

# CAHIERS D'ÉPISTÉMOLOGIE

Publication du *Groupe de Recherche en Épistémologie Comparée*  
Directeur: Robert Nadeau  
Département de philosophie, Université du Québec à Montréal

**Temps et rationalité selon Jean-Pierre Dupuy :**  
Critique et solution de rechange

**Michel B. Robillard**

**Cahier n° 2009**

**272<sup>e</sup> numéro**

**UQÀM**

<http://www.philo.uqam.ca>

**Cette publication, la deux cent soixante-douzième de la série, a été rendue possible grâce à la contribution financière du *Fonds pour la Formation de Chercheurs et l'Aide à la Recherche* du Québec ainsi que du *Programme d'Aide à la Recherche et à la Création* de l'UQAM.**

**Aucune partie de cette publication ne peut être conservée dans un système de recherche documentaire, traduite ou reproduite sous quelque forme que ce soit - imprimé, procédé photomécanique, microfilm, microfiche ou tout autre moyen - sans la permission écrite de l'éditeur. Tous droits réservés pour tous pays./ All rights reserved. No part of this publication covered by the copyrights hereon may be reproduced or used in any form or by any means - graphic, electronic or mechanical - without the prior written permission of the publisher.**

**Dépôt légal – 4<sup>e</sup> trimestre 2000  
Bibliothèque Nationale du Québec  
Bibliothèque Nationale du Canada  
ISSN 0228-7080  
ISBN: 2-89449-072-0**

**© 2000 Michel B. Robillard**

**Temps et rationalité selon Jean-Pierre Dupuy :**

Critique et solution de rechange

**Michel B. Robillard**

**Département de philosophie  
Université du Québec à Montréal  
Case postale 8888, succursale Centre-Ville  
Montréal (Québec) Canada H3C3P8**

**[michelrobillard@sympatico.ca](mailto:michelrobillard@sympatico.ca)**

Je suis très reconnaissant à mon directeur de recherche, le Professeur Robert Nadeau, pour son soutien, ses conseils judicieux et ses commentaires pertinents.

Il y a une trentaine d'années, Robert Nozick (1969) publiait un paradoxe qui a fait couler et qui continue à faire couler beaucoup d'encre chez les théoriciens de la décision : le paradoxe de Newcomb. Voici rapidement ce qu'est ce paradoxe : il y a deux boîtes, la boîte transparente A contenant 1000\$ et la boîte opaque B, laquelle *pourrait* contenir 1000000\$, et vous avez à choisir ou bien la seule boîte B ou bien les deux boîtes. On vous dit qu'un prédicteur infallible a placé le million de dollars dans la boîte B si, et seulement si, il a prédit que vous choisiriez la boîte B, et qu'il l'a laissé vide s'il a prédit que vous choisiriez les deux boîtes. Que faites-vous? Un nombre incroyable d'arguments ont été proposés en faveur de l'un ou l'autre de ces choix, mais ce que le paradoxe de Newcomb fait ressortir, ce sont les tensions au sein même de la rationalité dans la théorie de la décision. Quel est le choix le plus rationnel? Dans ce travail, nous étudierons un texte de Jean-Pierre Dupuy (1994) sur la rationalité en théorie de la décision. Nous verrons que la théorie de la décision comporte plusieurs enjeux, entraînant notamment les débats entre les causalistes et les évidentialistes et entre le compatibilisme et l'incompatibilisme du déterminisme et du libre-arbitre. Dans ce texte, Dupuy propose une nouvelle solution à ces enjeux; solution aussi compliquée qu'irrecevable pour plusieurs raisons. Cependant, comme il est nécessaire de bien comprendre ce dont on parle, et ce dont l'autre parle, lorsqu'on le critique, il nous faudra reconstruire l'argument de Dupuy pour ensuite attaquer ses points faibles. Nous verrons donc certaines irrationalités rattachées au choix évidentialiste, les problèmes de Newcomb avec cause commune, la mauvaise foi et le problème de la *backwards induction*.

### **Les irrationalités**

Dès le début de son texte, Dupuy examine quelques violations du principe métaphysique de fixité du passé par rapport à l'action libre. Notre étude de certaines de ces irrationalités ne sera que partielle et servira simplement à mettre en place les éléments du travail. Cependant, nous verrons en profondeur les problèmes de Newcomb avec cause commune et la mauvaise foi (*self-deception*). S'il est important de faire une telle démarche, c'est que la théorie de Dupuy repose sur l'idée que le principe de fixité du passé de Maurice Allais n'est pas universellement valide du point de vue normatif (de la raison). Donc, Dupuy, en montrant que le causalisme n'est pas universellement valide, tente de sauver l'évidentialisme, mais pour cela, il doit tout d'abord identifier les différents problèmes avec la rationalité qui lui est rattachée. Comme nous le disions, la formulation de ce principe, telle qu'utilisée par Dupuy, revient à Maurice Allais : « En matière de rationalité, la

maxime fondamentale est : seul compte l'avenir. » C'est-à-dire que le passé est ce qu'il est et que nous n'y pouvons rien, et que seules comptent les conséquences futures prévisibles de nos décisions ou actions. Bref, cette maxime n'est ni plus ni moins qu'une formulation de ce qu'on appelle le conséquentialisme en théorie du choix rationnel, c'est-à-dire une théorie des conséquences causales de nos décisions ou actions.

La première forme d'irrationalité s'intéresse à l'imitation. En effet, il est généralement admis, sauf dans les marchés financiers et la spéculation<sup>1</sup>, que la rationalité implique, entre autres, une autonomie des préférences et une absence d'influences mutuelles entre les agents. Par conséquent, quelqu'un qui imite le comportement d'autres est irrationnel. De façon à illustrer l'imitation et à faire le lien avec le second type d'irrationalité, reprenons l'exemple utilisé par Dupuy: « J'observe qu'un certain club privé à une cote élevée dans mon entourage. Pas mal de gens sont prêts à encourir des coûts non négligeables de toutes sortes pour s'y faire admettre. J'en infère que c'est là un objectif hautement désirable. Je me le fixe à moi-même et m'efforce de l'atteindre. » (Dupuy, 1994, p.70) Comme nous venons de le voir, il semble qu'un tel comportement soit irrationnel du fait qu'il est une imitation. Supposons maintenant, comme le fait Dupuy, qu'une fois mon objectif atteint, c'est-à-dire être admis dans le club, je me rende compte qu'il est banal et qu'il me déplaît au plus haut point. Cependant, en considérant qu'il m'en a coûté très cher pour y être admis, je me convaincs qu'il n'est peut-être pas si mal après tout. Ce qu'il y a d'irrationnel dans ce raisonnement, c'est le fait que le coût intervienne dans l'évaluation de la désirabilité de l'admission au club, alors qu'en fait les coûts que nous sommes prêts à payer pour obtenir quelque chose sont des signes de leur désirabilité, ils sont en quelques sortes des mesures de désirabilité. Tversky et Quattrone expliquent une telle irrationalité en affirmant qu'il s'agit d'une inversion du lien causal entre coût et désirabilité, ou encore, entre une problématique causale et une problématique d'effets de signes (ces deux notions sont importantes, car elles sont au cœur du débat que nous étudierons, le débat causalisme-évidentialisme). Il existe, selon Dupuy, une autre façon d'interpréter cette irrationalité de comportement. On peut dire que le fait que les autres aient encouru des coûts

---

<sup>1</sup>. Comme le fait remarquer Dupuy, dans ces deux cas, il est possible de raisonner ainsi : les autres disposent peut-être d'informations qui ne me sont pas accessibles et les imiter me permet de tirer parti de ce savoir. Évidemment, il se peut qu'ils soient aussi ignorants que moi, mais si tel est le cas, je n'ai rien à perdre en supposant qu'ils savent quelque chose. De plus, selon Dupuy, il semble que l'on puisse faire un raisonnement bayésien pour étayer cet argument.

importants pour être admis au club me donne une raison de croire que cela est désirable, sans cependant donner une raison au fait que j'aie cru cela. Bref, un peu à l'image de ce que Davidson affirme, ma croyance passée (la désirabilité d'une chose parce que les autres croient qu'elle l'est) ne constitue pas une raison pour ma croyance présente (le fait que je la désire) si j'ai par ailleurs une bonne et décisive raison de l'abandonner (le désenchantement face à cette chose pour laquelle j'ai encouru des coûts), et ce, même si elle peut en être la cause.

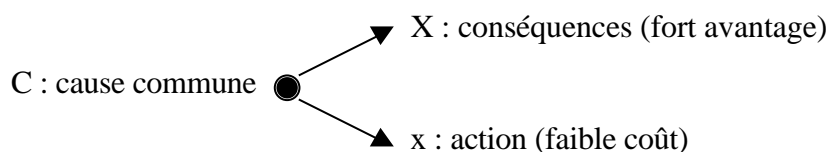
La troisième irrationalité rejoint celle dont nous venons de discuter, à ceci près : elle ne concerne pas la formation de la croyance, mais l'action elle-même ou comme le dit Dupuy «l'action au service de la croyance » (1994, p.73). Prenons un exemple très contemporain : il est très à la mode de nos jours de prendre des abonnements à des centres sportifs (des "gyms"). Donc, en jugeant moi-même de manière rationnelle de la désirabilité d'être en bonne santé, je m'abonne à grands coûts. Horreur! Je déteste faire de l'entraînement physique. Que dois-je faire ? Selon la théorie du choix rationnel, les dépenses doivent être considérées comme amorties au moment de leur effectuation, ce qui veut dire que le fait que j'aie déboursé une somme considérable pour m'abonner ne devrait plus entrer en ligne de compte à partir du moment où l'achat est fait. En d'autres mots, le passé est ce qu'il est et il faut désormais que je ne tienne compte que des conséquences futures de mes décisions ou actions (principe d'Allais). Dans cette optique, je devrais agir comme si l'abonnement m'avait été donné et prendre une décision en fonction des coûts futurs (pas nécessairement monétaires) et des avantages de toutes les options à ma disposition. Donc, en considérant que je n'aime pas m'entraîner, que je devrai me déplacer fréquemment et que le fait de faire quelque chose que je n'aime pas risque de miner mon moral, je pourrais conclure qu'il est préférable d'assumer les pertes et de juger que mon action de départ était irrationnelle. Or, la plupart d'entre nous ne raisonnons pas ainsi. En effet, cet abonnement m'a coûté cher et je dois l'utiliser, ce qui me permettra d'amortir ce coût (cette fois-ci, l'amortissement se fait après coup), et ce, même si d'autres options seraient préférables. C'est ce que nous appelons le sophisme de l'amortissement (*sunk cost fallacy*). Comme le fait remarquer Dupuy, nous sommes toujours dans la logique de l'imitation, mais plutôt que de copier les croyances des autres, ou mes propres croyances comme si elles étaient celles des autres, je copie un comportement qui serait mien si la situation avait été différente, si mes espérances s'étaient avérées justes. Encore une fois, si l'on se

---

fié à Tversky et Quattrone, il s'agit d'une confusion entre une problématique causale et une problématique en termes d'effets de signes. En commettant le sophisme de l'amortissement, mon action ne vise pas ses conséquences causales, mais l'évaluation qu'elle me permet de faire de la situation. J'essaie, d'une certaine façon, de me rassurer vis-à-vis ma rationalité en me disant que j'ai fait le bon choix en prenant cet abonnement, car je vais fréquemment au gymnase.

### **Les problèmes de Newcomb avec cause commune**

Maintenant que nous connaissons les bases du débat causalistes-évidentialistes, et que nous sommes en mesure de différencier les deux types de raisonnements (bien que ce ne soit encore que de façon grossière), nous sommes prêts à nous attaquer à des questions qui sont au cœur des discussions entre les partisans de chaque théorie. Ce que nous appelons les problèmes de Newcomb, ce sont toutes les situations qui impliquent une dépendance probabiliste et une indépendance causale entre l'action et un état du monde. En d'autres mots, il y a une forte dépendance probabiliste entre l'action et l'état du monde, sans toutefois que cette action soit la cause de cet état du monde, d'où l'indépendance causale. Prenons un exemple classique venant d'un statisticien américain du nom de R.A. Fisher<sup>2</sup>. Notons tout d'abord, que l'exemple de Fisher se situe dans une sous-classe de problèmes de Newcomb appelée problèmes de Newcomb avec cause commune (certains ne font pas cette différenciation) et que tout problème de cette sous-classe se présente ainsi : « un état du monde C, cause, tout à la fois, d'un état de choses très favorable X et d'une décision modérément coûteuse x » (Dupuy, 1994, p.76). Voici un schéma de la façon dont les problèmes de ce type se présentent :



On voit donc très bien qu'il n'y a pas de lien causal entre l'action x et les conséquences X. Notons aussi que non-C cause non-X et non-x, et que le coût de l'action x est relatif à l'avantage X.

---

<sup>2</sup>. Bien que l'idée de Fisher n'ait pas obtenu un grand succès dans le milieu médical, il n'en demeure pas moins qu'elle est aujourd'hui considérée comme un cas paradigmatique en théorie du choix rationnel.

Voyons l'exemple de Fisher. Nous savons qu'il existe une forte corrélation entre le fait de fumer et le cancer du poumon. Supposons que, contrairement à ce qu'on croyait, le cancer du poumon (la conséquence, non-X) n'est pas causé par la consommation du tabac (l'action, non-x), mais que la cause du cancer soit un gène (la cause commune, non-C) qui prédispose les gens à fumer. Est-ce qu'une personne qui éprouve un grand plaisir à fumer devrait désormais s'abstenir de le faire, même en sachant que sa décision n'aura aucune incidence sur le fait qu'elle aie ou non le gène? Vous admettez que s'arrêter de fumer de peur d'avoir le cancer est maintenant quelque chose qui semble irrationnel à première vue. Or, un problème surgit lorsque l'on applique la règle d'or de la théorie du choix rationnel : agis toujours de façon à maximiser ton espérance mathématique d'utilité. Selon cette règle, il serait rationnel de s'arrêter de fumer, puisque les probabilités conditionnelles (des états du monde par rapport à l'action) sont indifférentes au fait que ces liens soient causaux ou non. Mais puisque le premier raisonnement semble être le plus convaincant pour la plupart des auteurs contemporains, et puisque que nous savons que les liens entre les états du monde et l'action ne sont pas causaux dans un problème de Newcomb avec cause commune, il faut se demander ce qui ne va pas avec la théorie classique du choix rationnel. L'orthodoxie propose d'utiliser une règle qui, en bout de ligne, n'est ni plus ni moins que le principe d'Allais : peu importe ce que le sujet fera, il ne peut rien changer au fait passé de la présence ou de l'absence du gène. Ceux qui s'opposent à cette orthodoxie ont recours à une théorie évidentialiste de la décision, laquelle théorie est jugée inadéquate par les orthodoxes. En effet, les évidentialistes raisonnent comme si le fait d'arrêter de fumer était le prix à payer pour obtenir le signe de l'absence du gène. Voyons comment Dupuy résume l'argument évidentialiste, tel que vue par les orthodoxes :

« Lorsque l'agent maximise son espérance d'utilité de la façon habituelle, tout se passe comme s'il prenait sur lui-même un point de vue extérieur, ainsi qu'il le ferait sur quelqu'un d'autre, et découvrirait avec nous sa propre décision, sa stratégie consistant alors à agir de telle sorte à rendre aussi bonne que possible la nouvelle que cette découverte lui apporte. Si je faisais ceci, raisonne le sujet, mon action manifesterait la présence de tel trait désirable dans le monde. Je souhaite que ce soit le cas et j'agis donc ainsi. » (1994, p.76)

Bref, les sujets évidentialistes semblent agir comme s'ils avaient un pouvoir quelconque sur le passé, le pouvoir de le choisir, tandis que les tenants de la théorie causale ne s'intéressent qu'aux



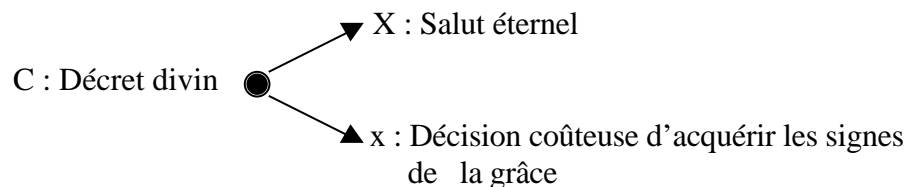
conséquences causales de leurs actions ainsi qu'à leur désirabilité.

Ce que Dupuy se propose de faire, c'est une défense et une illustration de l'évidentialisme, sans cependant que ce soit contre l'orthodoxie. Pour ce faire, il n'a d'autre choix que de supposer l'existence de deux formes de rationalités irréductibles l'une à l'autre, une rationalité causaliste et une rationalité évidentialiste. De plus, il admet que ces deux formes de rationalités sont rattachées à « deux conceptions du temps différentes bien qu'inséparables ». On se doute bien qu'il aura beaucoup à faire pour nous convaincre que (i) il existe deux formes de rationalités humaines, (ii) nous avons deux *expériences* du temps différentes et (iii) que la théorie évidentialiste est justifiable même si elle est associée à toutes les irrationalités que nous avons vues jusqu'à maintenant.

Avant d'avancer ses arguments, Dupuy nous présente un cas paradigmatique de problème de Newcomb avec cause commune, l'exemple de Max Weber sur les conséquences éthiques de la doctrine de la prédestination, et l'« esprit du capitalisme ». Il est important de noter ici que Dupuy, comme il le fait remarquer, ne s'intéresse qu'à la forme logique de l'argument et non pas à sa validité empirique. Voici l'exemple tel que tiré du texte de Dupuy (1994, p.77-78) :

« En vertu d'une décision divine prise de toute éternité, chacun appartient à un camp, celui des élus ou celui des damnés, sans savoir lequel. Les hommes ne peuvent absolument rien à ce décret. Il n'y a rien qu'ils puissent faire pour gagner ou mériter leur salut. La grâce divine, cependant, se manifeste par des *signes*. La chose importante, ici, est que ces signes ne s'observent pas par introspection, ils *s'acquièrent par l'action*. Le principal d'entre eux est le succès que l'on obtient en *mettant à l'épreuve* sa foi dans une activité professionnelle (*Beruf*). Cette épreuve est coûteuse, elle exige de travailler sans relâche, méthodiquement, sans jamais se reposer dans la possession, sans jamais jouir de la richesse [...] La répugnance au travail est le symptôme d'une absence de la grâce. »

La structure logique du problème est la suivante :



Du point de vue de l'orthodoxie, le choix rationnel serait le fatalisme, puisque quoi qu'un individu fasse, il ne peut rien changer au passé, c'est-à-dire au fait qu'il soit ou non élu, il est donc préférable qu'il vive dans l'oisiveté. Selon la doctrine calviniste, chacun devait se considérer comme élu et tout doute à ce sujet consistait en une confiance en soi déficiente, et le seul moyen d'obtenir cette confiance était le travail sans relâche dans un métier. Donc, le travail assurait les gens d'être parmi les élus. Le problème avec ce type de raisonnement, nous l'avons vu, est que les Calvinistes agissaient comme si  $x$  était la cause de  $X$ , ce qui est irrationnel, car cela consiste à prendre le signe  $x$  pour la chose  $C$ . Or, Dupuy nous met en garde contre l'étiquetage trop rapide de la théorie évidentialiste, car il semble, selon Weber, que l'évidentialisme soit à la base du rationalisme économique moderne, lequel est considéré comme le summum de la Raison dans l'histoire. Par conséquent, rejeter l'un consisterait, en quelque sorte, à rejeter l'autre.

### **Déterminisme, compatibilisme et mauvaise foi**

Avant d'en arriver à l'argument de Dupuy, voyons comment la mauvaise foi (*self-deception*) joue un rôle important dans le débat, en tant qu'elle est le summum de l'incohérence personnelle. La question que se pose Dupuy est de savoir si effectivement les Calvinistes se mentaient à eux-mêmes, et il entend montrer que non, de façon à faire apparaître la rationalité du choix évidentialiste. Notons que Tversky et Quattrone ont proposé une interprétation acceptable du choix évidentialiste. Cependant, puisque cette interprétation s'en tient à la mauvaise foi, elle ne permet pas, selon Dupuy, de faire ressortir la rationalité du choix. Nous reviendrons à cette interprétation à la fin de la présente section, car elle sera la base de ma solution pour sauver la rationalité du choix évidentialiste. Selon Dupuy, les Calvinistes ont les deux croyances suivantes (lesquelles, selon lui, ne sont pas nécessairement incompatibles) :

- (1) Les Calvinistes croient qu'ils n'ont pas procédé eux-mêmes à leur élection, parce qu'ils croient que c'est Dieu qui les a élus;
- (2) Les Calvinistes croient qu'ils étaient libres de choisir  $x$  ou Non- $x$  lorsqu'ils ont choisi  $x$ .

Il faut donc résoudre le problème de l'incohérence entre le déterminisme de (1) et le libre-arbitre de (2). Bien que (2) ne fasse pas partie des données du problème, Dupuy nous demande

« quel sens cela aurait-il de s'interroger sur la rationalité de l'agent si celui-ci n'était pas doté, et ne se dotait pas lui-même, de la faculté de libre-arbitre, entendue dans son sens minimal de pouvoir agir autrement qu'on ne le fait? » (Dupuy, 1994, p.81) Dupuy nous propose alors deux manières de défendre le compatibilisme suivant le degré de la menace que le déterminisme fait peser sur le libre-arbitre. Nous verrons que les deux formes de rationalité et de temporalité qui leur sont associées s'enracinent, selon Dupuy, dans la distinction entre deux conceptions du déterminisme.

Voyons tout d'abord ce qu'est la théorie incompatibiliste et ce qu'elle implique. C est un événement se produisant à  $t_1$ , x est l'action en  $t_2$  que C annonce, S désigne le sujet et  $\Box$   $S_t$  est l'opérateur logique de nécessité tel que:  $\Box_t (p)$  signifie que p est vraie et S n'est pas libre en t de faire une action qui rendrait p fausse. Nous avons donc l'argument incompatibiliste suivant :

- N1 :  $\Box S_{t_2}$  (C s'est produit à  $t_1$ )  
 N2 :  $\Box S_{t_2}$  (Si C s'est produit à  $t_1$ , alors S fait x à  $t_2$ )  
 D'où N3 :  $\Box S_{t_2}$  (S fait x à  $t_2$ )

Il est important de noter, ainsi que le fait Dupuy (1994, p.83-84), que N1 exprime le principe de fixité du passé (elle est nécessaire de façon contingente) et que N2 est une proposition exprimant la fixité de la chaîne temporelle qui relie C à x, c'est-à-dire qu'elle donne à C le rôle annonciateur de x. Notons aussi, pour bien suivre le raisonnement de Dupuy, que N2 est équivalente à :

$$(3) \quad (\text{Si C s'est produit à } t_1) < (\text{S fait x à } t_2)$$

où < représente l'implication stricte. Par conséquent, (3) ne fait qu'exprimer le fait que dans tous les mondes possibles où C s'est produit à  $t_1$ , S fait x à  $t_2$ . Il est extrêmement important ici de bien comprendre que le caractère nécessaire de N2 est essentiel. En effet, une proposition comme :

$$A2 : \quad \text{Si C s'est produit à } t_1, \text{ alors S fait x à } t_2$$

ne saurait exclure le fait qu'il soit possible dans un monde autre que le nôtre que même si C s'est produit à  $t_1$ , que S ne fasse pas x à  $t_2$ . C'est là, selon Dupuy, qu'origine la bifurcation entre les deux formes de temporalité qu'il utilise dans son argument. Notons finalement que N3 ne fait qu'exprimer le fait que S fait effectivement x à  $t_2$ , tel qu'annoncé en  $t_1$ , mais que S n'a pas d'autre choix (il est nécessaire que S fasse x en  $t_2$  étant donné N1 et N2).

Dupuy affirme qu'il y a deux possibilités de contrer cet argument. Soit l'on admet N1 et l'on se doit de rejeter N2, soit l'on admet N2 et l'on se doit de rejeter N1. La seconde solution est celle qu'utilise Dupuy, et c'est sur cette dernière que nous nous concentrerons. Disons seulement que la première solution correspond à l'orthodoxie et qu'elle repose sur le principe de *fixité du passé par rapport à l'action présente* (principe d'Allais). Donc, il va de soi que Dupuy se propose de démontrer que le passé n'est pas nécessairement fixe et qu'il doit être possible de le changer. Évidemment, le pouvoir que possède l'agent sur le passé (si une telle chose est possible) ne peut être un pouvoir causal. Selon Dupuy, il s'agit d'un pouvoir contrefactuel d'invalider le passé qui n'est possible que si l'on suppose que l'agent est libre d'agir. Voici le raisonnement que l'agent doit faire selon Dupuy :

« Je suppose que C a eu lieu à  $t_1$  et, donc, que je vais décider de faire x à  $t_2$ . Bien que faisant x à  $t_2$ , je sais qu'il est dans mon pouvoir de faire Non-x. Il est donc dans mon pouvoir à  $t_2$  de faire quelque chose (à savoir Non-x) tel que, *si je le faisais*, C ne se serait pas produit à  $t_1$  puisque le lien entre le passé et mon action étant fixe, si C s'était produit à  $t_1$ , c'est x que je ferais. En agissant autrement que je le fais, j'accéderais à un autre monde *possible* (puisque je suis libre), dans lequel le passé serait différent de ce qu'il est dans notre monde. » (p.85)

Bien que cela ne soit que le début de l'argument de Dupuy, une première critique s'impose. Au plan logique, le raisonnement ci-dessus est impeccable. Cependant, je crois qu'il ne s'agit ni plus ni moins que d'une belle machine à résoudre les problèmes de Newcomb avec cause commune : on entre le problème dans la machine et une solution évidentialiste est produite de façon (supposément) rationnelle. Or, bien qu'elle soit bien roulée, je crois que cette machine est tout simplement superflue et qu'elle a été construite de toutes pièces pour satisfaire les thèses de Dupuy. Prenons une analogie : imaginez que vous faites la visite d'une usine de pâtes et papier et qu'au bout de la ligne de production vous apercevez une énorme machine. Vous interrogez alors le

contremaître et lui demandez à quoi elle sert et comment elle fonctionne. Sa réponse est la suivante : « Au prix que nous avons payé pour l'acquérir, personne ne l'a encore ouverte, de peur de l'endommager, mais je sais qu'elle peut couper une feuille de papier au format qu'on lui indique en une heure. » Vous lui demandez alors pourquoi ils n'utilisent pas un simple travailleur muni de ciseaux qui pourrait facilement faire le même travail, à moindre coût et beaucoup plus rapidement. La seule chose que le contremaître vous répond alors est : « Ça fonctionne, non? » L'argument de Dupuy possède le même problème que cette machine : il fonctionne selon les données qu'on lui fournit (c'est-à-dire des problèmes de Newcomb correspondant avec les thèses de Dupuy), mais il est extrêmement coûteux au plan ontologique. En effet, il oblige Dupuy à supposer deux types de rationalité humaine, deux expériences de la temporalité et un pouvoir contrefactuel sur le passé. L'argument que je vais maintenant proposer vise à démontrer que l'appareillage conçu par Dupuy pour rendre rationnel le choix évidentialiste n'est pas nécessaire et qu'il est facilement remplaçable à moindre coût ontologique.

La vie de l'agent se déroule sans qu'il ait conscience de ce qui s'est produit (par exemple, le décret divin). À un moment donné de sa vie, l'agent se demande ce qu'il doit faire (Est-il élu ou non? A-t-il le fameux gène ou non? Doit-il travailler sans relâche ou non? Doit-il arrêter de fumer ou non?). À cet instant précis, lequel je nomme  $t_{1,5}$ , commence le raisonnement, la prise de décision (si C se produit alors je fais x ou Non-x). C'est aussi à ce moment que, dans un certain sens,  $t_1$  et l'événement lui étant relié (c'est-à-dire C, en ce qui nous concerne) surviennent *pour l'agent*. Le raisonnement ne commence pas, comme semble le laisser croire Dupuy, lorsque Dieu prend sa décision (ou lorsque C se produit), il commence lorsque l'agent se remet en question, instant qui correspond à l'occurrence de  $t_1$  du point de vue de l'agent, où Dieu prend sa décision selon l'agent. Bref, j'affirme qu'un événement n'a de conséquence sur le processus décisionnel d'un agent qu'à partir de l'instant où ce dernier réalise que cet événement a eu lieu. Ainsi, c'est comme si l'événement avait lieu à l'instant où l'agent se questionne sur l'occurrence de cet événement, c'est-à-dire à l'instant où commence la prise de décision en rapport à cet événement.

Il me semble que Dupuy assimile  $t_{1,5}$  et  $t_2$ , où  $t_{1,5}$  représente l'instant auquel, selon moi, l'agent commence son raisonnement, et  $t_2$  représente l'instant où Dupuy fait débiter le raisonnement de son agent, c'est-à-dire à l'extérieur de la prise de décision elle-même. Donc, à partir

de ce moment,  $t_{1,5}$ , l'agent a le choix de faire  $x$  ou Non- $x$ , et ce, même s'il croit, ou s'il veut croire, qu'il n'a d'autre choix que de faire  $x$ . Il est primordial de comprendre que  $t_{1,5}$  est, dans mon raisonnement, un amalgame des instants  $t_1$  et  $t_2$  chez l'agent de Dupuy, et que le  $t_2$  de mon argument est l'instant où l'agent agit. Ce que Dupuy fait, c'est mettre son agent hors de la temporalité, comme s'il observait le déroulement des choses lors de la prise de décision: « le sujet se dotant d'une forme de pouvoir sur le passé, et se regardant de l'extérieur, comme à distance de lui-même » (Dupuy, 1992). Mais à quoi cela peut-il servir, puisqu'il ne sera jamais en mesure de connaître ce qui s'est produit en  $t_1$  sauf, selon Dupuy, lorsqu'il agira? Je m'explique : dans un tel raisonnement, *S ne peut tenir compte* du fait que C s'est produit en  $t_1$  que s'il le suppose ou en acquière le savoir en  $t_{1,5}$ , autrement il ne pourrait se faire une idée complète du raisonnement tel qu'il semble le faire dans le passage cité plus haut. Si effectivement il conçoit le raisonnement complet seulement en  $t_2$ , comme semble le croire Dupuy, c'est comme s'il se donnait un point de vue extérieur sur lui-même lui permettant de savoir ce qu'il doit faire en  $t_2$  si C s'est produit en  $t_1$ .

Imaginons que S est à  $t_2$ , il réfléchit et suppose que C s'est produit en  $t_1$ , réfléchissant encore, il se demande quels sont ses choix maintenant en  $t_2$ . Il a le choix de faire  $x$  ou Non- $x$ . Cependant, il ne peut savoir qu'en principe il devrait faire  $x$  que s'il se l'est lui-même imposé en  $t_{1,5}$ , sinon il ne pourrait connaître la proposition nécessaire «si C s'est produit à  $t_1$ , alors je fais  $x$  à  $t_2$  », et s'il se l'est effectivement imposé en  $t_{1,5}$ , alors il savait en  $t_{1,5}$  qu'il aurait le choix en  $t_2$  de faire  $x$  ou Non- $x$ . Il semble donc impossible qu'un agent puisse se dire en  $t_2$  « si C s'est produit en  $t_1$ , alors (*de façon nécessaire*) je fais  $x$  en  $t_2$  », car il n'a aucun moyen lors de la prise de décision de savoir ce qui s'est effectivement produit en  $t_1$  (n'oubliez pas le cadre déterministe de l'argument). Autrement dit, son raisonnement ne peut être que le suivant en  $t_{1,5}$  (non pas en  $t_2$ ): « Si C se produit en  $t_{1,5}$  [l'agent supposant ici l'occurrence de C], alors je ferai  $x$  en  $t_2$ , mais je sais que j'aurai le choix de faire Non- $x$  en  $t_2$  », qu'il fasse Non- $x$  en  $t_2$  ne signifie plus qu'il invalide le lien temporel entre C et  $x$ , car il sait en  $t_{1,5}$  qu'il pourra faire  $x$  en  $t_2$ . Bref, c'est comme si en  $t_{1,5}$  l'agent effectuait le raisonnement « [si C se produit en  $t_{1,5}$ , alors je ferai Non- $x$  en  $t_2$ ] et [si C se produit en  $t_{1,5}$ , alors je ferai  $x$  en  $t_2$ ] ». Voilà qui est irrationnel, me direz-vous, car l'agent possède deux croyances contradictoires. Cependant, nous pouvons traduire de façon logiquement valide<sup>3</sup> ces deux

---

<sup>3</sup>. On sait que  $((P \sim Q) \& (P \ Q)) \rightarrow (P \ (Q \vee \sim Q))$ .

propositions par : « si C se produit en  $t_{1,5}$ , alors je ferai x ou Non-x en  $t_2$  ». Cela n'a plus rien de bien extraordinaire. Ce qu'il faut retenir de ma critique, c'est que l'agent ne peut, à  $t_2$ , concevoir le raisonnement de Dupuy, car il ne possède aucun moyen de vérifier ce qui s'est effectivement produit en  $t_1$ . La seule façon dont il peut le concevoir ainsi est s'il se l'est imposé à lui-même en  $t_{1,5}$ . En d'autres mots, l'agent en  $t_{1,5}$  ne fait rien de plus que de formuler ce que nous appelons un futur contingent. Mais la proposition de l'agent de Dupuy « je suppose que C a eu lieu à  $t_1$  et, donc, que je vais décider de faire x à  $t_2$  » ne saurait être considérée comme un futur contingent, car elle porte sur un événement passé, et sa valeur est dès lors déterminée, c'est-à-dire vraie, car elle valide la première partie de la disjonction. Or, Dupuy affirme que l'agent est en mesure de se donner un pouvoir contrefactuel sur cette proposition et de faire en sorte que s'il faisait Non-x elle serait invalidée, mais selon ce que nous avons dit, elle serait tout de même vraie, car elle s'appliquerait alors à la deuxième partie de la disjonction. Voilà pourquoi j'affirme que Dupuy assimile  $t_{1,5}$  et  $t_2$ . En  $t_{1,5}$  lorsque nous avons un futur contingent, l'agent possède un tel pouvoir, mais lorsque nous sommes en  $t_2$  Dupuy se doit de postuler les principes que nous voulons éliminer pour arriver à un résultat semblable. Une autre remarque s'impose : le seul pouvoir qu'un agent peut posséder sur le futur est celui de faire en sorte qu'il soit différent de ce qu'il *aurait été autrement*, ce qui ne veut pas dire qu'il peut *changer* le futur. Alors quoi d'autre? me direz-vous. Il ne s'agit pas de modifier le cours du futur ou de modifier le cours d'un futur en devenir déjà écrit sur le Grand Rouleau. Il s'agit simplement du fait qu'un futur possible survienne plutôt qu'un autre selon que l'agent agisse d'une façon plutôt que d'une autre.

Finalement, si l'on accepte mon argument, il n'est plus nécessaire de faire entrer en ligne de compte le déterminisme, car cela n'a plus aucune importance. En effet, Dupuy se devait de postuler le déterminisme afin que son agent agisse, de façon nécessaire, en fonction de ce qui s'est produit en  $t_1$ . Or, nous avons vu que les événements n'entrent en ligne de compte lors de la prise de décision d'un agent qu'à partir de l'instant où se réalise l'occurrence de l'événement en question ou qu'il en suppose l'occurrence.

Il n'est plus nécessaire non plus de postuler deux temporalités, les évidentialistes raisonnent désormais en fonction du futur, tout comme les causalistes; et, par conséquent, il n'est plus besoin de postuler deux rationalités. Nous verrons dans ce qui suit comment alors il est

possible de justifier le choix évidentialiste (x) sans avoir recours aux postulats de Dupuy.

Dupuy poursuit : « On peut alors montrer que, se dotant de ce pouvoir sur le passé, l'agent, dans le problème de Newcomb avec cause commune, est rationnellement conduit à choisir x, c'est-à-dire le choix évidentialiste, ou calviniste [...] la démonstration, toutefois, requiert que Non-C cause Non-x. » (1994, p.85) Selon ma critique, le choix rationnel est maintenant celui des causalistes, ou des conséquentialistes. Or, nous verrons maintenant une autre façon d'interpréter le choix évidentialiste pour le rendre rationnel.

Comme nous l'avons mentionné, c'est à Tversky et Quattrone (1986) que revient l'idée de base de cette interprétation. Voici ce qu'ils nous proposent :

P1 : Les Calvinistes croient qu'ils ont procédé eux-mêmes à leur élection en choisissant x;

P2 : Les Calvinistes croient qu'ils n'ont pas procédé eux-mêmes à leur élection.

P1 et P2 sont évidemment des propositions contradictoires et Tversky et Quattrone concluent que les Calvinistes se cachent P1 parce qu'ils veulent croire que Dieu les a élus. En postulant que la première croyance est la cause de la seconde, on obtient un cas de *self-deception*. Voilà qui semble une interprétation tout à fait justifiable du choix évidentialiste, et Tversky et Quattrone ont démontré empiriquement qu'une majorité de sujets font le choix évidentialiste sur des bases semblables. Cependant, comme nous le disions plus haut, cette interprétation ne permet pas de faire ressortir la rationalité du choix évidentialiste. Au contraire, elle le fait paraître encore plus irrationnel, et c'est pourquoi je vais maintenant proposer une nouvelle interprétation basée sur la *self-deception* qui, je l'espère, nous permettra de faire ressortir cette rationalité.

La notion de *self-deception* que j'utilise est essentiellement la même que Kent Bach utilise dans "(Apparent) Paradoxes of Self-Deception and Decision" (1998) dont en voici les traits principaux :

- elle n'est pas constituée de croyances contradictoires
- elle n'est pas considérée comme intentionnelle
- on écarte l'hypothèse de l'homoncule et de la partition de l'esprit



Comment alors expliquer celle-ci? Mis à part la rationalisation, laquelle entraîne trop de problèmes, Bach identifie deux processus permettant à la *self-deception* d'empêcher certaines pensées de venir à l'esprit de l'agent :

« This can be accomplished by evasion : one avoids the thought *that p* by avoiding the thought *of p* (avoiding the thought of *p* does not count as self-deception unless one would be disposed to avoid the thought that *p* even if one did not avoid the thought of *p*). There is also what I call “jamming”, so-called because of the radio/radar analogy. Whereas with evasion one keeps one’s attention off the touchy subject by focusing it elsewhere, in jamming one clutters one’s mind with thoughts contrary to the unpleasant belief or contrary to evidence one has in support of that belief [...] For example, the rich old man thinks of his mistress’s sweet words and other displays of affection but not of her motives, which are obvious to others.» (Bach, 1998, p.169)

Ayant caractérisé la *self-deception* ainsi, voici l'interprétation que je fais du choix évidentialiste. Selon moi, on peut schématiser la conduite des Calvinistes ainsi :

P3 : Les Calvinistes croient qu'ils sont élus ;

P4 : Les Calvinistes croient que les élus doivent travailler sans relâche (faire *x*) ;

P5 : Les Calvinistes ne veulent pas croire qu'ils ont le choix de faire *x* ou Non-*x*.

Donc, les Calvinistes travaillent sans relâche, non pas pour se donner les signes de la grâce divine, mais parce qu'ils ne voient pas (ou ne veulent pas voir) qu'ils ont une autre option (*self-deception*).

En effet, puisqu'ils sont élus par Dieu, pourquoi voudraient-ils avoir à faire un tel choix ? Ce qui est désirable à leurs yeux, c'est de penser qu'ils n'ont pas le choix de faire *x*, *parce qu'ils* sont élus.

Donc, ils se cachent non pas P1, comme l'affirme Tversky et Quattrone, mais plutôt le fait d'avoir un libre-arbitre. Or, comme nous l'avons dit plus haut, Dupuy pose la question suivante : « Quel sens cela aurait-il de s'interroger sur la rationalité de l'agent si celui-ci n'était pas doté, et ne se dotait pas lui-même, de la faculté de libre-arbitre, entendue dans son sens minimal de pouvoir agir autrement qu'on ne le fait? » En interprétant le choix évidentialiste comme nous l'avons fait, nous ne rejetons pas la rationalité des agents, nous ne faisons que la cacher d'eux-mêmes selon les

processus cognitifs proposés par Bach, ce qui n'empêche en rien les sujets d'être rationnels. Ce que nous disons, c'est que *dans ce cas particulier*, les Calvinistes se cachent (pas de manière intentionnelle) leur libre-arbitre, donc leur rationalité, parce que cela est désirable. Autre conséquence de mon interprétation: la position compatibiliste est toujours possible et nous n'avons pas besoin de postuler deux temporalités comme le fait Dupuy, ce qui constitue un avantage considérable.

Devons-nous conclure que les évidentialistes sont tous des *self-deceivers*? Que tous ces gens dans les expériences de Tversky et Quattrone sont des *self-deceivers*? Je ne saurais répondre de façon générale, mais pour ce qui est des problèmes de Newcomb, je crois que nous devons admettre cela. Qu'est-ce que se cachent les *one-boxers* dans le problème de Newcomb originel? Leur rationalité (au sens où celle-ci implique le libre-arbitre). Bach (1998) croit que la décision de choisir la seule boîte opaque est une décision que l'on ne peut considérer comme rationnelle du début à la fin. Elle semble rationnelle lorsqu'elle est prise, mais lorsque vient le temps d'agir, elle nous paraît irrationnelle et une tentation survient de choisir les deux boîtes à la dernière minute. Donc, ce que les *one-boxers* se cachent, c'est le fait qu'ils aient ce choix de dernière minute, tout comme les Calvinistes se cachent le fait d'avoir le choix de faire x ou Non-X, en se disant quelque chose comme « Si je change ma décision, le prédicteur l'aura prédit et je n'aurai pas le million de dollars ».

## Références bibliographiques

- Bach, K. (1998), « (Apparent) Paradoxes of Self-Deception and Decision », in Dupuy, J.-P., ed., *Self-deception and paradoxes of rationality*. Stanford, CA : CSLI Publications, p.163-189.
- Dupuy, J.-P. (1990), « Temps du projet et temps de l’histoire », *Cahiers d’épistémologie* N° 9017, Montréal, Presses de l’Université du Québec à Montréal, 59 p.
- Dupuy, J.-P. (1994), « Temps et rationalité », in Roger Frydman, éd., *Quelles hypothèse de rationalité pour la théorie économique ?* Cahiers d’économie politique, Paris, l’Harmattan, p.69-104.
- Dupuy, J.-P. (1999), « Éthique et rationalité », *Éthique et philosophie de l’action*, Paris, Ellipses, École Polytechnique, p.9-28.
- Ferejohn, J. (1998), « Cooperation and Time », in Dupuy, J.-P., ed., *Self-deception and paradoxes of rationality*. Stanford, CA : CSLI Publications, p.151-161.
- Nozick, R. (1969), « Newcomb’s Problem and Two Principles of Choice », in N. Rescher et al., éd., *Essays in Honor of Carl G. Hempel*, D. Reidel, Dordrecht, Pays-Bas, P.114-146.
- Quattrone, G. A. and Tversky A. (1986), « Self-Deception and the Voter’s Illusion », in J.Elster, éd., *The Multiple Self*, Cambridge University Press.

Les numéros parus à compter de l'année 1996 peuvent être téléchargés en format PDF à partir du site Internet du département de philosophie de l'UQÀM [<http://www.philo.uqam.ca>]. On trouvera également à ce lien une liste complète de tous les numéros parus depuis le début de la collection en 1981.

Tous les cahiers de recherche parus dans cette série sont par ailleurs disponibles à la Bibliothèque centrale ainsi qu'au Centre de documentation des sciences humaines.

## NUMÉROS RÉCENTS

**François Blais** : *L'allocation universelle et la réconciliation de l'efficacité et de l'équité* (No. 9901);

**Michel Rosier** : *Max U versus Ad hoc* (No. 9902);

**Luc Faucher** : *Émotions fortes, constructionnisme faible et éliminativisme* (No. 9903);

**Claude Panaccio** : *La philosophie analytique et l'histoire de la philosophie* (No. 9904);

**Jean Robillard** : *L'analyse et l'enquête en sciences sociales : trois problèmes* (No. 9905);

**Don Ross** : *Philosophical aspects of the Hayek-Keynes debate on monetary policy and theory, 1925-1937* (No. 9906);

**Daniel Vanderveken** : *The Basic Logic of Action* (No. 9907);

**Daniel Desjardins** : *Aspects épistémologiques de la pensée de J.A. Schumpeter* (No 9908);

**Daniel Vanderveken** : *Success, Satisfaction and Truth in the Logic of Speech Acts and Formal Semantics*  
(No 9909);

**Luc Faucher** : *L'histoire de la folie à l'âge de la construction sociale: Étude critique de L'âme réécrite de Ian Hacking* (No 9910);

**Jean-Pierre Cometti** : *Activating Art followed by « Further remarks on art and “arthood” in contemporary French aesthetics »* (No 9911);

**Daniel Vanderveken** : *Illocutionary Logic and Discourse Typology* (No 9912);

**Dominique Lecourt** : *Sciences, mythes et éthique* (No 2001);

**Claude Panaccio** : *Aquinas on Intellectual Representation* (No 2002);

**Luc Faucher, Ron Mallon** : *L'autre en lui-même : psychologie zombie et schizophrénie* (No 2003) ;

**Luc Faucher, Pierre Poirier** : *Psychologie évolutionniste et théories interdomaines* (No 2004) ;

**Christian Arnsperger** : *De l'altruisme méthodologique à l'animisme transcendantal : le capitalisme comme pathologie du corps et de l'âme* (No 2005) ;

**Claude Panaccio** : *Subordination et singularité. La théorie ockhamiste des propositions singulières* (No 2006) ;

**Philippe Nemo** : *Miettes pour une philosophie de l'histoire post-historiciste* (No 2007);

**Pierre Milot** : *Nuages interstellaires déformés par des jets de matière – Culture scientifique et culture littéraire*  
(No 2008);

**Michel B. Robillard** : *Temps et rationalité selon Jean-Pierre Dupuy : critique et solution de rechange* (No 2009).