

L'INTERVENTIONNISME PERMET-IL LA CAUSALITÉ « DESCENDANTE » ?

Alexandre Marcellesi

*Department of Philosophy
University of California, San Diego, 9500 Gilman Drive, La Jolla
CA 92093-0119
alexandre@ucsd.edu*

RÉSUMÉ

De nombreux tenants du physicalisme non-réductif défendent l'idée qu'adopter la théorie interventionniste de la causalité développée par James Woodward permet de résoudre le problème de l'exclusion causale. Je m'intéresse ici à l'argument proposé par Panu Raatikainen, selon lequel les propriétés survenantes peuvent avoir des pouvoirs causaux indépendants de ceux des propriétés les réalisant, et y oppose plusieurs objections. Je montre en particulier que lorsque la théorie interventionniste est correctement appliquée, elle interdit les cas de causalité descendante.

ABSTRACT

Several advocates of non-reductive physicalism have argued that adopting Woodward's interventionist account of causation affords them a way out of the causal exclusion problem. I present an argument put forward by Panu Raatikainen according to which supervenient properties can have causal powers not possessed by their realizers, and raise several objections against it. I show, in particular, that when properly applied, interventionism precludes cases of downward causation.

MOTS-CLÉS

Interventionnisme, Physicalisme non-réductif, Causalité descendante, Panu Raatikainen, James Woodward

1 INTRODUCTION

Paul Raatikainen a récemment défendu la thèse selon laquelle la théorie interventionniste de la causalité développée par James Woodward (Woodward, 2003) fournit au physicalisme non-réductif un cadre dans lequel développer une solution au problème de l'exclusion causale (Raatikainen, *à paraître*). Des arguments similaires ont été proposés par d'autres physicalistes non-réductifs cherchant à préserver la pertinence causale des propriétés survenantes face à l'argument de l'exclusion causale de Kim (Kistler, *à paraître*; Menzies & List, *à paraître*; Shapiro, *à paraître*; Shapiro & Sober, 2007; Weslake, Ms.). Selon ces différents arguments, rejeter la conception nomologique de la causalité à l'œuvre dans l'argument de l'exclusion causale¹, et adopter à la place la conception interventionniste, permet au physicaliste non-réductif de résister à la conclusion selon laquelle les propriétés survenantes ne peuvent être causalement pertinentes que si elles sont identiques aux propriétés sur lesquelles elles surviennent. Bien que l'argument avancé par Raatikainen soit à première vue convaincant, un examen détaillé de la théorie interventionniste révèle que, loin d'aider le physicalisme non-réductif, celle-ci rend impossible la causalité « descendante ». Après une brève présentation de l'interventionnisme, et de ses aspects les plus pertinents pour la question de la causalité « descendante » (§2), je présente l'argument avancé par Raatikainen en faveur de la possibilité de cas de causalité « descendante » (§3). Après avoir clarifié certains aspects de l'argument, j'y oppose une objection décisive, et conclus l'incompatibilité de l'interventionnisme et de la causalité « descendante » (§4).

2 UNE BRÈVE INTRODUCTION À LA THÉORIE INTERVENTIONNISTE DE WOODWARD

Une brève présentation de l'approche interventionniste est nécessaire afin d'évaluer l'argument de Raatikainen en connaissance de cause. L'intuition à la base de l'interventionnisme est qu'une cause constitue un moyen fiable de manipuler son effet. Voici la manière dont Woodward caractérise les concepts de cause *directe* et de cause *contributive*² :

(M) Une condition nécessaire et suffisante pour que X soit une cause *directe* de Y par rapport à un ensemble de variables V est qu'il soit

1. Voir par exemple (Kim, 1998, 33, trad. A.M.) où Kim accepte explicitement « la condition nomologique [...] selon laquelle deux événements qui sont dans une relation causale doivent instancier une loi causale ».

2. Le concept que j'utiliserai ici est celui de cause contributive, puisqu'il n'est pas nécessaire que les propriétés survenantes, pour être causalement pertinentes, soient des causes directes, mais seulement qu'elles soient des causes contributives. Dans ce qui suit, les expressions « X cause Y », « X est une cause contributive de Y » et « X est pertinent causalement vis-à-vis de Y » sont donc utilisées comme des synonymes.

possible d'intervenir sur X de manière à modifier Y , ou la distribution de probabilité sur Y , lorsque l'on fige à une valeur particulière les autres variables Z_i dans V . Une condition nécessaire et suffisante pour que X soit une cause *contributive* de Y par rapport à un ensemble de variables V est que (i) il existe un chemin dirigé de X vers Y tel que chaque arc dans ce chemin est une relation de causalité directe ; c'est-à-dire, un ensemble de variables $Z_1 \dots Z_n$ telles que X est une cause directe de Z_1 , qui est elle-même une cause de Z_2 , qui est elle-même une cause de $\dots Z_n$, qui est elle-même une cause de Y , et (ii) il existe une intervention sur X qui modifie Y lorsque les autres variables dans V qui ne sont pas sur ce chemin sont figées à une valeur particulière. (Woodward, 2003, 59, trad. A. M.)

Trois aspects importants de la conception interventionniste sont à souligner. Tout d'abord, l'interventionnisme appartient à la famille des conceptions contrefactuelles de la causalité. Dire que X est une cause contributive de Y , c'est dire que le contrefactuel « interventionniste » suivant est vrai :

(IC) Si l'on intervenait sur X de manière à changer sa valeur de $X = x_1$ à $X = x_n$, tout en maintenant figées les valeurs des variables pertinentes causalement vis-à-vis de Y mais ne se trouvant pas sur le chemin dirigé $X \rightarrow Y$, alors la valeur de Y changerait de $Y = y_1$ à $Y = y_n$ ³

Bien qu'il s'agisse d'une conception contrefactuelle de la causalité, l'interventionnisme diffère de manière importante des conceptions contrefactuelles traditionnelles telles que celles proposées par David Lewis (Lewis, 1973, 2006)⁴.

Une autre caractéristique importante de l'interventionnisme est qu'il s'agit d'une théorie dite « contrastive » de la causalité. Les théories contrastives de la causalité prennent au sérieux, et rendent explicites, les contrastes implicites dans la plupart des énoncés causaux. Quand je dis que j'ai attrapé un rhume parce que j'ai promené mon chien sous la pluie, ce que je veux dire c'est que promener mon chien sous la pluie, plutôt qu'être resté à la maison, est ce qui a causé mon rhume. Si la classe de contraste pour la cause putative de mon rhume incluait un cas dans lequel je promène mon chien sous la neige, mais pas le cas dans lequel je reste chez moi, alors l'énoncé causal cesserait d'être vrai. Bien que les théories contrastives de la causalité semblent saisir un aspect important du concept de cause, la manière dont doivent être fixées les classes de contraste est rarement évidente. Si le contexte se charge souvent de les déterminer, ce n'est pas toujours le cas⁵.

3. Ici, x_n et y_n dénotent n'importe quel membre de la classe de contraste pour X et Y autres que x_1 et y_1 . Dans les applications à venir de (IC), j'omettrai la clause « tout en maintenant figées les valeurs des variables pertinentes causalement vis-à-vis de Y mais ne se trouvant pas sur le chemin dirigé $X \rightarrow Y$ » pour des raisons de brièveté.

4. Voir (Woodward, 2003, 133-145) pour une comparaison détaillée entre l'interventionnisme et les théories contrefactuelles de Lewis.

5. Voir par exemple (Northcott, 2008) pour une défense des conceptions contrastives et une discussion des principaux problèmes qui y sont liés.

Un troisième point important à relever a trait au concept d'*intervention*. Toutes les manipulations de la valeur d'une variable ne constituent pas des interventions. Voici les quatre conditions qu'une manipulation I de X doit satisfaire pour être une intervention sur X par rapport à Y ⁶ :

- I1 I cause X .
- I2 I agit comme un interrupteur vis-à-vis des autres variables causant X . En d'autres termes, certaines valeurs de I sont telles que lorsque I les atteint, X cesse de dépendre des valeurs des autres variables causant X pour ne plus dépendre que de la valeur de I .
- I3 Tous les chemins dirigés de I vers Y passent par X . En d'autres termes, I ne cause pas directement Y et n'est une cause d'aucune cause de Y autre que X elle-même, exceptées celles des causes de Y , s'il en existe, se trouvant sur le chemin I - X - Y lui-même : c'est-à-dire exceptées
 - (a) les causes de Y qui sont des effets de X (i.e. les variables situées causalement entre X et Y) et
 - (b) toute cause de Y située entre I et X et n'ayant pas d'effet sur Y indépendamment de X .
- I4 I est (statistiquement) indépendant de toute variable Z causant Y et ne se trouvant pas sur un chemin passant par X . (Woodward, 2003, 98, trad. A.M.)

Mon objection principale contre Raatikainen est directement liée à une application illégitime du concept d'intervention (cf. §4.2). Mais avant d'examiner les objections auxquelles il se heurte, il convient dans un premier temps de présenter cet argument et de tirer au clair les conséquences qu'entend en tirer Raatikainen.

3 L'ARGUMENT DE RAATIKAINEN

Raatikainen décrit la thèse qu'il entend défendre de la manière suivante :

Dans le cadre interventionniste, un état mental peut être la cause véritable d'un comportement, par exemple ; et [...] l'état physique sous-jacent peut ne pas être la cause. (Raatikainen, à *paraître*, 7, trad. A. M.)

Les états (ou événements) mentaux peuvent être des causes authentiques, c'est-à-dire qu'il y a, en un sens, de la causalité descendante (*ibid.*, 10)

Raatikainen développe son argument en utilisant l'exemple suivant : imaginez que John et Peter sont colocataires. Par une belle soirée d'été, John forme le désir de boire une bière. John croit également qu'il y a de la bière

6. Pour être plus précis, ces quatre conditions définissent ce que c'est pour X que d'être une variable d'intervention par rapport à Y . Une intervention est une manipulation de X satisfaisant ces conditions.

dans le frigo. Appelons cette croyance P. Étant donné son désir et sa croyance, John se dirige vers ce frigo pour y prendre une bière. Appelons ce comportement Q. Les choses auraient néanmoins pu se passer différemment. Peter aurait pu dire à John qu'il n'y avait plus de bière. Intuitivement, dans ce scénario contrefactuel, John ne se serait pas dirigé vers le frigo pour y chercher une bière. Il aurait plutôt été à l'épicerie ou au bar pour satisfaire son désir. Les physicalistes, qu'ils soient réductifs ou non-réductifs, acceptent la thèse selon laquelle la croyance de John, P, est réalisée par un état de son cerveau, appelons-le B. La question suivante se pose alors : Est-ce que P a causé Q ? Ou bien est-ce que B a causé Q ? Ou bien P et B sont-ils deux causes contributives de Q ? Le diagramme suivant permet de représenter la situation à laquelle nous sommes confrontés (figure 1).

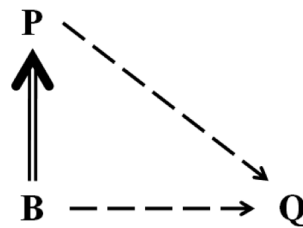


Fig. 1

Ici les flèches simples représentent des relations causales tandis que les flèches doubles représentent des relations de survenance. La thèse de Raatikainen est qu'il peut exister une relation causale « descendante » et non-redondante entre P et Q, c'est-à-dire que P peut être causalement pertinent vis-à-vis de Q alors que B ne l'est pas. Pour déterminer la valeur de vérité de l'énoncé selon lequel P cause Q, la première étape est de sélectionner un ensemble de variables V. Ici les variables seront X, W et Y représentant respectivement P, B et Q. L'étape suivante est de déterminer les différentes valeurs possibles de ces variables. Raatikainen détermine les classes de contraste pour X, W et Y de la manière suivante :

1. $X = x_1$ lorsque John a la croyance P et $X = x_2$ lorsque ce n'est pas le cas.
2. $W = w_1$ lorsque John a l'état cérébral B et $W = w_2$ lorsque ce n'est pas le cas.
3. $Y = y_1$ lorsque John a le comportement Q et $Y = y_2$ lorsque ce n'est pas le cas.

Comme je l'ai indiqué plus haut, B réalise la croyance P de John, de sorte que lorsque $X = x_1$, $W = w_1$. Raatikainen justifie la manière dont il choisit les classes de contraste de la manière suivante :

Le contraste le plus naturel, le contraste « par défaut » [...] est – à moins qu'il existe une raison spécifique de le choisir différemment – que c'est la présence plutôt que l'absence de la propriété en question [i.e. l'effet] qui

est causée par la présence d'une propriété appropriée [i.e. la cause] plutôt que son absence.

Voici une version « interventionniste » de la figure 1 :

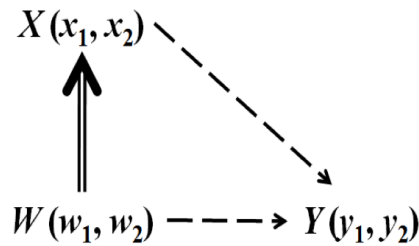


Fig. 2

Maintenant que le problème est posé en termes interventionnistes, il est possible de répondre de manière plus précise à la question de savoir si X est causalement pertinent vis-à-vis de Y , c'est-à-dire à la question de savoir si la croyance de John est pertinente causalement vis-à-vis de son comportement. Pour que ce soit le cas, le contrefactuel interventionniste suivant doit être vrai :

- (1) Si l'on intervenait sur X de manière à changer sa valeur de $X = x_1$ à $X = x_2$, la valeur de Y changerait de $Y = y_1$ à $Y = y_2$.

Selon ce contrefactuel, une intervention sur la croyance de John de manière à l'éliminer résulterait en une modification de son comportement. Et, intuitivement, il semble vrai que si John n'avait pas cru qu'il y avait de la bière dans le frigo, il ne se serait pas dirigé vers le frigo pour y prendre une bière et satisfaire son désir. La vérité de (1) implique donc la pertinence causale de X vis-à-vis de Y .

Nous ne sommes cependant qu'à mi-chemin. Montrer qu'une propriété survenante peut être pertinente causalement vis-à-vis d'une propriété de base et montrer que cette pertinence causale n'est pas redondante sont deux entreprises distinctes. En d'autres termes, après avoir montré que X pouvait être pertinent causalement vis-à-vis de Y , il nous faut à présent montrer que ce n'est pas le cas de W . La pertinence causale de W vis-à-vis de Y est liée au contrefactuel suivant :

- (2) Si l'on intervenait sur W de manière à changer sa valeur de $W = w_1$ à $W = w_2$, la valeur de Y changerait de $Y = y_1$ à $Y = y_2$.

Ce qu'exprime ce contrefactuel, c'est que si l'on intervenait de manière à modifier l'état cérébral dans lequel se trouve John, alors son comportement changerait, et il ne se dirigerait pas vers le frigo pour y chercher une bière. Ce contrefactuel est-il vrai ou faux? Raatikainen pense qu'il est faux, et que sa fausseté dérive de la réalisabilité multiple de la croyance de John selon laquelle il y a de la bière dans le frigo. Voici la manière dont il justifie la fausseté de (2) :

Étant donnée notre acceptation de la thèse de la réalisabilité multiple, il devrait être possible qu'un état cérébral B', distinct de B, soit également un réalisateur de la croyance selon laquelle il y a de la bière dans le frigo. En conséquence, il existe un monde possible w dans lequel une intervention modifie l'état cérébral de John de B à B', et dans ce monde John se dirige néanmoins vers le frigo, au lieu d'aller à l'épicerie. Il s'agit d'un monde où P & non-Q [i.e. un monde où $W = w2$ mais $Y = y1$]. De plus, w semble beaucoup plus proche du monde actuel que ne l'est un monde dans lequel John ne croit pas qu'il y a de la bière dans le frigo et se dirige, en conséquence, vers l'épicerie plutôt que vers le frigo. (*ibid.*, 9, trad. A. M.)

En d'autres termes, dans le monde possible le plus proche dans lequel la valeur de W est modifiée de $w1$ à $w2$, la valeur de Y ne change pas, et le contrefactuel (2) est donc faux. La fausseté de (2), combinée avec la vérité de (1), implique donc que X est pertinent causalement vis-à-vis de Y , et ceci de manière non-redondante, c'est-à-dire que c'est la croyance de John, plutôt que l'état cérébral réalisant cette croyance, qui a causé son comportement. La conclusion de Raatikainen est que, dans la mesure où elle permet à un état mental d'être une cause non-redondante d'un état physique tel que le comportement de John, la théorie interventionniste montre qu'il peut y avoir des cas de causalité « descendante ». Et la possibilité de ce type de cas permet de résoudre le problème de l'exclusion causale de Kim : un état mental peut être pertinent causalement sans être identique à l'état physique qui le réalise. Bien qu'à première vue convaincant, l'argument de Raatikainen n'est en pas moins incorrect.

4 POURQUOI L'INTERVENTIONNISME NE PERMET PAS LA CAUSALITÉ « DESCENDANTE »

Si l'argument de Raatikainen semble à première vue convaincant, il se heurte néanmoins à une objection décisive. Mais avant de se tourner vers cette objection, il convient de clarifier la conclusion que Raatikainen entend tirer de son argument.

4.1 QUEL CONCEPT DE NIVEAU ?

Si j'utilise ici des guillemets pour le terme « descendante » dans l'expression « causalité "descendante" », c'est parce que déterminer le concept de niveau que Raatikainen utilise est loin d'être évident, et ceci rend la conclusion qu'il tire de son argument ambiguë. Il conçoit sa conclusion comme établissant l'existence de cas de causalité « descendante ». Mais est-il correct de dire que le comportement de John est situé à un niveau inférieur à celui auquel est située sa croyance ? Si l'on adopte, par exemple, un concept météorologique de niveau, alors pour que l'on ait affaire à un cas de causalité « descendante », il faut que le comportement de John soit une propriété d'une partie propre de

l'objet instanciant la propriété mentale correspondant à la croyance de John. On pourrait imaginer soutenir que son comportement est une propriété du corps de John alors que sa croyance est une propriété de la personne qu'est John. Mais il semble étrange de soutenir que le corps de John soit une partie *méréologique* de sa personne. Selon un tel concept de niveau, on pourrait également imaginer soutenir que la croyance de John, étant une propriété de son cerveau (ou de son système nerveux), est située à un niveau inférieur à son comportement, puisque le comportement est une propriété du corps de John dans son ensemble, y compris son cerveau. Il me semble que, si l'on adopte un concept strictement méréologique de niveau, la solution la plus raisonnable soit de voir la croyance de John et son comportement comme tous deux situés au même niveau.

Une autre possibilité est que Raatikainen utilise l'expression « descendante » pour qualifier les relations causales entre propriétés survenantes et propriétés de base. Bien que Kim lui-même insiste sur le fait que les relations de survenance ne sont pas des relations inter-niveaux, cela semble être la manière dont la plupart des philosophes impliqués dans le débat sur la causalité mentale utilisent le terme. Mais encore une fois, on peut se demander s'il est vrai que, selon ce concept de niveau « de survenance », le comportement de John est situé à niveau inférieur à celui auquel est située sa croyance. La thèse de survenance la plus souvent acceptée est celle de la survenance des états mentaux d'un individu sur ses états microphysiques. Mais le comportement de John ne fait pas partie de ses états microphysiques. Et l'on pourrait même soutenir que le comportement de John, tout comme sa croyance, survient sur ses états microphysiques, et est tout autant multiplement réalisable que l'est sa croyance. Mais si c'est le cas, alors ce que montre l'argument de Raatikainen, c'est qu'il peut y avoir des relations causales entre propriétés survenantes, et non pas entre une propriété survenante et une propriété de la base. Si telle est la conclusion que Raatikainen souhaite tirer de son argument, alors son argument est essentiellement identique à celui avancé par Peter Menzies et Christian List (Menzies & List, à *paraître*) et qui prétend résoudre le problème de l'exclusion causale en montrant qu'il peut exister des relations causales entre propriétés survenantes qui soient autonomes vis-à-vis des relations causales entre propriétés de base. L'interprétation la plus charitable de Raatikainen me semble être celle menant à voir son argument comme prétendant montrer qu'il peut exister des relations causales entre propriétés survenantes, par exemple la croyance de John qu'il y a de la bière dans le frigo, et propriétés de base, par exemple la propriété microphysique réalisant le comportement de John consistant à se diriger vers le frigo pour y chercher une bière. La situation est représentée par le diagramme suivant (figure 3, page suivante).

Dans ce diagramme, Z représente la propriété de base sur laquelle survient le comportement de John, et cette variable peut avoir les deux valeurs suivantes :

$Z = z_1$ lorsque John est dans l'état microphysique S et $Z = z_2$ quand ce n'est pas le cas.

Maintenant que la conclusion que Raatikainen tire de son argument a été clarifiée, tournons-nous à présent vers l'objection la plus importante à laquelle cet argument se heurte.

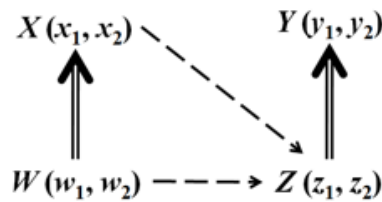


Fig. 3

4.2 INTERVENTIONS ET CAUSALITÉ « DESCENDANTE » ⁷

L'objection principale contre l'argument de Raatikainen est que, lorsqu'elle est correctement appliquée, la théorie interventionniste interdit les cas de causalité « descendante » plutôt qu'elle ne les permet. Pour qu'il existe une relation causale « descendante » entre X et Z , il est nécessaire qu'il soit possible d'intervenir sur X par rapport à Z . Or, en supposant comme le font les physicalistes non-réductifs, qu'il existe une relation de survenance entre X et W , et qu'il existe une relation causale entre W et Z , on peut aisément montrer qu'une intervention sur X par rapport à Z est impossible. Il est important de noter que cette objection ne suppose pas la vérité de la thèse de la clôture causale du monde physique, mais suppose simplement la pertinence causale de W vis-à-vis de Z . La question qui se pose est celle de savoir si la façon dont la croyance de John est réalisée influence la façon dont son comportement est réalisé. Il ne me semble pas y avoir de raison *a priori* de nier que ce soit le cas. La charge de la preuve incombe donc au physicaliste non-réductif qui souhaiterait rejeter mon argument en niant cette prémisse. Les conditions I1-I4 qu'une manipulation d'une variable doit satisfaire pour être une authentique intervention ont été indiquées plus haut (§2). Raatikainen affirme que le fait que Peter dise à John qu'il n'y a plus de bière constitue une intervention en bonne et due forme. Représentons cette manipulation par la variable I , cette variable pouvant avoir deux valeurs : $I = i_1$ lorsque Peter ne dit pas à John qu'il n'y a plus de bière et $I = i_2$ lorsque Peter dit à John qu'il n'y a plus de bière. La situation est représentée dans le diagramme suivant (figure 4) :

7. Un argument similaire a été indépendamment développé par Michael Baumgartner dans (Baumgartner, 2009)

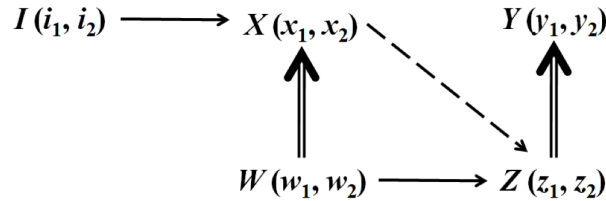


Fig. 4

La question qui se pose est de savoir si les conditions I1-I4 peuvent être satisfaites par I . Dans la mesure où il existe un chemin dirigé entre I et X , et puisque ce chemin est le seul qui aboutisse à X , les conditions I1 et I2 sont satisfaites par I . La condition I3 est également satisfaite, puisque le seul chemin dirigé entre I et Z est le chemin passant par X , c'est-à-dire I - X - Z . Qu'en est-il de la condition I4 ? J'ai supposé que W était pertinent causalement vis-à-vis de Z . Or, W ne se situe pas sur un chemin dirigé passant par X . La condition I4 requiert donc que I et W soient statistiquement indépendants. Considérons le cas dans lequel I est une cause déterministe de X ⁸. Dans ce cas, lorsque $I1 = i1$, alors $P(X = x1) = 1$. En d'autres termes, dans la situation considérée, John croit systématiquement qu'il y a de la bière dans le frigo à moins que Peter ne lui dise que ce n'est pas le cas. Supposons que lorsque John croit qu'il y a de la bière dans le frigo, cette croyance est réalisée par B avec une probabilité de .75 et un autre état cérébral B^* avec une probabilité de .25. Lorsque $I = i1$, on a donc $P(X = x1) = 1$ et $P(W = w1) = .75$. Que se passe-t-il lorsque X prend une autre valeur que $x1$, par exemple $x2$? Et ce changement dans la valeur de X modifie-t-il la distribution de probabilité sur W ? Il semble évident que si Peter révèle à John qu'il n'y a plus de bière, alors John arrêtera de croire qu'il y a de la bière dans le frigo. En tout état de cause, si $I = i2$ alors $P(X = x1) < 1$. En d'autres termes, la probabilité que John continue à croire qu'il y a de la bière dans le frigo étant donnée l'information contradictoire fournie par Peter sera strictement inférieure à 1. Mais si c'est le cas, alors lorsque $I = i2$, $P(W = w1) < .75$ et la valeur prise par I modifie donc la distribution de probabilité sur W . I et W ne sont donc pas statistiquement indépendants, la condition I4 n'est pas satisfaite, et I n'est donc pas une authentique intervention sur X vis-à-vis de Z . Mais s'il n'est pas possible d'intervenir sur X vis-à-vis de W , alors X ne peut pas être dit pertinent causalement vis-à-vis de Z ⁹.

8. Le même argument peut être formulé dans le cas où I n'est pas une cause déterministe de X .

9. Un physicaliste non-réductif désirant développer l'idée de causalité « descendante » dans le cadre interventionniste pourrait être tenté de se débarrasser la condition I4. Il existe cependant de très bonnes raisons, liées au principe de la cause commune et aux méthodes de l'inférence causale, pour conserver une condition telle que I4 dans une théorie de la causalité.

5 CONCLUSION

Malgré le fait qu'il soit à première vue convaincant, l'argument proposé par Raatikainen repose donc sur une application illégitime de la théorie interventionniste de la causalité au physicalisme non-réductif. Loin de permettre la causalité « descendante », l'approche interventionniste la rend au contraire impossible. Adopter la théorie interventionniste bloque donc ce qui est souvent conçu comme la principale solution au problème de l'exclusion causale pour les tenants du physicalisme non-réductif. La leçon à tirer me semble être la suivante : les physicalistes non-réductifs ont tort de situer la force de l'argument de l'exclusion causale exclusivement dans l'adoption par Kim d'une conception nomologique de la causalité. Il ne suffit pas d'adopter la théorie interventionniste de la causalité pour résoudre le problème de la causalité mentale ¹⁰.

BIBLIOGRAPHIE

- Baumgartner, M. (2009), « Interventionist Causal Exclusion and Non-reductive Physicalism », *International Studies in the Philosophy of Science*, 23, 161-178.
- Kim, J. (1998), *Mind in a Physical World : An Essay on the Mind-Body Problem and Mental Causation*, Cambridge, MA : MIT Press.
- Kistler, M. (à paraître), « Mechanisms and Downward Causation » *Philosophical Psychology*.
- Lewis, D. (1973), « Causation » *Journal of Philosophy*, 70, 556-567.
- Lewis, D. (2000), « Causation as influence » dans Collins J., Hall N., & Paul (L. A.) (éds.), *Causation and Counterfactuals*, Cambridge, MIT Press.
- Menzies, P., & List, C. (à paraître), « The Causal Autonomy of the Special Sciences » dans Macdonald C. & Macdonald G. (éds.), *Emergence and Causation*, Oxford, Oxford University Press.
- Northcott, R. (2008), « Causation and contrast classes », *Philosophical Studies*, 139, 111-123.
- Raatikainen, P. (à paraître), « Causation, Exclusion, and the Special Sciences », *Erkenntnis*.
- Shapiro, L. (à paraître), « Lessons from Causal Exclusion », *Philosophy & Phenomenological Research*.
- Shapiro, L., & Sober, E. (2007), « Epiphenomenalism – the Do's and the Don'ts » dans Wolters G. & Machamer P. (éds.), *Thinking about Causes : From*

10. Des versions antérieures de cet article ont été présentées à l'université d'Édimbourg, au congrès de la SOPHA à Genève, et à l'université de Nottingham. Je remercie en particulier Sven Walter, Jonas Christensen, Jim Woodward, Cyrille Imbert, Pascal Ludwig et Daniel Andler pour leurs commentaires et critiques. Je suis particulièrement reconnaissant envers Michael Baumgartner pour ses multiples commentaires et encouragements.

Greek Philosophy to Modern Physics, Pittsburgh, Pittsburgh University Press, 235-264.

Weslake, B. (Ms.), *Exclusion Excluded*.

Woodward, J. (2003), *Making Things Happen : A Theory of Causal Explanation*, Oxford, Oxford University Press.

Woodward, J. (2008), « Mental Causation and Neural Mechanisms » dans Hohwy J. & Kallestrup J. (éds.), *Being Reduced : New Essays on Reduction, Explanation and Causation*, Oxford, Oxford University Press.