, M. (1998), *La part des gènes*, Paris, O. Jacob. Trad. anglaise par M. Cobb, *The Misunderstood Gene*, Harvard University Press, 2001.
- Nagel, E. (1961), *The Structure of Science*, London: Routledge and Kegan Paul.
- Papineau, D. (1989), Pure, Mixed, and Spurious Probabilities and Their Significance for a Reductionist Theory of Causation, in: *Minnesota Studies in the Philosophy of Science, vol XII: Scientific Explanation*, Minneapolis, University of Minnesota Press, p. 307-356.
- Popper, K. (1934), *Logik der Forschung*, 3ème 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.
- Price, H. (1992), Agency and Causal Asymmetry, *Mind* 101, pp. 501-520.
- Putnam, H. (1984), Is the Causal Structure of the Physical Itself Something Physical ?, *Midwest Studies in Philosophy 9 : Causation and Causal Theories*, p. 3-16; repr. in H.P., *Realism with a Human Face*, ed. by James Conant, Cambridge, MA: Harvard University Press, 1990; trad. par Claudine Tiercelin, *Le réalisme à visage humain*, Paris, Ed. du Seuil, 1994.
- Putnam, H. (1992), *Renewing Philosophy*, Cambridge, MA: Harvard University Press; trad. par Luc Biquez, sous la dir. de Pascal Marin,

- Pour une nouvelle philosophie*, Lyon, Université catholique de Lyon (mémoire de DEA), 2000.
- Reichenbach, H. (1956), *The Direction of Time*, Berkeley, University of California Press, 1991.
- Russell, B. (1912), On the Notion of Cause, repr. in B.R., *Mysticism and Logic* (1917), London, Routledge, 1986.
- Russell, B. (1914), *Our Knowledge of the External World*, London, Routledge, 1993.
- Sachs, R. G. (1987), *The Physics of Time Reversal*, Chicago, University of Chicago Press.
- Salmon, W. (1984), *Scientific Explanation and the Causal Structure of the World*, Princeton University Press.
- Salmon, W. (1990), *Four Decades of Scientific Explanation*, Minneapolis, University of Minnesota Press.
- Salmon, W. (1994), Causality Without Counterfactuals, *Phil. of Science* 61, p. 297-312.
- Schaffer, J. (2000), Trumping Preemption, *Journal of Philosophy* 97, p. 165-181.
- Schaffer, J. (2001), Causes as probability raisers of processes, *Journal of Philosophy* 98, p. 75-92.
- Schiven, M. (1962), Explanation, Prediction, and Laws, in H. Feigl et G. Maxwell (eds.), *Minnesota Studies in the Philosophy of Science*, III, Minneapolis, University of Minnesota Press, p. 170-230.
- Shepard, R. (1962), The Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function, *Psychometrika* 27, p. 125-140 et 219-246.
- Simpson, E.H. (1951), The Interpretation of Interaction in Contingency Tables, *J. of the Royal Statistical Society, Ser. B*, 13, p. 238-241.
- Sosa, E. et Tooley, M. (1993), Introduction, in: Sosa E. et Tooley M. (eds.), *Causation*, Oxford University Press.
- Stalnaker, R. (1968), A Theory of Conditionals, in: *Studies in Logical Theory*, Nicholas Rescher (ed.), Oxford, Blackwell.
- Stegmüller, W. (1969), *Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie, Vol. I.: Erklärung, Begründung, Kausalität*. Berlin, Springer, 1983.
- Suppes, F. (1970), *A Probabilistic Theory of Causality*, Amsterdam, North Holland.
- von Wright, G. H. (1971), *Explanation and Understanding*, Ithaca (N.Y.), Cornell University Press.
- Yakira, E. (1994), *La causalité de Galilée à Kant*, PUF.
- Zimmermann, D. W. (2000), Shoemaker's Argument for his Theory of Properties, *Facta philosophica* 2, p. 271-290.