

Université de Montréal

**Les contre-exemples de Frankfurt ratent leur cible**  
**Un nouveau heaume pour le principe des possibilités alternatives**

par

Simon-Pierre Chevarie-Cossette

Département de philosophie

Faculté des arts et sciences

Mémoire présenté à la Faculté des arts et sciences  
en vue de l'obtention du grade de Maîtrise Ès Arts (M. A.)  
en philosophie  
option recherche

Le mercredi 25 juin 2014

© Simon-Pierre Chevarie-Cossette, 2014



## Résumé

Les contre-exemples de Frankfurt sont inoffensifs contre l'argument de la conséquence (*consequence argument*), l'argument qui, à partir du principe des possibilités alternatives et du déterminisme, montre que nous ne pouvons être tenus moralement responsables de nos actions. En effet, ils sont formulés soit dans un cadre déterministe, soit dans un cadre indéterministe. S'ils sont formulés dans un cadre indéterministe, ils sont inoffensifs parce qu'ils contreviennent à un principe méthodologique que nous défendons : le principe de non-négation des prémisses (PNNP). En fait, nous montrons que pour tout argument donné, il est proscrit de supposer la négation d'une prémisses afin de réfuter une autre prémisses à moins que l'attaque réussisse à réfuter les deux prémisses en question. Or, d'une part, les contre-exemples de Frankfurt indéterministes supposent explicitement qu'une prémisses de l'argument de la conséquence – le déterminisme est vrai – est fausse; et d'autre part, ils ne peuvent pas nous donner de raisons de croire en l'indéterminisme, ce que nous montrons grâce à des considérations sur la transmission de la justification. Construire des contre-exemples de Frankfurt indéterministes est donc incorrect pour des raisons méthodologiques et logiques. S'ils sont formulés dans un cadre déterministe, les contre-exemples de Frankfurt font face à une autre accusation d'entorse argumentative, présentée dans la défense du dilemme (*Dilemma Defence*) de Kane-Ginet-Widerker : celle de la pétition de principe. Nous inspectons et nuancions cette accusation, mais concluons qu'elle tient puisque les contre-exemples de Frankfurt déterministes supposent au final une analyse des agents contrefactuels dans les mondes déterministes et de la relation « rendre inévitable » que ne peuvent endosser ni les incompatibilistes de la marge de manœuvre (*leeway incompatibilists*), ni les incompatibilistes de la source (*source incompatibilists*) ni non plus les semicompatibilistes. Conséquemment, les contre-exemples de Frankfurt ne peuvent plus soutenir la forme de compatibilisme à laquelle ils ont donné naissance. L'incompatibilisme de la source ne peut plus être préféré à l'incompatibilisme de la marge de manœuvre ni non plus rejeter toute participation des possibilités alternatives dans l'explication de la responsabilité morale sur cette seule base.

**Mots-clés** : Contre-exemples de Frankfurt, principe des possibilités alternatives, principe de non-négation des prémisses, incompatibilisme, compatibilisme, expériences de pensée, pétition de principe, responsabilité morale, libre-arbitre, PAP.

## Abstract

Frankfurt-type examples are inoffensive against the Consequence argument, which purports to show that from both the principle of alternative possibilities and determinism, we can deduce that we are not morally responsible for our actions. Indeed, they require either a deterministic context or an indeterministic one. If they require *indeterminism*, they are harmless because they violate a methodological principle that we defend: the no-premise-negation principle (PNNP). In fact, we show that for each given argument, we cannot legitimately suppose the negation of a premise in order to refute another, unless the attack succeeds to refute both premises. Yet, on the one hand, indeterministic Frankfurt-type examples explicitly suppose that a premise of the Consequence argument – determinism holds – is false; and on the other hand, Frankfurt-type examples do not give us reasons to reject determinism, which we show with considerations on the transmission of justification. To build indeterministic Frankfurt scenarios is therefore incorrect for methodological and logical reasons. If they require *determinism*, Frankfurt-type examples are facing a different, yet very serious, argumentative accusation (presented in the Kane-Ginet-Widerker famous Dilemma defence): begging the question. We inspect and qualify this accusation. However, we ultimately claim that it still holds because a deterministic Frankfurt-type example supposes an analysis of the role of counterfactual agents and of the relation “render inevitable” in deterministic worlds that is acceptable neither for a leeway incompatibilist, a source incompatibilist or a semicompatibilist. Thus, Frankfurt-type examples no longer successfully support the form of compatibilism they contributed to give birth to. Also, source incompatibilism may not be preferred to leeway incompatibilism, nor reject all use of alternative possibilities in the explanation of moral responsibility on the sole basis of Frankfurt scenarios.

**Keywords:** Frankfurt-type examples, principle of alternative possibilities, no-premise negation principle, incompatibilism, compatibilism, thought experiments, *petitio principii*, moral responsibility, freewill, begging the question.

# Table des matières

Introduction.....	8
Chapitre I : les prémisses ne sont pas des atomes.....	19
1. Un contre-exemple à une prémisse ne devrait pas supposer la négation d'une autre prémisse.....	19
2. Deux exemples supplémentaires.....	24
3. Caractère général du PNNP.....	30
4. Réponse à quelques objections contre le PNNP.....	35
A. L'utilité du PNNP.....	35
B. La légitimité dialectique de la construction du PNNP.....	38
C. Les contre-exemples et leur nature.....	39
Chapitre II : le caractère inoffensif des contre-exemples de Frankfurt contre l'argument de la conséquence.....	42
1. L'illégitimité des contre-exemples de Frankfurt indéterministes.....	42
A. Immunisation du PAP contre un contre-exemple indéterministe.....	42
B. Satisfaction des clauses du PNNP : défense de (2).....	45
C. Satisfaction des clauses du PNNP : l'impossibilité de réfuter le déterminisme par un contre-exemple de Frankfurt.....	47
2. L'illégitimité des contre-exemples de Frankfurt déterministes.....	59
A. La défense du dilemme ( <i>dilemma defence</i> ) de Kane-Ginet-Widerker.....	59
B. Deux réponses à l'accusation de pétition de principe.....	61
C. Analyse critique des solutions de Fischer et de Haji-McKenna.....	67
D. Constats et nuances sur la question de la pétition de principe.....	75
E. La relation « rendre inévitable » et les agents contrefactuels dans un monde déterministe.....	81
Chapitre III : réponse aux répliques de Frankfurt, et de Dennett et Taylor.....	98
1. La réponse de Frankfurt.....	99
2. La critique de Dennett et de Taylor sur la relation entre déterminisme et causalité ...	102
3. La critique de Dennett et de Taylor sur la relation entre déterminisme et la possibilité	

Conclusion .....	110
1. Récapitulatif .....	110
2. Les conséquences de notre analyse sur l'incompatibilisme de la source .....	114
Bibliographie .....	118

*À ma mère, qui m'a donné le goût de la lecture,  
À mon père qui m'a donné celui d'argumenter et  
à Justine et Geneviève qui m'en ont donné  
l'occasion.*

## **Remerciements**

Je remercie Thomas Mekhaël, Marc-André Caron et David Rocheleau Houle pour leur lecture attentive et leurs commentaires, qui auront rendu le tout plus précis et compréhensible; Oisín Deery et Mathieu Bélanger pour leur validation de mes intuitions les plus fondamentales; Geneviève Griffin et Thiago Hutter pour leur discipline que j'ai peiné tant bien que mal à calquer; le département de philosophie de l'Université de Montréal et son association étudiante pour avoir aiguisé mon esprit critique; enfin et surtout, mon directeur, Daniel Laurier, pour sa minutie, son attention et sa sagacité.



Même la connaissance intuitive d'une aussi grande force logique que le principe de symétrie employé par Archimède peut nous égarer [...]. Ce qui est instinctif est tout aussi faillible que ce qui est distinctement conscient. Sa seule valeur repose dans les provinces qui nous sont familières. (Mach 1889)

# Introduction

Pouvons-nous être tenus moralement responsables d'actions si elles s'inscrivent dans une chaîne causale ininterrompue? Autrement dit, si le déterminisme nomologique<sup>1</sup> est vrai, disposons-nous de la liberté nécessaire pour être dotés de la responsabilité morale? C'est une question qui appelle des réponses à la fois diverses et inscrites dans des traditions riches et pérennes; des traditions d'autant plus riches qu'elles s'abreuvent à des débats en philosophie de l'action, en métaphysique et en métaéthique; d'autant plus pérennes que l'ampleur de leurs conséquences sur la philosophie politique, sur la philosophie du droit et sur l'éthique ramène constamment cette question, tantôt avec force, tantôt implicitement, à l'esprit de ceux qui se hasardent à douter du fondement de nos comportements et de nos institutions<sup>2</sup>. Ce mémoire n'entend pas aborder la question de l'existence du libre arbitre à proprement parler ou celle de la vérité du déterminisme. C'est strictement à celle de leur compatibilité que nous nous intéresserons ici et c'est pourquoi déterministes et libertariens<sup>3, 4</sup> ont autant l'un que l'autre leur place dans notre analyse.

---

<sup>1</sup> Chaque fois que nous ferons référence au déterminisme, nous entendrons par là le déterminisme nomologique, c'est-à-dire le déterminisme total des lois de la nature, par opposition à un déterminisme divin ou logique (c.f. Hofer 2010). Le déterminisme nomologique suppose que la conjonction des lois de la nature et de l'état du monde à un certain moment implique l'état du monde à un autre moment donné (avant ou après).

<sup>2</sup> Débattre de la question du libre arbitre dépend largement de la supposition que nos pratiques et croyances sociales et individuelles concernant la moralité sont possiblement très sérieusement erronées (Elzein 2013). C'est peut-être à cause de cette conséquence que la majorité des philosophes anglo-saxons sont compatibilistes (Bourget et Chalmers 2013) et qu'il y a une très grande différence entre les intuitions des gens sur le compatibilisme et leur façon de résoudre des problèmes en utilisant celles-ci (Deery, et al. 2014, Nahmias, et al. 2005).

<sup>3</sup> La plupart des termes que nous employons sont largement utilisés dans la littérature. Nous mettons à profit la nomenclature proposée par van Inwagen (2008).

<sup>4</sup> Ceux qui croient à une forme de libre arbitre et rejettent du même coup une forme de déterminisme total. Il est évident que le libertarien n'a pas besoin, pour défendre une forme ou une autre de libre arbitre, de postuler qu'aucun processus déterministe n'existe.

On peut bien sûr se demander pourquoi des libertariens tiendraient à défendre l'une ou l'autre des formes d'incompatibilismes<sup>5</sup> (p. ex. Ginet 1990, Van Inwagen 1975, Kane 2001)<sup>6</sup>. La question se pose en effet puisque pour se pencher sur la compatibilité entre le déterminisme et la responsabilité morale, il faut, provisoirement du moins, faire la supposition que le déterminisme est vrai. Au-delà de la simple curiosité intellectuelle, il y a fort à parier que nombreux sont ceux qui y voient tantôt une manière de justifier l'importance d'un modèle de libre arbitre qui viendrait sauver la responsabilité morale; tantôt une *raison suffisante* pour soutenir leur libertarisme<sup>7</sup>. Quant à ceux qui adhèrent au déterminisme ou qui croient que c'est une position probable<sup>8</sup> ou simplement raisonnable, ils sont presque contraints de se positionner sur la question de la compatibilité ou du moins d'en envisager les conséquences. Si le compatibiliste<sup>9</sup> peut aisément s'extraire par la suite du débat, l'incompatibiliste déterministe (qu'on appelle aussi « déterministe dur » (*hard determinist*)) devra suggérer une manière d'accepter l'absence du libre arbitre. Certains déterministes durs croient que nous *devrions* entretenir l'illusion du libre arbitre (p. ex. l'illusionisme de Smilansky 2002); d'autres suggèrent que nous y sommes *psychologiquement contraints* (Strawson 1974), encore que cela n'empêche pas de modifier nos institutions (Elzein 2013). Enfin, les éliminativistes y voient l'occasion de nous délester de notre trouble intérieur en prenant conscience de nos déterminations ou d'éliminer certaines de nos attitudes réactionnelles néfastes<sup>10</sup> issues d'une

---

<sup>5</sup> L'incompatibilité du déterminisme et de la liberté nécessaire à la responsabilité morale.

<sup>6</sup> Pour plus d'exemples, voir (Clarke 2008).

<sup>7</sup> C'est dire qu'il existe des auteurs (p. ex. Lycan 2003, Van Inwagen 2008) qui croient au libre arbitre parce qu'ils n'acceptent ni les conséquences du déterminisme ni la solution compatibiliste.

<sup>8</sup> Pereboom, par exemple, croit que certaines théories libertariennes qui laissent place à la responsabilité morale sont parfaitement cohérentes avec l'existence de la responsabilité morale, mais qu'elles s'accordent mal avec nos meilleures théories physiques (Pereboom 2011).

<sup>9</sup> Celui qui croit que le déterminisme est compatible avec la liberté nécessaire à la responsabilité morale.

<sup>10</sup> Il existe bien sûr tout un continuum théorique sur la question. Certains rejettent seulement une partie de nos attitudes réactionnelles ainsi qu'une forme tronquée de responsabilité morale (Bonicalzi 2013).

conception erronée de l'étendue de notre liberté, par exemple les remords ou la colère (Spinoza - *Éthique* II SP49, Pereboom 2009, 2011)<sup>11</sup>. L'absence de libre arbitre ne serait pas dramatique, précisément parce que la notion de responsabilité morale pourrait être avantageusement remplacée.

À la question de la compatibilité entre le déterminisme et le libre arbitre (la liberté nécessaire à la responsabilité morale), question qui, nous venons de le voir, a toujours une certaine importance, plusieurs réponses contradictoires font autorité et il convient ici de les classer. Nous avons dit qu'il y a d'une part les incompatibilistes, qui croient que si le déterminisme était vrai, nous ne serions pas réellement moralement responsables de nos actions. Parmi eux, un désaccord subsiste sur la raison fondamentale pour laquelle le déterminisme, s'il était vrai, éclipserait la responsabilité morale. Certains que l'on nomme « incompatibilistes de la marge de manœuvre » (*leeway incompatibilists*) (p. ex. Van Inwagen 1975), croient que c'est parce que nous sommes dépourvus de possibilités alternatives que nous ne pouvons être libres et responsables si le déterminisme est vrai. Être responsable d'une action implique la possibilité de l'éviter. D'autres (p. ex. Pereboom 2003, Kane 1985), qu'on appelle les « incompatibilistes de la source » (*source incompatibilists*), avancent plutôt que le déterminisme nous empêche d'être la source ultime de nos actions. À aucun moment dans notre histoire personnelle nous n'aurions eu un réel choix permettant d'asseoir une éventuelle

---

<sup>11</sup> Quoique Pereboom se réclame directement de Spinoza lorsqu'il dit embrasser un déterminisme dur (Pereboom 2011), il n'est peut-être pas très clair si Spinoza est compatibiliste ou incompatibiliste, mais il est certainement révisionniste à l'égard de ce que devraient être nos attitudes face aux actions d'autrui. Clairement, pour lui, la vérité du déterminisme change notre approche au monde et à notre façon de concevoir notre liberté et celle des autres. Or, s'il n'y a pas de libre arbitre chez Spinoza (*Éthique*, II, P48), il y a bel et bien une liberté, qu'on acquiert par la connaissance de nos déterminations (*Éthique*, V, SX). En ce sens, le déterminisme n'efface pas la liberté, mais en change la définition.

responsabilité morale<sup>12</sup>. Bref, nous pourrions être responsables d'une action inévitable, si nous pouvions être responsables de *qui nous sommes*, ce que le déterminisme empêche néanmoins. Ceux des incompatibilistes qui accordent une valeur aux possibilités alternatives (tous les incompatibilistes de la marge de manœuvre et les incompatibilistes de la source qui croient malgré tout qu'il doit exister, pour qu'il y ait de la responsabilité morale, certaines situations où il y a bel et bien des possibilités alternatives) s'entendent généralement sur une conception de la possibilité. Pour eux, avoir une possibilité alternative présente signifie, *grosso modo*, qu'il existe un monde possible dont le passé est exactement le même que celui du monde réel, mais où j'agis présentement différemment. Autrement dit, pour que je sois moralement responsable, il faut qu'il y ait plusieurs futurs *métaphysiquement accessibles*, des futurs donc, qui pourraient exister sans changer pour ce faire le passé<sup>13</sup>, y compris des futurs où j'agis différemment<sup>14</sup>. Quant aux incompatibilistes de la source qui ne font pas du tout appel aux possibilités alternatives, jamais ne les voit-on endosser une façon différente de définir les possibilités.

D'autre part, il y a les « compatibilistes », qui croient que même si le déterminisme est vrai, nous sommes libres et donc responsables. Certains compatibilistes adhèrent au déterminisme et d'autres non, mais encore une fois ce ne sont pas ces catégories qui nous intéressent ici. Les « compatibilistes classiques » (p. ex. Hume 2000/1748, Ayer 1954, Schlick

---

<sup>12</sup> Pour plusieurs incompatibilistes de la source (c.f. Kane 1989) les possibilités alternatives sont nécessaires à la responsabilité morale, mais n'en sont pas l'ingrédient principal. En effet, il faut qu'à un moment décisif de son existence, l'agent ait accès à de réelles possibilités. Néanmoins, il est très clair que pour l'incompatibiliste de la source, il n'est pas nécessaire que, pour qu'un agent soit moralement responsable de n'importe quelle action *x*, il ait la possibilité de faire autrement que *x*.

<sup>13</sup> De nombreux compatibilistes sont au contraire enclins à croire que ce qui est possible est ce qui est compatible avec les lois de la nature et un monde dont le passé est *similaire* au nôtre (Taylor et Dennett 2011) ou est ce qui est à la fois métaphysiquement possible et non empêché par des facteurs extérieurs (Chrysippe, tel qu'analysé par Bobzien 1997). Le fait que nous soyons des éléments essentiels de la chaîne causale est donc source de liberté.

1966, Berofsky 2003) proposent que la possibilité de faire autrement se définisse au conditionnel, c'est-à-dire comme le fait que si les désirs d'un individu eussent été différents (ce qui est peut-être impossible, si les désirs sont déterminés), celui-ci eut agi différemment<sup>15</sup>. Il y a donc chez les compatibilistes classiques une importance fondamentale attachée à la notion de possibilité ou de pouvoir. D'autres (p. ex. Frankfurt 1971, McKenna 2011, Dennett 1984a, Fischer 2002) ont préféré postuler l'absence de lien entre possibilités (au sens incompatibiliste ou au sens conditionnel) et responsabilité morale. La responsabilité morale dépend des propriétés de notre vie mentale, qu'il s'agisse de notre rationalité, de notre capacité à reconnaître et à évaluer les raisons et les conséquences d'une action; qu'il s'agisse encore d'absence de contraintes intérieures (comme la maladie) et extérieures (comme la manipulation) ou encore de conformité des volontés et des actes. Ce qui permet la responsabilité morale n'a donc rien à voir avec les possibilités alternatives. C'est dire que même si personne n'avait de possibilité alternative, la responsabilité morale serait toujours possible.

Bien sûr, ces catégories répondent surtout à un classement des théories *contemporaines* sur la compatibilité du libre arbitre et du déterminisme. Par exemple, bien que l'on considère largement que les stoïciens étaient compatibilistes, il semble que la très grande majorité d'entre eux ne se soient même pas intéressés au problème du libre arbitre, mais seulement à celui de l'action non contrainte (« ce qui dépend de nous ») (Bobzien 1997). Le cas de Kant, bien que très différent, reflète aussi l'imperfection de ces catégories, cette fois par la difficulté

---

<sup>14</sup> Cet énoncé dépend de ce qu'on a appelé le *poss-ability principle* (Spencer 2013), qui stipule que si  $x$  est possible pour un agent, alors  $x$  est métaphysiquement possible.

<sup>15</sup> Par exemple, Hume définit la liberté ainsi : « [b]y liberty, then, we can only mean a power of acting or not acting, according to the determinations of the will; that is, if we choose to remain at rest, we may; if we choose to move, we also may » (Hume 2000/1748, §73).

à laquelle se heurte celui qui entend le classifier : il admet d'une part que si les phénomènes sont des objets, le déterminisme élimine la liberté (Kant 2001/1781/1787, pp. 497-498); et d'autre part, que la liberté est compatible avec le déterminisme si ce ne sont que nos représentations qui sont soumises à la causalité (Langsam 2000). C'est seulement grâce à la sphère nouménale et pratique<sup>16</sup> que la liberté est préservée. Et pourtant, Kant rejette la vision des compatibilistes comme Hume, la décrivant comme un « misérable subterfuge » (Kant 2002/1788, p. 112)! On peut donc dire que Kant partage les conclusions des compatibilistes contemporains tout en endossant les justifications des incompatibilistes.

Quoi qu'il en soit, la division que nous avons présentée a l'avantage de marquer une distinction importante à même celle des écoles de pensée compatibilistes et incompatibilistes : il y a d'une part ceux qui croient que la notion de possibilité est un ingrédient central et nécessaire à la responsabilité morale pour une action précise (incompatibilistes de la marge de manœuvre, compatibilistes classiques) et ceux qui le refusent (incompatibilistes de la source, compatibilistes à la Frankfurt). Les premiers, même s'ils en donnent des interprétations différentes, acceptent le principe des possibilités alternatives, principe clef qui permet de lier des considérations normatives à la vérité du déterminisme (Elzein 2013) :

PAP : Une personne est moralement responsable d'une action seulement si elle avait pu agir autrement.<sup>17 18</sup>

---

<sup>16</sup> Il est loin d'être clair que la sphère pratique et la sphère nouménale sont une seule et même chose.

<sup>17</sup> « Agir autrement » est interprété par certains comme « poser une action différente », alors que pour d'autres cela inclut aussi « ne pas agir ». De même, il y a un consensus au sein des défenseurs du PAP à propos du fait que la possibilité nécessaire à la responsabilité morale doit avoir du « omph », un caractère « robuste » (Fischer 2002) c'est-à-dire qu'elle doit être accessible à l'agent. Pour que cette possibilité soit accessible, il faut que l'agent sache qu'il peut vouloir une possibilité alternative et que cette possibilité l'absoudra probablement de sa responsabilité morale actuelle (Pereboom 2011). Ainsi, si la seule façon d'éviter la mort de ma mère d'un accident de la route est de crever les pneus de sa voiture, mais que cela m'est inconnu, je ne peux compter cette possibilité comme une possibilité suffisante pour satisfaire l'interprétation la plus commune du PAP.

<sup>18</sup> L'interprétation compatibiliste du PAP postule qu'un agent a la possibilité de faire autrement si, dans l'hypothèse où il a d'autres désirs, états mentaux, volontés, etc., il agit différemment.

Les seconds, quant à eux, le rejettent. Ils sont largement inspirés des idées qu'a avancé Frankfurt dans un article de 1969 dans lequel il rejette le PAP, peu importe son interprétation, un article dont l'influence est comparée à celle de Gettier en épistémologie (Fischer 2010).

Pour attaquer le PAP, Frankfurt propose une expérience de pensée :

Jones hésite à poser une action. Or, à son insu, Black, un agent contrefactuel<sup>19</sup>, s'est assuré que Jones finira par décider de poser l'action. En effet, si Jones manifeste qu'il *s'apprête* à décider du contraire<sup>20</sup>, Black déclenchera un dispositif qui forcera Jones à changer d'avis. Finalement, Jones pose son action sans que Black intervienne. (Frankfurt 1969)

Frankfurt croit que cet exemple devrait nous convaincre qu'il est possible d'être moralement responsable d'une action inévitable. En effet, Jones n'a pas de possibilités alternatives (il ne peut pas ne pas poser son action et ne peut pas prendre la décision de ne pas la poser), mais semble néanmoins responsable de son action.

De nombreuses attaques ont été formulées contre les scénarios de Frankfurt et notre propos n'est pas ici d'en faire l'énumération et encore moins l'évaluation exhaustive<sup>21</sup>. Nous comptons plutôt proposer une attaque originale contre lui, laquelle s'appuiera bien sûr en partie sur des critiques antérieures, mais qui, pour une large part, repose sur un principe méthodologique de notre cru. Nous estimons que de nombreux auteurs, y compris des critiques de Frankfurt, se sont laissés piéger dans l'élaboration des expériences de pensée

---

<sup>19</sup> Par agent contrefactuel, on entend un agent qui n'intervient pas, mais qui serait intervenu si le passé avait été différent. « Contrefactuel » réfère à « conditionnel », à « ce qui serait arrivé si les choses avaient été différentes ».

<sup>20</sup> De très nombreuses versions du contre-exemple de Frankfurt ont été élaborées, aussi ne voulons-nous pas en décrire une version trop précise.

<sup>21</sup> Pour un excellent compte-rendu des différents contre-arguments au PAP, voir (Widerker et McKenna 2006b). Après un tel ouvrage, s'attaquer encore une fois aux contre-exemples de Frankfurt peut paraître abusif, mais non seulement les compatibilistes à la Frankfurt sont tenaces, mais certains incompatibilistes comme Widerker se sont rendus aux arguments de Frankfurt (Widerker 2006a). Par ailleurs, la critique des contre-exemples de Frankfurt est une occasion admirable de se pencher sur toutes sortes de questions connexes qui sont loin de faire l'unanimité. De toute façon, pour reprendre les mots de Lycan : « Think about it. On *any* philosophical topic, the person who propounds an analysis is going to get creamed. Philosophical analyses virtually never work, but are lacerated by counterexample after counterexample. » (Lycan 2003, p. 112)



qu'ils mobilisent, qu'il s'agisse des expériences de Frankfurt ou de ses versions cousines (par exemple, les contre-exemples de Fischer ou de Pereboom). Il nous est avis que le principe dont nous nous réclamons, s'il est susceptible d'être débattu, devrait recueillir un assentiment général. Le considérer comme de peu d'importance parce que trivial serait faire abstraction de la quantité non négligeable de fois où on l'a enfreint dans la littérature sur le libre arbitre et dont l'écho se répercute probablement hors du champ de notre enquête. Le rejeter comme trop contraignant ferait fi des intuitions solides qui l'appuient.

En réalité, nous sommes persuadé que Frankfurt et bon nombre de ses critiques sont coupables d'un double oubli : d'abord, oubli du contexte dans lequel le contre-exemple a été proposé et reste le plus important; et ensuite, oubli du fait que les agents contrefactuels ont un rôle diamétralement différent selon que le déterminisme est vrai ou faux. Le contexte principal d'utilisation du principe des possibilités alternatives (PAP) est l'argument des incompatibilistes pour montrer que le déterminisme supprime notre responsabilité morale, l'argument de la conséquence (*Consequence argument*) (Ginet 1966, Van Inwagen 1975)<sup>22</sup>. Une fois qu'on a montré que de la fixité des lois de la nature, il découle que nul n'a de possibilités alternatives, le principe des possibilités alternatives permet de nier toute forme de responsabilité morale. Cet argument traditionnel des « hard determinists » (Carroll et Markosian 2010) peut être ainsi présenté :

P1 : Le déterminisme est vrai, ce qui signifie que la conjonction des lois de la nature et de l'état du monde à un temps  $t$  implique l'état du monde à un temps  $t_{+/-n}$  (antérieur ou postérieur à  $t$ ), incluant chacune de mes actions.

P2 : Si mes actions sont impliquées par une chose sur laquelle je n'ai aucun pouvoir, alors je n'aurais pas pu agir autrement.

---

<sup>22</sup> De nombreuses versions de cet argument ont été proposées. Nous en proposerons nous-mêmes une plus loin (III, 3).

P3 : Je n'ai de pouvoir ni sur un état du monde précédant ma naissance ni sur les lois de la nature.

PAP : Je suis moralement responsable d'une action seulement si j'avais pu agir autrement.

---

C1 : Je ne suis responsable d'aucune de mes actions.

Un long débat a fait rage depuis la publication de l'article de Frankfurt : certains ont cru que celui-ci avait réfuté le PAP et ont utilisé ce résultat pour asseoir leurs théories du libre arbitre; plusieurs autres ont identifié des failles dans les contre-exemples qui ont été réaménagés à maintes reprises par leurs défenseurs. Nous nous rangerons dans cette seconde catégorie et tenterons d'exposer des raisons originales de rejeter les contre-exemples de Frankfurt.

Cependant, tenants comme défenseurs du PAP ont toujours pensé que si les contre-exemples de Frankfurt au PAP étaient concluants, l'argument canonique contre la responsabilité morale tomberait, lui aussi. En réalité, les contre-exemples de Frankfurt échouent à mettre en échec l'argument contre la responsabilité morale, *et ce, même s'ils minent le PAP*.

Notre analyse s'articulera en plusieurs étapes de nature très différentes, tantôt méthodologiques, tantôt métaphysiques. Le premier chapitre sera dédié à la formulation et à la défense d'un principe – le principe de non-négation des prémisses (PNNP) – permettant d'encadrer la construction de contre-exemples philosophiques en général. Après l'avoir défini (I, 1) et montré en quoi il est intuitif dans trois cas distincts (I, 2), nous en prouverons la généralité (I, 3) et répondrons à des objections auxquelles il pourrait être confronté (I, 4).

Le second et principal chapitre aura pour objectif de montrer que les contre-exemples de Frankfurt sont mis en échec. D'une part, s'ils supposent l'indéterminisme – et c'est ainsi que de nombreux auteurs les ont construits (Pereboom 2003, 2009a, 2001, 2012, Widerker 2006a, 2009, Mele et Robb 1998, Haji et McKenna 2006, Hunt 2007) –, ils enfreignent le

PNNP (II, 1). En effet, d'un côté les contre-exemples de Frankfurt indéterministes supposent explicitement qu'une prémisse de l'argument de la conséquence – le déterminisme est vrai – est fausse; et d'un autre côté, ils ne peuvent pas nous donner de raisons de croire en l'indéterminisme, ce que nous montrons grâce à des considérations sur la transmission de la justification. D'autre part, s'ils supposent le déterminisme, les contre-exemples de Frankfurt sont coupables d'une pétition de principe. Nous inspecterons et nuancerons cette accusation, avant de montrer qu'elle tient au final parce que les contre-exemples de Frankfurt déterministes supposent une analyse des agents contrefactuels et de la relation « rendre inévitable » qui est inacceptable pour un incompatibiliste de la marge de manœuvre, pour un incompatibiliste de la source et même pour un semicompatibiliste<sup>23</sup> (II, 2). Les deux parties du deuxième chapitre correspondront donc aux deux branches d'un dilemme : soit les contre-exemples de Frankfurt sont indéterministes (II, 1), soit ils sont déterministes (II, 2).

Notre troisième chapitre abordera deux réponses à notre analyse, celle de Frankfurt et celle de Dennett et Taylor. Cela nous permettra d'abord de montrer que Frankfurt semble avoir exagéré la portée de ce que doit soutenir le tenant du principe des possibilités alternatives. Puis, nous nous pencherons sur la critique de Dennett et de Taylor, laquelle illustre à merveille l'*outréculance intuitive* de la thèse compatibiliste, mais surtout la nécessité d'un cadre d'analyse entièrement compatibiliste pour admettre les contre-exemples de Frankfurt. Nous en profiterons pour critiquer leur conception de la notion de possibilité, laquelle contredit directement notre analyse du deuxième chapitre.

---

<sup>23</sup> C'est-à-dire pour celui qui s'abstient de se prononcer sur le lien entre le déterminisme et les possibilités alternatives, mais qui suppose que le déterminisme ne supprime pas la responsabilité morale parce que les possibilités alternatives n'en sont pas des ingrédients essentiels.

Nous clorons notre parcours par un récapitulatif et une analyse des impacts possibles sur l'incompatibilisme de la source. Un débat de quatre décennies risque-t-il de prendre fin au terme de ce mémoire? Probablement pas, même s'il nous semble que la part du lion des raisons que nous y avancerons relève d'un terrain, s'il en est un, où un consensus devrait être envisageable, celui de la méthode et de la logique. Si tous s'entendent pour contester la suffisance d'une bonne méthode et d'une argumentation valide, il y a peu de nos contemporains qui se sont hasardés à les juger non nécessaires. Si notre verdict ne peut demeurer sans appel, il reste que renverser constamment le fardeau de la preuve est certainement la seule façon d'avancer et c'est à cette tâche que nous nous attarderons dès à présent.

# Chapitre I : les prémisses ne sont pas des atomes

## 1. Un contre-exemple à une prémisse ne devrait pas supposer la négation d'une autre prémisse

Décomposer un argument en prémisses et s'attarder à chacune d'entre elles séparément a certainement des avantages non négligeables : cerner les problèmes clairement et ne pas, pour ainsi dire « s'empêtrer dans les mots »; avoir une discussion qui, parce que ses participants s'entendent sur ses composantes, peut évoluer; enfin, pouvoir catégoriser nettement des positions en fonction du rejet ou de l'acceptation de telle ou telle prémisse d'un argument important. Le débat contemporain sur le libre arbitre est paradigmatique de cette pratique puisque, comme nous l'avons vu précédemment, les différentes positions canoniques défendues sur la conciliation entre libre arbitre et déterminisme peuvent s'articuler – et l'ont souvent été, à tort ou à raison<sup>24</sup> – autour de différentes réactions possibles face à l'argument de la conséquence ou face à ses avatars (entre autres, l'argument direct (*direct argument*), plus simple).

Force est de constater qu'il y a néanmoins des désavantages à procéder d'une telle façon. En analysant en isolation une prémisse, le risque d'omettre son contexte est latent et cette omission n'est pas inoffensive. À de nombreuses reprises, des philosophes ont avancé la nécessité d'étudier les thèses dans leur contexte historique – cette importance a d'ailleurs été à l'occasion tempérée, donnant lieu à une panoplie « d'historicismes » (cf. Glock 2008) –, mais

---

<sup>24</sup> C'est ce que déplore Tognazzini quand il défend l'idée que l'incompatibilisme de la source ne devrait pas être défini fondamentalement comme rejetant le principe des possibilités alternatives (PAP) (Tognazzini 2011).

cette question en a peut-être éclipsé une autre, plus modeste, et c'est celle de l'oubli du contexte non pas historique, mais formel, c'est-à-dire l'argument dont la prémisse fait partie.

Nous croyons que ce genre d'oubli a entraîné une erreur majeure, c'est-à-dire le non-respect d'un principe argumentatif que nous défendrons : il ne faut pas, en général, rejeter une prémisse d'un argument en vertu d'un contre-exemple qui entre en contradiction avec l'une des autres prémisses de l'argument concerné. Sinon, cela supposerait que cette autre prémisse soit rejetée, ce qui n'est pas toujours souhaité. Le principe de non-négation d'une prémisse (PNNP) que nous proposons mérite d'être clairement énoncé :

PNNP : Un contre-exemple à une prémisse  $x$  ne doit pas supposer la négation d'une prémisse  $y$  si (1)  $x$  et  $y$  sont deux prémisses différentes d'un même argument et si (2)  $y$  n'est pas réfutée elle aussi.

Illustrons d'emblée le principe par un exemple (dont nous avons simplifié la formalisation). Nous reviendrons sur ses clauses ultérieurement.

Un électeur dit à un second, qui partage les mêmes convictions progressistes que lui : « demain je dois voter Québec solidaire; en effet,

P4 : Demain, je vote et c'est dans Hochelaga que je peux exercer mon vote.  
( $p \wedge \forall h$ )

P5 : Or, si demain je vote, je dois voter pour le parti le plus progressiste.  
( $p \rightarrow r$ )

P6 : Si je dois voter pour le parti le plus progressiste et qu'un candidat de Québec solidaire se présente dans mon comté, je dois voter pour Québec solidaire. ( $(\forall x (Vx \rightarrow Qx)) \rightarrow (r \rightarrow s)$ )

P7 : Un candidat de Québec solidaire se présente dans Hochelaga. ( $Qh$ )

C2 : Je dois donc voter Québec solidaire. ( $s$ ) »<sup>25</sup>

---

<sup>25</sup> Nous poursuivons la notation entamée précédemment pour éviter toute confusion entre différentes P1, d'autant plus que nous revenons parfois sur des arguments avancés plusieurs dizaines de pages plus tôt.

Le second lui répond : « écoute, P5 est fausse; ce n'est pas un principe général acceptable. En effet, dans les comtés où le parti le plus progressiste n'a aucune chance de se faire élire, tu devrais plutôt voter pour le deuxième parti en lice qui puisse détrôner les Libéraux. Bien sûr, je ne parle pas de comtés comme Mercier ou Hochelaga où un candidat solidaire a des chances de l'emporter, mais plutôt de comtés à Laval, par exemple. Il existe donc une situation où  $p$  est vrai et où  $r$  est faux. P5 est donc fausse et par voie de conséquence, ton argument n'est pas probant parce que ses prémisses ne sont pas toutes vraies. »

Il semble très intuitif de considérer que ce contre-argument n'est pas une attaque sérieuse à l'argument dans son ensemble, précisément *parce qu'il enfreint le PNNP*. D'abord, établissons qu'il y a bel et bien infraction du PNNP : il s'agit d'un contre-exemple à une prémisses  $x$  (P5) qui suppose la négation d'une prémisses  $y$  (P4) et qui satisfait les deux conditions d'application du PNNP : (1)  *$x$  et  $y$  sont deux prémisses différentes d'un même argument* (c'est le cas) et (2)  *$y$  n'est pas réfutée elle aussi*. Évidemment, (1) et (2) sont satisfaits. Il serait en effet absurde de croire que le contre-exemple à P5 réfute P4 : le second électeur ne prouvera pas à l'aide de considération stratégique que le premier n'est pas inscrit sur la liste électorale d'Hochelaga!

Ensuite, montrons que le contre-exemple échoue à bloquer la déduction en vertu de la transgression du PNNP. En d'autres mots, faisons la démonstration qu'une personne qui adhère aux prémisses de l'argument *et* à l'explication du contre-exemple doit tout de même déduire la conclusion C2. Force est de constater que si cette personne admet l'explication du contre-exemple, elle admet que ce contre-exemple mine P5 en isolation. Mais cela ne l'oblige pas à admettre de nouvelles prémisses si elle veut déduire C2. Il se trouve en effet que P5 fait partie d'un argument plus général. En effet, P4 nous fournit un contexte : l'argument ne

concerne qu'une circonscription où Québec solidaire a une chance de l'emporter, Hochelaga. Dans Hochelaga, Québec solidaire est dans la course et le Parti libéral, lui, n'a aucune chance de l'emporter. Des considérations sur le vote stratégique contre le Parti libéral sont donc impertinentes pour l'argument. Pour que le contre-exemple soit efficace, il faut supposer que l'électeur est dans une circonscription où Québec solidaire n'a aucune chance de l'emporter. Or, cela implique que l'électeur n'est pas dans Hochelaga (pour des raisons conjoncturelles, bien entendu) et donc que  $p$  est faux. Enfin, si  $p$  est faux,  $p \wedge q$  est faux, bien sûr. À cause de cela, le contre-exemple n'empêche aucunement la déduction de s'effectuer : le premier électeur peut tout à fait concevoir que P5 est faux en général sans pour autant délaissé son argument.

Cette intuition argumentative a des assises logiques solides. On peut, dans les faits, immuniser trivialement P5 contre le contre-exemple :

P4 : Demain, je vote et c'est dans Hochelaga que je peux exercer mon vote.  
 $(p \wedge Vh)$

P5\* : **Si c'est dans Hochelaga que je peux exercer mon vote, alors** si demain je vote, je dois voter pour le parti le plus progressiste.  $Vh \rightarrow (p \rightarrow r)$ <sup>26</sup>

P6 : Bien sûr, si je dois voter pour le parti le plus progressiste, je dois voter Québec solidaire, à moins qu'aucun candidat de cette faction ne se représente dans ma circonscription.  $((\forall x (Vx \rightarrow Qx)) \rightarrow (r \rightarrow s))$

P7 : Un candidat de Québec solidaire se présente dans Hochelaga.  $(Qh)$

C2 : Je dois donc voter Québec solidaire.  $(s)$

Le contre-exemple évoqué précédemment est évidemment inefficace pour réfuter P5\* puisqu'il stipule  $\neg Vh$ , alors que pour rejeter P5\*, il faut qu'il existe une assignation  $\alpha$  telle que  $\alpha(p) = \top$ ,  $\alpha(Vh) = \top$  et  $\alpha(r) = \perp$ . Autrement dit, pour que le contre-exemple ait réfuté

---

<sup>26</sup> On aurait pu aussi adopter la formulation « **Si c'est dans Hochelaga que je peux exercer mon vote et que** demain je vote, je dois voter pour le parti le plus progressiste  $(Vh \wedge p) \rightarrow r$  » (qui est l'équivalent logique de  $Vh \rightarrow (p \rightarrow r)$ ), mais la formulation conditionnelle met davantageusement en relief le processus d'immunisation.



P5\*, il aurait fallu que le premier électeur ne vote pas dans Hochelaga ( $\neg Vh$ ). Le second électeur ne peut donc pas espérer influencer son voisin puisque la déduction jusqu'à C2 est ininterrompue.

Le changement apporté à l'argument n'en altère pas du tout les tenants et aboutissants, mais il rend compte du fait que P5 s'inscrit dans un argument complet. Cette transformation précise P5. Il n'est pas nécessaire de poser  $p \rightarrow r$ ; seul  $(Vh \rightarrow (p \rightarrow r))$ , est suffisant. Bien entendu, il suffit uniquement que P6 soit vraie *dans le contexte de l'argument* pour permettre la déduction et c'est précisément ce qu'apporte l'ajout de la conditionnelle  $(Vh \rightarrow)$ .

Bien sûr  $(Vh \rightarrow (p \rightarrow r))$  est moins fort que  $p \rightarrow r$  : cela revient à affirmer  $p \rightarrow r$ , mais seulement pour un contexte restreint, c'est-à-dire les assignations où  $Vh$  est vrai. Plus généralement, si une proposition est vraie pour toute situation, elle est vraie dans une situation particulière. C'est la trivialité dont rend compte la transformation, laquelle est justement logiquement triviale :

1	$B$	Hypothèse
2	<div style="border-left: 1px solid black; padding-left: 5px; border-bottom: 1px solid black;"><math>A</math></div>	Hypothèse
3	<div style="border-left: 1px solid black; padding-left: 5px; border-bottom: 1px solid black;"><math>B</math></div>	Rétération 1
4	$A \rightarrow B$	Introduction de l'implication

Nous avons jusqu'à présent de bonnes raisons de maintenir le PNNP : un premier exemple éveille les intuitions sur lesquelles ce principe s'appuie et cette situation semble pouvoir se généraliser – ce que nous démontrerons bientôt – puisque c'est en vertu de propriétés logiques que l'argument peut être sauvé contre le type de contre-exemple qui transgresse le PNNP.

## 2. Deux exemples supplémentaires

Parce que nous faisons nôtre la mise en garde d'Ersnt Mach (1919) (l'intuition est faillible et n'a donc de valeur que dans un contexte familier) et pour satisfaire d'avance certaines critiques formulées à l'encontre des expériences de pensée (p. ex. Wilkes 1988, Peijnenburg et Atkinson 2003), nous fournirons deux exemples *réalistes* supplémentaires qui pourraient militer en faveur du PNNP.

Imaginons que Geneviève soit une Montréalaise abonnée à un programme de collecte de compost. Chaque semaine, elle accumule un certain nombre de rébus végétaux qu'elle dispose dans un bac qu'elle laisse dans la cuisine. Or, comme elle dispose d'un système de climatisation et qu'elle ouvre peu les fenêtres l'été, son appartement commence à empester. Son colocataire lui propose alors de laisser son bac dehors, puisque, prétend-il, les odeurs seront évacuées – ils sont, en effet, au 2<sup>e</sup> étage d'un appartement montréalais où la densité de population relativement faible permet de laisser des déchets dehors sans que les voisins en subissent les conséquences – en même temps qu'il n'y a aucun risque à l'y laisser puisque les animaux sont incapables d'ouvrir le bac en question (il dispose d'une petite serrure). Imaginons maintenant que Geneviève réponde que son colocataire a tort de lui proposer de laisser le bac à l'extérieur. Son argument, prétendrait-elle, ne fonctionnerait pas puisqu'une des prémisses sur lesquelles il repose – il est impossible pour un animal d'ouvrir le bac – est fausse : elle sait que certaines formes de singes et de pieuvres sont à même d'ouvrir ce genre de bac. Il est assez évident que son colocataire peut à juste titre se moquer de cette réponse puisqu'il n'y a aucune chance qu'un singe ou encore moins une pieuvre fasse irruption sur leur terrain. De manière plus générale, le contre-argument échoue parce qu'il enfreint le PNNP : même s'il réfute effectivement la prémisse isolée « il est impossible pour un animal d'ouvrir le

bac », le contre-exemple ne tient absolument pas compte du contexte et nie une autre des prémisses de l'argument en supposant implicitement que l'appartement ne se trouve pas à Montréal, mais dans un pays où des singes se promènent dans les rues ou, pire encore, sous les océans. Cette illustration peut sembler enfantine, mais c'est exactement de la même *structure d'erreur argumentative* que nous accuserons tout à l'heure ceux qui construisent des contre-exemples de Frankfurt indéterministes. Notons que le schéma argumentatif problématique employé dans la situation par Geneviève aurait été le même si elle avait mentionné des animaux habitant Montréal capables d'ouvrir le bac, mais dépourvus d'un tel désir. C'est dire qu'enfreindre le PNNP peut s'effectuer à des degrés de subtilité différents.

Un troisième exemple appuyant le PNNP, cette fois-ci tiré de la littérature philosophique, mérite d'être avancé. D'une part, le contre-exemple analysé ne semble pas, à première vue, contredire le PNNP et montre pour cette raison l'étendue de notre principe. D'autre part, David Copp utilise implicitement et naturellement le PNNP, ce qui appuie notre thèse qu'il s'agit d'un principe intuitif en plus de nous laisser croire qu'il est fréquemment mobilisé. Il s'agit d'un argument proposé par Copp (2006) pour prouver que le PAP peut être dérivé de la maxime kantienne *devoir implique pouvoir* :

P8 : Si quelqu'un est blâmable de *A*, alors il est mal de faire *A*.

P9 : S'il est mal de faire *A*, alors l'agent a l'obligation de ne pas faire *A*.

K : Si quelqu'un a l'obligation de ne pas faire *A*, alors il est capable de ne pas faire *A*.

---

PAP1<sup>27</sup> : Si quelqu'un est blâmable de *A*, alors il est capable de ne pas faire *A*.

L'on peut formaliser l'argument de la façon suivante :

P8 :  $BA \rightarrow MA$

---

<sup>27</sup> Il s'agit en fait d'une forme restreinte du PAP, lequel inclut également la responsabilité morale à propos des devoirs accomplis.

$$\begin{array}{l}
\text{P9 : } MA \rightarrow O \neg A \\
\text{K : } O \neg A \rightarrow P \neg A \\
\hline
\text{PAP1 : } BA \rightarrow P \neg A
\end{array}$$

Copp doit répondre à un contre-argument (*cf.* Copp 2006, section 7) contre P9 qui stipule qu'on peut très bien imaginer qu'il est mal de faire quelque chose ( $MA$ ) sans qu'on ait pour autant l'obligation de ne pas accomplir cette action ( $\neg O \neg A$ ). Dans un cas ( $\gamma$ ) où Madeleine ne pouvait pas s'empêcher de mentir, on pourrait dire qu'elle n'avait pas l'obligation de ne pas mentir ( $\gamma(O \neg A) = \perp$ ), mais que son action était mauvaise ( $\gamma(MA) = \top$ ). En effet, il est mal d'avoir menti parce que le mensonge est une sorte d'action qui est mauvaise en soi ou encore il est mal d'avoir menti parce que dans un monde parfait, personne ne mentirait. L'assignation  $\gamma$  constitue bien sûr un contre-exemple à P9. Voici la réponse de Copp :

The thought that Madeleine acted badly, or that there was something objectionable about her lying, or that lying is a wrong kind of act, does not support the idea that she is blameworthy for lying. (Copp 2006, p. 286)

Cette réponse semble au premier abord insatisfaisante, puisqu'elle ne paraît pas sauver P9. On voit mal en effet en quoi dire que Madeleine n'est pas blâmable dans le contre-exemple change quoi que ce soit au fait qu'elle a accompli une action mauvaise  $A$  sans avoir l'obligation de  $\neg A$ .  $MA$  reste vrai;  $O \neg A$  reste faux;  $MA \rightarrow O \neg A$  (P9) reste donc faux.

Néanmoins, si on analyse les propos de Copp dans le contexte de l'argument, on saisit très bien en quoi sa réponse bloque le contre-argument. Puisque l'argument principal vise à déduire  $BA \rightarrow P \neg A$ , il est évident qu'un contre-exemple qui impliquerait  $\neg BA$  n'empêchera pas la déduction d'être effectuée. Or, Copp pense précisément que le contre-exemple suppose  $\neg BA$ , c'est-à-dire que dans les cas où le mensonge de Madeleine n'est mal que parce que mentir est un mal en soi ou parce que dans un monde parfait le mensonge n'existerait pas, Madeleine n'est pas blâmable de ce mensonge.

Le réaménagement suivant de l'argument – le même que celui utilisé dans le cas du vote dans Hochelaga – montre bien que l'objection, si Copp a raison, ne fait pas mouche :

$$\begin{array}{l}
 \text{P8 : } BA \rightarrow MA \\
 \text{P9* : } \mathbf{BA} \rightarrow (MA \rightarrow O \neg A)^{28} \\
 \text{K : } O \neg A \rightarrow P \neg A \\
 \hline
 \text{PAP1 : } BA \rightarrow P \neg A
 \end{array}$$

Ce changement reste trivial. L'argument préserve sa validité; P9\* est toujours aussi intuitive, mais a l'avantage de préciser qu'elle ne concerne que les actes blâmables. Le contre-argument est alors inoffensif s'il suppose  $\neg BA$  (un contre-argument à P9\* devrait supposer l'assignation  $\delta$  telle que  $\delta(BA) = \top$ ,  $\delta(MA) = \top$  et  $\delta(O \neg A) = \perp$ ). Bref, il est tout à fait possible qu'une action mauvaise n'entraîne pas l'obligation de l'éviter, mais s'agit-il d'actions non seulement mauvaises, mais blâmables, cela est beaucoup moins clair.

Cela dit, peut-être que Copp se trompe et qu'en fait le contre-exemple n'implique pas que Madeleine n'est pas blâmable. Le contre-exemple serait dans ce cas compatible avec le fait que Madeleine soit blâmable (c'est-à-dire qu'il serait un contre-exemple à P9\*). Un contre-exemple à P9\* doit stipuler que Madeleine n'avait pas l'obligation de tenir sa promesse ( $\neg O \neg A$ ), que ne pas tenir sa promesse était mal ( $MA$ ) et que ne pas tenir sa promesse était blâmable ( $BA$ ). L'ajout de cette nouvelle condition ( $BA$ ) rend l'identification d'un contre-exemple beaucoup plus difficile. Il nous semble même que Copp aurait jusqu'à preuve du contraire, écarté l'objection à son argument. P9\* ne peut être rejetée qu'en vertu d'un contre-argument plus complet ou différent de celui proposé initialement.

À première vue, avons-nous dit, le contre-exemple proposé à Copp n'entre pas en conflit avec le PNNP (même s'il peut sembler contrevenir aux intuitions qui motivent le

PNNP) : en attaquant P9, il ne nie ni P8 ni K. En revanche, s'arrêter là serait faire l'économie d'un théorème auquel nous aurons à faire appel par la suite : le théorème de la déduction (sémantique)<sup>29</sup>. Celui-ci permet en effet de transformer trivialement tout argument dont la conclusion est une implication : l'antécédent de la conclusion devient la première prémisse et la conclusion est remplacée par le conséquent. Par exemple, un argument dont la conclusion serait « si seules sont bonnes les sensations, alors la mort n'est rien pour nous » peut être transformé trivialement en un argument dont la première prémisse affirme que *seules sont bonnes les sensations* et qui conclut que *la mort n'est rien pour nous*. Cela ne veut pas dire que les deux arguments sont identiques : le premier ne concerne que le lien entre deux thèses; le second affirme la deuxième thèse étant donné que la première est vraie. Cela signifie plutôt que si l'on accepte le premier cheminement logique, on ne peut pas rejeter le second<sup>30</sup>.

Tout argument de type

$$\frac{\begin{array}{l} p \rightarrow q \\ q \rightarrow r \end{array}}{p \rightarrow r}$$

peut donc être exprimé sous la forme

$$\frac{\begin{array}{l} p \\ p \rightarrow q \\ q \rightarrow r \end{array}}{r}$$

et vice versa. On pourrait nommer la première forme « conditionnelle »; la seconde, « affirmative ». La forme conditionnelle est moins forte que la forme affirmative : si celle-là

---

<sup>28</sup> On pourra préférer la forme  $(BA \wedge MA) \rightarrow O \neg A$ , ce qui revient au même, puisque  $[(A \wedge B) \rightarrow C] \equiv [(A \rightarrow (B \rightarrow C))]$ .

<sup>29</sup> Le théorème de la déduction affirme que  $\Gamma, A \vdash B$  si et seulement si  $\Gamma \vdash A \rightarrow B$ . La version sémantique de ce théorème pose que  $\Gamma, A \models B$  si et seulement si  $\Gamma \models A \rightarrow B$ .

affirme, par exemple, que *si Dieu n'existe pas, tout est permis*, celle-ci avance que *Dieu n'existe pas* et *que tout est permis*. Le tenant de la forme affirmative défend dans ce cas *et* le nihilisme, *et* l'athéisme, *et* leur lien; le tenant de la forme conditionnelle se contente de soutenir la vérité du lien entre les deux positions.

Celui qui défend la position affirmative défend donc nécessairement la position conditionnelle également, mais l'inverse n'est pas vrai. Conséquemment, (i) si une attaque fait mouche contre un argument de type conditionnel, son pendant affirmatif est aussi miné ( $\neg C \models \neg A$ ). Au contraire, (ii) si une attaque fait mouche contre un argument de type affirmatif, son pendant conditionnel n'est pas nécessairement miné ( $\neg A \not\models \neg C$ ) ou épargné ( $\neg A \not\models C$ )<sup>31</sup>. Par contraposition de (i), (iii) si une attaque ne fait pas mouche contre un argument affirmatif, elle ne fonctionne pas contre son pendant conditionnel ( $A \models C$ ). Et, bien sûr, par contraposition de (ii), (iv) si une attaque ne fait pas mouche contre un argument conditionnel, son pendant affirmatif n'est pas nécessairement miné ( $C \not\models \neg A$ ) ou épargné ( $C \not\models A$ )<sup>32</sup>.

Appliquons ces principes à l'exemple précédent. (i) Si on a raison de répliquer à la position conditionnelle qu'il peut exister une moralité sans Dieu parce qu'il est possible de *construire* une morale, la position affirmative tombe aussi. (ii) Si on a raison de répliquer à la position affirmative que Dieu existe, la position conditionnelle demeure intacte; et si on a

---

<sup>30</sup> Ce n'est pas par hasard si en déduction naturelle, les implications prennent la forme d'une conclusion identifiée sous une hypothèse affirmée.

<sup>31</sup> Si une attaque contre un argument affirmatif fonctionne, cette attaque menace soit la prémisse affirmative (normalement la première), soit une des prémisses conditionnelles. Si elle attaque la prémisse affirmative, alors l'argument conditionnel tient. Si elle attaque l'une des prémisses conditionnelles, alors l'argument conditionnel ne tient pas.

<sup>32</sup> Si une attaque contre un argument conditionnel ne fonctionne pas, alors soit cette attaque menace l'antécédent de la première implication (ce qui pourrait être la prémisse affirmative), soit non. (Par exemple, dans l'argument  $A \rightarrow B; B \rightarrow C \models A \rightarrow C$ , la potentielle prémisse affirmative serait  $A$ .) Si elle menace la potentielle prémisse affirmative, alors l'argument affirmatif est miné. Sinon, l'argument affirmatif est sain et sauf.

raison de répliquer à la position affirmative qu'il peut exister une moralité sans Dieu, la position conditionnelle est touchée. (iii) Si la tentative de réfuter la position affirmative par l'idée de morale construite échoue, cette tentative échoue aussi contre la position conditionnelle. (iv) Si c'est en réfutant l'existence de Dieu qu'on échoue à montrer que la position conditionnelle est fausse, la position affirmative est minée; enfin, si c'est par l'idée de morale construite qu'on échoue à montrer que la position conditionnelle est fausse, la position affirmative est préservée.

Dans le cas de Copp, la conjonction de (iii) (« si une attaque ne fait pas mouche contre un argument affirmatif, elle ne fonctionne pas contre son pendant conditionnel ») et du PNNP réfute le contre-exemple. En effet, une fois l'argument conditionnel de Copp transformé par le théorème de la déduction en version affirmative, il devient :

$P^* : BA$

$P8 : BA \rightarrow MA$

$P9 : (MA \rightarrow O \neg A)$

$K : O \neg A \rightarrow P \neg A$

---

$PAP1 : P \neg A$

Bien sûr, le contre-exemple mentionné plus haut, où  $BA$  est faux, appliqué à cette version de l'argument, contrevient directement au PNNP en contredisant  $P^*$  pour attaquer  $P8$  (dans le respect des deux clauses). En vertu de (iii), si une attaque ne fait pas mouche contre la version affirmative de l'argument de Copp, elle est tout aussi inefficace contre sa version conditionnelle. L'argument de Copp est donc sécurisé.

### 3. Caractère général du PNNP

Plus généralement, nous défendons la thèse selon laquelle pour tout contre-exemple qui attaque *une* prémisses et qui enfreint le PNNP, il existe une procédure d'immunisation logique



analogue à celle que nous avons présentée qui rend inefficace le contre-exemple. Ce résultat général empêche que nos exemples précédents soient considérés comme corroborant notre principe pour des raisons contingentes. La procédure est toujours appliquée de la même manière.

1. Soit un argument *valide*  $A$  tel que  $PA, PB \dots PN$  sont les prémisses et  $C$  est la conclusion. Soit  $B$ , un contre-exemple direct à une des prémisses  $PA, PB \dots PN$  (par exemple,  $PD$ ) qui suppose implicitement qu'une autre des prémisses  $PA, PB \dots PN$  (par exemple,  $PI$ ) est fausse<sup>33</sup>.
2. La prémisses directement attaquée par le contre-exemple est remplacée par une implication où la prémisses attaquée est le conséquent et où la prémisses niée pour ce faire est l'antécédent (par exemple  $PI \rightarrow PD$ ).
3. L'implication est simplifiée (remplacée par une formulation moins lourde qui occupe le même rôle argumentatif) si possible pour plus de clarté.
4. L'argument reste valide et ses prémisses sont compatibles avec l'argument donné.

Appliquons méthodiquement cette procédure aux deux exemples que nous avons précédemment formalisés :

### ***L'argument du vote dans Hochelaga***

P4 :  $p \wedge \forall h$

P5 :  $p \rightarrow r$

P6 :  $((\forall x (Vx \rightarrow Qx)) \rightarrow (r \rightarrow s))$

P7 :  $Qh$

---

C2 :  $s$

1. Le contre-exemple attaque P5 en montrant qu'il existe une assignation  $\alpha$  telle que  $\alpha(p) = \top$  et que  $\alpha(r) = \perp$ , mais pour ce faire il suppose que  $\alpha(\forall h) = \perp$ , ce qui fait que P4 est fausse par définition de la conjonction. Pourtant, celui qui élabore le contre-exemple ne veut pas attaquer P4 (le PNNP ne serait, dans ce cas, pas transgressé selon la seconde clause). Le PNNP est donc transgressé.
2. P5 est remplacée par P5\*\* :  $(p \wedge \forall h) \rightarrow (p \rightarrow r)$ .

---

<sup>33</sup> Comme nous l'avons vu précédemment, il est possible que la situation initiale soit un peu moins claire et qu'il faille procéder à des ajustements pour arriver à 1. On pense notamment à l'application du théorème de la déduction et des thèses (i)-(iv).

3. P5\*\* est simplifiée et devient P5\* :  $(\forall h \rightarrow (p \rightarrow r))$ .<sup>34</sup>
4. L'argument reste valide et aucune assignation proposée ne rend fausse l'une de ses prémisses.

### *L'argument de Copp*

P8 :  $BA \rightarrow MA$

P9 :  $MA \rightarrow O \neg A$

K :  $O \neg A \rightarrow P \neg A$

---

PAP1 :  $BA \rightarrow P \neg A$

1. Le contre-exemple attaque P9 en montrant que  $MA$  est vrai et  $O \neg A$  est faux, mais pour ce faire il suppose que  $BA$  est faux. Après l'application du théorème de déduction (et par (iii)), le PNNP est transgressé.
2. P9 est remplacée par P9\* :  $BA \rightarrow (MA \rightarrow O \neg A)$ .
3. P9\* est suffisamment simplifiée.
4. L'argument reste valide et aucune assignation proposée ne rend fausse l'une de ses prémisses.

Voici deux variantes supplémentaires qu'il est utile d'analyser pour mieux comprendre le fonctionnement du processus d'immunisation.

#### *Première variante de l'offensive :*

1. Le contre-exemple attaque P9 en montrant que  $MA$  est vrai et  $O \neg A$  est faux, mais pour ce faire il suppose que P8 est fausse. Le PNNP est donc transgressé.
2. P9 est remplacée par P9\*\* :  $(BA \rightarrow MA) \rightarrow (MA \rightarrow O \neg A)$ .
3. P9\*\* est simplifiée et devient P9\* :  $BA \rightarrow (MA \rightarrow O \neg A)$ .<sup>35</sup>
4. L'argument est probant : aucune assignation proposée ne rend une prémisse fausse ou n'altère sa validité.

#### *Deuxième variante de l'offensive :*

---

<sup>34</sup> En voici la preuve : par hypothèse,  $q$ . Par hypothèse,  $p$ . Donc,  $p \wedge q$  par introduction de la conjonction. Donc,  $p \rightarrow r$  par *modus ponens* parce que  $(p \wedge q) \rightarrow (p \rightarrow r)$  et  $p \wedge q$ . Donc,  $r$  par *modus ponens* parce que  $(p \rightarrow r)$  et  $p$ . Puisque sous l'hypothèse  $p$  on a  $r$ , alors on a  $p \rightarrow r$  par introduction de l'implication. Puisque  $p \rightarrow r$  sous l'hypothèse  $q$ , alors  $q \rightarrow (p \rightarrow r)$  par introduction de l'implication.

<sup>35</sup> La preuve se déroule ainsi : 1. Nous disposons de  $(BA \rightarrow MA) \rightarrow (MA \rightarrow O \neg A)$  et voulons obtenir  $BA \rightarrow (MA \rightarrow O \neg A)$ . 2. Par *reductio ad absurdum*, nous faisons l'hypothèse que  $BA$  et  $\neg (MA \rightarrow O \neg A)$ . Comme  $(BA \rightarrow MA) \rightarrow (MA \rightarrow O \neg A)$  et  $\neg (MA \rightarrow O \neg A)$ , par *modus tollens*, nous obtenons  $\neg(BA \rightarrow MA)$ . Par la règle de négation de l'implication, nous déduisons  $\neg MA$  de  $\neg(BA \rightarrow MA)$  et  $MA$  de  $\neg (MA \rightarrow O \neg A)$ , ce qui est absurde.

1. Le contre-exemple attaque P8 en montrant que  $BA$  est vrai et  $MA$  est faux, mais pour ce faire il suppose que P9 est faux. Le PNNP est donc transgressé.
2. P8 est donc remplacée par P8\*\* :  $(MA \rightarrow O \neg A) \rightarrow (BA \rightarrow MA)$ .
3. P8\*\* est simplifiée et devient P8\* :  $\neg MA \rightarrow (BA \rightarrow MA)$ <sup>36</sup>.
4. L'argument est probant : aucune assignation proposée ne rend une prémisse fausse ou n'altère sa validité.

Il ne devrait pas être étonnant que la procédure fonctionne puisque tout argument de la forme générale  $A1, A2, \dots, Ao \models B$  peut être soumis à la procédure avec succès. En effet,

1. Le contre-exemple attaque  $An$  et pour ce faire suppose  $\neg A6$ .
2.  $An$  est remplacé par  $A6 \rightarrow An$ .
3. Il est impossible de simplifier la formule.
4. L'argument est probant. En voici la preuve.

### **Première partie : la validité se transmet<sup>37</sup>**

Il faut d'abord prouver que si le premier argument est valide, le second l'est aussi. Plus précisément, il faut montrer que si  $A1, A2, \dots, Ao \models B$ , alors  $A1, A2, \dots, A6, A6 \rightarrow An, A7 \dots Am, Ao \dots \models B$ <sup>38</sup>.

1. Par définition, pour toute assignation  $a$ , si  $a(A1) = a(A2) = \dots = a(Ao) = \top$ , alors  $a(B) = \top$ .
2. Par *reductio ad absurdum*,  $a(A1) = a(A2) = \dots = a(A6) = a(A6 \rightarrow An) = a(A7) \dots = a(Am) = a(Ao) = \dots = \top$ , mais  $a(B) = \perp$ .
3. Par *modus ponens*, si  $a(A6) = \top$  et que  $a(A6 \rightarrow An) = \top$ ,  $a(An) = \top$ .
4. Par les règles de la conjonction, si  $a(A1) = a(A2) = \dots = a(A6) = a(A6 \rightarrow An) = a(A7) \dots = a(Am) = a(Ao) = \dots = \top$  et que  $a(An) = \top$ , alors  $a(A1) = a(A2) = \dots = a(Ao) = \top$ .

---

<sup>36</sup> La preuve se déroule ainsi : 1. Nous disposons de  $(MA \rightarrow O \neg A) \rightarrow (BA \rightarrow MA)$  et voulons obtenir  $\neg MA \rightarrow (BA \rightarrow MA)$ . 2. Par *reductio ad absurdum*, nous faisons l'hypothèse que  $\neg MA$  et  $\neg (BA \rightarrow MA)$ . 3. Comme  $(MA \rightarrow O \neg A) \rightarrow (BA \rightarrow MA)$  et  $\neg (BA \rightarrow MA)$ , nous obtenons  $\neg (MA \rightarrow O \neg A)$  par *modus tollens*. 4. Par la négation de l'implication, nous déduisons  $\neg MA$  de  $\neg (BA \rightarrow MA)$  et  $MA$  de  $\neg (MA \rightarrow O \neg A)$ , ce qui est absurde.

<sup>37</sup> Ce résultat peut aussi être obtenu par la version sémantique du théorème de la déduction.

<sup>38</sup> La notation a pour but de mettre en relief que nous ne postulons pas d'emblée que  $An$  est vraie pour effectuer la déduction. On peut faire une preuve encore plus simple en supposant que  $An$  est fausse. On remplace alors les étapes 4. et 5. par « comme  $a(An) = \top$  et  $a(An) = \perp$ , il y a une contradiction. CQFD.

5. Par *modus ponens*, comme  $a(A1) = a(A2) = \dots = a(Ao) = \top$  et si  $a(A1) = a(A2) = \dots = a(Ao) = \top$ , alors  $a(B) = \top$ , on obtient  $a(B) = \top$ . Mais, par hypothèse,  $a(B) = \perp$ . Donc, il y a une contradiction. CQFD.

***Deuxième partie : la vérité des prémisses se transmet***

Ensuite, il faut montrer que si toutes les prémisses du premier argument sont vraies, sauf  $An$  qui est fausse seulement si  $A6$  l'est aussi, toutes celles du second sont vraies. En fait, toutes les prémisses du deuxième argument sont identiques sauf : 1.  $An$  est absente, ce qui élimine la possibilité d'une prémisse fausse; 2.  $A6 \rightarrow An$  est présente. Comme toutes les autres prémisses sont identiques, elles sont forcément vraies dans les deux arguments. Il reste à voir si  $A6 \rightarrow An$  est vraie. À cause de la nature du contre-exemple, pour que  $An$  soit fausse, il faut que  $A6$  soit fausse aussi. Mais si  $A6$  et  $An$  sont fausses, alors  $A6 \rightarrow An$  est vraie. Et si  $An$  est vraie, alors  $A6 \rightarrow An$  est vraie. Donc  $A6 \rightarrow An$  est vraie. CQFD.

Conséquemment, si toutes les prémisses du premier argument sont vraies (sauf la prémisse qui est attaquée,  $An$ ), il en va de même de celles du second. Et comme en plus si le premier argument est valide, le second l'est (comme nous l'avons démontré), alors si le premier argument est probant, le deuxième l'est aussi (parce qu'un argument probant est un argument valide dont les prémisses sont vraies). La procédure d'immunisation est donc véritablement générale<sup>39</sup>.

---

<sup>39</sup> Avec de telles démonstrations, il peut sembler difficile d'admettre que le PNNP repose sur une position épistémologique comme le holisme. La procédure d'immunisation ne se fonde que sur les théorèmes de la logique classique et sur la définition d'un contre-exemple.

## 4. Réponse à quelques objections contre le PNNP

### A. L'utilité du PNNP

Trois contre-arguments possibles méritent d'être à présent examinés. Le premier consiste à remettre en question l'utilité du PNNP. Le principal argument en faveur de ce principe méthodologique, la procédure d'immunisation défendue plus haut, serait, selon cette façon de voir, tout à fait suffisante. Il serait en fait du devoir du philosophe d'immuniser ses prémisses à l'avance contre les attaques potentielles en les contextualisant. Ainsi, si un philosophe soutient que  $A \rightarrow B; B \rightarrow C \models A \rightarrow C$ , alors il doit user de prévoyance et modifier son argument ainsi :  $A \rightarrow B; A \rightarrow (B \rightarrow C) \models A \rightarrow C$  pour éviter un contre-exemple qui supposerait  $\neg A, B$  et  $\neg C$ . Or, il peut également le modifier ainsi :  $\neg C \rightarrow (A \rightarrow B); B \rightarrow C \models A \rightarrow C$  pour éviter un contre-exemple qui supposerait  $A, \neg B$  et  $C$ . À première vue, on pourrait penser que le PNNP a donc uniquement pour rôle d'alléger légèrement les arguments, mais que ce faisant on alourdit les exigences méthodologiques de la formation de contre-exemples.

Il en est tout autrement. Déjà, on peut se douter que les deux procédures d'immunisation sont mutuellement exclusives. Il faudrait que le philosophe soit en mesure de prévoir laquelle des attaques sera portée contre lui. Inspecter un autre exemple confirmera cette impression. Soit l'argument  $A; B \models A \wedge B$ . Nous avons deux méthodes mutuellement exclusives qui nous permettent de protéger notre argument : 1)  $A; A \rightarrow B \models A \wedge B$  et 2)  $B \rightarrow A; B \models A \wedge B$ . Nous disons qu'elles sont mutuellement exclusives parce que nous ne pourrions avoir 3)  $B \rightarrow A; A \rightarrow B \models A \wedge B$ . Cette dernière déduction est en effet invalide, contrairement à 1) et 2). Comment choisir entre ces dernières? Il peut être très difficile de le savoir à l'avance. Peut-être même que la question doit rester ouverte pour accommoder autant

ceux qui croient d'abord que *B* et ceux qui croient d'abord que *A*. C'est exactement ce que permet le PNNP. N'y a-t-il pas alors un risque que tantôt on immunise l'argument de la façon 1) et tantôt de la façon 2) de sorte qu'on se rendrait coupable de 3)? Autrement dit, le PNNP protège-t-il abusivement un argument qui devrait être abandonné? La deuxième clause prévient cette situation.

Nous en venons justement aux clauses. La clause (1) a simplement pour rôle de délimiter l'étendue de l'application du PNNP à un même argument. Nous pourrions très bien admettre qu'il est nécessaire de ne jamais supposer que tout énoncé que nous avons précédemment accepté soit nié dans la construction d'un contre-exemple. C'est d'une certaine manière ce que défendent les philosophes qui refusent les expériences de pensée qui impliquent le physiquement impossible. Cependant, pour les besoins de ce mémoire, nous nous contenterons d'un principe plus restreint, qui ne s'applique qu'aux prémisses d'un même argument.

Le rôle de la clause (2) (la deuxième prémisse *y* n'est pas réfutée elle aussi) mérite de plus amples précisions. Nous reviendrons au deuxième chapitre sur une critique de cette clause qui la juge trop restreinte (il faudrait plutôt qu'elle se lise (2)\* : « la deuxième prémisse n'est pas rejetée, indépendamment ou non »), mais nous nous attarderons ici essentiellement à en montrer l'utilité. Reprenons l'argument sur l'obligation de voter pour Québec solidaire. Pour que la clause (2) ne soit pas satisfaite, il faudrait que le contre-exemple soit, primo, destiné entre autres à montrer que le premier électeur ne votera pas dans Hochelaga, secundo, suffisant à le montrer. Dans ce cas, aucune des deux conditions n'est satisfaite, bien sûr : il n'y a ni intention de réfuter P4, ni succès à le faire. D'abord, le contre-exemple formulé (il existe des situations où il est préférable de ne pas voter pour le parti le plus progressiste) évoque des

intuitions morales pour invalider un principe moral, non pour accorder des caractéristiques factuelles au monde; ensuite, des considérations morales ne sauraient prouver un quelconque état du monde (nous en discuterons en II, 1). Il faut admettre que dans la plupart des cas, un contre-exemple qui est destiné à réfuter une prémisse a rarement pour fonction d'en réfuter une deuxième. Le cas échéant, on pourra généralement affirmer à juste titre qu'il y a en fait une erreur et que c'est la deuxième et non la première prémisse qui est visée. La clause (2) a donc une utilité très restreinte, mais tout de même existante comme l'atteste l'argument suivant :

P10 : *a* a toujours été généralement respectueuse de *b*.

P11 : *b* a toujours été généralement respectueuse de *a*.

P12 : si *a* et *b* ont été généralement mutuellement respectueuses, alors *a* et *b* pourraient être de véritables amies.

C3 : Donc, *a* et *b* pourraient être de véritables amies.

Or, il y a deux mois, *b* s'est rendue coupable d'un faux témoignage à l'endroit de *a*, ce qui a fait que *b* est actuellement emprisonnée. C'est que, *a* avait, quelques mois plus tôt, trahi la confiance de *b* d'une manière particulièrement cruelle. En vertu de cette considération nous rejetons P11. Supposons alors qu'on nous objecte qu'en vertu du PNNP (sans clause (2)), il faut ajuster P11 pour obtenir P11\* qui désigne  $P10 \rightarrow P11$ . Il faudrait alors identifier une assignation  $\varepsilon$  où  $\varepsilon(P10) = \top$  et  $\varepsilon(P11) = \perp$ . Notre contre-exemple, qui suppose que P10 est faux, ne fonctionnerait plus. Pareillement, si nous avions prétendu que notre contre-exemple réfutait P10, nous aurions fait face au même problème. Dans les deux cas, nous aurions pu par la suite nous attaquer à la prémisse restante (P10 ou P11), qu'on n'aurait pas pu, cette fois, réaménager pour qu'elle soit immunisée. Avec la clause (2), nous pouvons immédiatement considérer que le contre-exemple fonctionne (du moins selon les critères du PNNP) puisqu'il est à la fois suffisant et destiné à réfuter aussi bien la prémisse à laquelle il s'attaque

directement (P11) que celle qu'il nie pour ce faire (P10). Si l'on reprend le cas de  $A; B \models A \wedge B$ , on constate que (2) permet d'éviter que l'argument soit tantôt réaménagé comme en 2), tantôt comme en 1) contre celui qui voudrait aussi bien réfuter 1) que 2). En effet, si on présente un argument qui réfute  $A$  et  $B$  (ou un argument qui réfute  $A$  et un autre qui réfute  $B$ ), la clause (2) n'est pas satisfaite et l'on n'est donc pas contraint de ne pas nier  $A$  quand on réfute  $B$  et vice-versa. La relation d'immunisation entre les prémisses est donc non symétrique; les arguments sont comme des pions qui ne peuvent protéger leur protecteur. On se fraie un passage entre le Charybde de la circularité et le Scylla de l'oubli du contexte.

## **B. La légitimité dialectique de la construction du PNNP**

Une fois face à l'utilité démontrée du PNNP, il est possible de rétorquer qu'il est en réalité illégitimement défendu. Plus précisément, le deuxième contre-argument qui pourrait nous être opposé est que le PNNP a été tiré par induction à partir de seulement trois cas, ce qui bien sûr aurait été illégitime. Il ne s'agit pas d'induction pour au moins deux raisons. D'abord, il a été prouvé – et l'intuition y est pour peu – que la procédure d'immunisation était générale et c'est cette procédure qui constitue le principal argument en faveur du PNNP. L'écart entre la procédure d'immunisation et le PNNP était mince et nous croyons l'avoir adéquatement franchi dans la section précédente. Ensuite, le rôle de nos contre-exemples a été « d'élucider une certaine intuition conceptuelle et logique » (Schnall et Widerker 2012, p. 10). Nous avons vu dans les trois exemples que ce qui rendait précisément problématiques les contre-exemples était le *fait même* qu'ils fassent abstraction des autres prémisses de l'argument, le *fait même* qu'ils attaquent les prémisses en isolation. Nous employons un procédé dialectique similaire à celui de Van Inwagen (1983) lorsque, tentant de faire la démonstration de son principe de transfert  $\beta$  (si nous ne pouvons pas changer un événement qui en implique un second, alors



nous ne pouvons pas changer ce dernier), une variante de P2, il emploie deux exemples à la fois intuitifs et éloignés du débat sur la compatibilité entre le déterminisme et la responsabilité morale. À partir de la structure logique de ses exemples, Van Inwagen fait la démonstration que son principe est vrai *prima facie* et que le fardeau de la preuve repose du côté de ses adversaires. Cette stratégie semble tout à fait valable et demander à ce que le principe semble immédiatement intuitif même dans ses applications les plus problématiques constitue une exigence excessive (Schnall et Widerker 2012). Ces deux raisons (la procédure d’immunisation et le procédé dialectique similaire à celui de Van Inwagen) sont indépendantes et doivent être toutes les deux réfutées pour réfuter le PNNP.

### C. Les contre-exemples et leur nature

À partir du moment où notre procédé dialectique est légitime et utile, il reste trois méthodes pour invalider notre propos que nous ne ferons que mentionner brièvement. La première est de donner des raisons de rejeter le PNNP dans sa totalité, par un contre-exemple et des raisons adéquates, tâche que nous confierons à nos potentiels contradicteurs<sup>40</sup>. La seconde méthode est de proposer une version rivale du PNNP, qui capturerait nos intuitions mises en lumière dans

---

<sup>40</sup> Voici un exemple.

- i) Le Soleil approche de la Terre rapidement et inexorablement.
- ii) Les humains ne disposent pas de la téléportation.
- iii) Si le Soleil approche de la Terre rapidement et inexorablement et que les humains ne disposent pas de la téléportation, alors huit millions d’humains seront tués d’un coup.
- iv) Ce qui constitue l’identité personnelle des gens est un substrat mental cartésien nommé « âme » (chaque personne en a une).
- v) Si un humain meurt, alors ce qui constitue son identité personnelle s’éteint.
- vi) Donc, huit millions d’âmes seront tuées d’un coup.

Imaginons les expériences de pensée de téléportation de Parfit pour invalider intuitivement iv). Il semblerait bizarre dans cette situation de soutenir que ces expériences de pensée ne tiennent pas la route parce qu’elles contredisent la seconde prémisse. C’est peut-être parce que les situations que décrit Parfit ne sont pas tout à fait des contre-exemples, mais des illustrations de sa théorie. C’est d’ailleurs à peu près la réponse que formule Frankfurt à Widerker, réponse sur laquelle nous reviendrons.

les exemples précédents, mais qui ne serait pas suffisamment forte pour nous être utile dans notre prochain chapitre dans notre attaque des contre-exemples de Frankfurt<sup>41</sup>. C'est aussi une tâche que nous laisserons en plan<sup>42</sup>. La dernière méthode est de défendre une conception alternative ce que peut être un contre-exemple. Un contre-exemple ne serait pas forcément une assignation qui rend faux un énoncé, mais ce serait une mise en situation servant à éveiller des intuitions pour effectuer des distinctions conceptuelles ou appuyer des théories plus générales (Frankfurt 2003). Nous y reviendrons en III, 1.

Si nous nous sommes étendu longuement sur la défense du PNNP, c'est que nous sommes enclin à vibrer à l'unisson avec Lycan (2003), lorsqu'il avance que le fardeau de la preuve est toujours du côté de celui qui veut montrer que quelque chose est impossible (et on pourrait ajouter « interdit ») :

A theorist who maintains of something that is not obviously impossible that nonetheless that thing *is* impossible owes us an argument. And since entailment claims are claims of necessity and impossibility, the same applies to them. Anyone who insists that a sentence  $S_1$  entails another sentence  $S_2$  must defend that thesis if it is controversial. If I tell you that "Pigs have wings" entails "It snows every day in Chapel Hill", you need not scramble to show how there might be a world in which the first was true but the second false; rather, you would rightly demand that I display the alleged modal connection. (Lycan 2003, p. 109)

En effet, bien qu'elle porte sur le principe de transfert  $\beta$  de Van Inwagen, l'attaque de Lycan pourrait tout aussi bien concerner le PAP que le PNNP (qui est aussi une implication

---

<sup>41</sup> C'est exactement ce que tente McKenna (2008) avec l'argument de van Inwagen. Il s'agit de proposer une version du principe critiqué, version qui en capture les bénéfices et en rejette les inconvénients. Pour sa part, McKenna propose une version modifiée de  $\beta$  où le transfert de non-responsabilité ne peut « passer » par un agent fonctionnant normalement (McKenna 2008, p. 376), une suggestion qui semble plutôt *ad hoc*.

<sup>42</sup> Il s'agirait probablement d'ajouter une troisième condition au PNNP. D'emblée, nous devrions écarter toute forme de condition du genre « si la supposition de la négation de la prémisse  $y$  n'a qu'un rôle dialectique et donc non nécessaire », car si tel était le cas, il devrait être facile de modifier les conditions du contre-exemple. C'est en

matérielle). D'ailleurs, nous croyons que s'il est vrai que le fardeau de la preuve est du côté de celui qui veut montrer qu'un énoncé est impossible, il est également du côté de celui qui veut établir qu'un procédé dialectique est illégitime, surtout lorsque ce dernier a été fréquemment utilisé. Défendre le principe de non-négation des prémisses en long et en large était d'autant plus pertinent que le fardeau de la preuve est souvent présumé reposer du côté qui avantage le plus la position métaphysique de l'auteur (c.f. Lycan 2003, Van Inwagen 2008, lesquels défendent respectivement que le fardeau de la preuve est du côté des incompatibilistes et des compatibilistes).

---

y travaillant qu'on constate parfois que certaines suppositions qui nous semblaient tout à fait contingentes sont absolument nécessaires au contre-exemple.

## **Chapitre II : le caractère inoffensif des contre-exemples de Frankfurt contre l'argument de la conséquence**

Maintenant que nous disposons du principe de non-négation des prémisses (PNNP), qui nous oblige à analyser les prémisses d'un argument dans leur contexte, nous pouvons passer au cœur de notre propos. Ce chapitre se déroulera en deux temps : d'abord, nous montrerons que les contre-exemples de Frankfurt, pour répondre aux exigences du PNNP, ne doivent pas supposer l'indéterminisme; ensuite, nous analyserons l'accusation de pétition de principe contre la supposition du *déterminisme* dans un contre-exemple de Frankfurt et en concluons que cette accusation est juste. Nous pourrions alors conclure que les contre-exemples de Frankfurt doivent être rejetés.

### **1. L'illégitimité des contre-exemples de Frankfurt indéterministes**

#### **A. Immunisation du PAP contre un contre-exemple indéterministe**

Les contre-exemples de Frankfurt s'attaquent au principe des possibilités alternatives (PAP), lequel est présent dans plusieurs formes d'arguments incompatibilistes. Nous commencerons par nous attarder à l'argument contre la responsabilité morale (version affirmative) pour nous pencher ensuite sur l'argument incompatibiliste à proprement parler (version conditionnelle). Tout argument qui suppose le déterminisme et le PAP sera concerné par notre analyse. Revoici l'argument contre la responsabilité morale :

P1 : Le déterminisme est vrai, ce qui signifie que la conjonction des lois de la nature et de l'état du monde à un temps  $t$  implique l'état du monde à un temps  $t +/ - n$ , incluant chacune de mes actions. ( $D$ )

P2 : Si mes actions sont impliquées par une chose sur laquelle je n'ai aucun pouvoir, alors je n'aurais pas pu agir autrement.

P3 : Je n'ai de pouvoir ni sur un état du monde précédant ma naissance ni sur les lois de la nature.

PAP : Je suis moralement responsable d'une action seulement si j'avais pu agir autrement. ( $RA \rightarrow PA$ )

---

C1 : Je ne suis responsable d'aucune de mes actions. ( $\neg RA$ )

Selon le PNNP, il serait absurde de remettre en question l'une ou l'autre des prémisses P2, P3, ou le PAP sur la base d'un contre-exemple où le déterminisme serait faux (c'est-à-dire sur la base de  $\neg D$ ), à moins de *réussir* du même coup à remettre en question P1 (selon la clause (2) du PNNP)). Donc, à partir du moment où P1 est acceptée ou supposée, les autres prémisses ne peuvent être contestées par un contre-exemple déterministe. Le contre-exemple auquel nous nous intéresserons est bien sûr celui de Frankfurt contre le PAP. En suivant la procédure d'immunisation que nous avons exposée précédemment, nous devrions effectuer les aménagements suivants :

1. Le contre-exemple attaque PAP en montrant que  $RA$  est vrai et  $PA$  est faux, mais pour ce faire il suppose que  $D$  est faux. Après l'application du théorème de déduction, Le PNNP est transgressé.
2. PAP est remplacée par PAP\* :  $D \rightarrow (RA \rightarrow PA)$ . PAP devient donc

PAP\* : **Si le déterminisme est vrai**, alors je suis moralement responsable d'une action seulement si j'avais pu agir autrement.

3. PAP\* est suffisamment simplifiée.
4. L'argument est probant : aucune assignation proposée ne rend une prémisse fausse ou n'altère sa validité.

De cette manière, on réitère que le PAP n'a à être défendu que dans un contexte déterministe.

S'il est vrai que les contre-exemples de Frankfurt nécessitent une situation dans laquelle un

agent dispose du libre arbitre au sens où certains de ses actes ou choix ne sont pas déterminés, alors ces contre-exemples ne réfutent en aucun cas l'argument précédent, à moins qu'ils donnent des raisons suffisantes de rejeter le déterminisme (ils mineraient P1 et ne satisferaient pas la deuxième condition du PNNP). Par ailleurs, il est possible qu'ils soient assez puissants pour réfuter le PAP dans un contexte non déterministe, même s'ils sont inoffensifs contre l'argument de la conséquence.

Voyons ce qu'il en est de l'argument en faveur de l'incompatibilisme à proprement parler. L'argument est similaire, mais ne se prononce pas sur la vérité du déterminisme (c'est sa version conditionnelle) :

P2 : Si mes actions sont impliquées par une chose sur laquelle je n'ai aucun pouvoir, alors je n'aurais pas pu agir autrement.

P3 : Je n'ai de pouvoir ni sur un état du monde précédant ma naissance ni sur les lois de la nature.

PAP : Je suis moralement responsable d'une action seulement si j'avais pu agir autrement. ( $RA \rightarrow PA$ )

---

C1\* : Si le déterminisme est vrai (anciennement, P1), je ne suis responsable d'aucune de mes actions (anciennement C1).

On peut tout de suite affirmer que le résultat sera le même que dans l'argument contre la responsabilité morale. En effet, nous avons tout à l'heure  $P1, P2, P3, PAP \models C$ ; nous avons ici  $P2, P3, PAP \models P1 \rightarrow C1$ , ce qui est permis trivialement par la version sémantique du théorème de déduction, comme nous l'avons vu précédemment. Ainsi que nous l'avons soutenu, si un contre-exemple ne remet pas en cause le premier argument, il ne devrait pas non plus remettre en cause le second en vertu de (iii) : si une attaque ne fait pas mouche contre un argument affirmatif, elle ne fonctionne pas contre son pendant conditionnel. Il y a en effet moins de points objectifs (c'est-à-dire de cibles que l'on peut atteindre) à un argument qui affirme moins. Cela est évident dans le cas présent : le contre-exemple de Frankfurt, s'il suppose un

monde non déterministe, remet P1 en cause, comme dans l'argument contre la responsabilité morale. Or, supposer que P1 est faux revient à dire que l'implication  $P1 \rightarrow C1$  est trivialement vraie et donc que  $C1^*$  est vrai. Par ailleurs, il est possible d'effectuer la même modification triviale que celle utilisée précédemment en modifiant le PAP pour obtenir le PAP\* : **si le déterminisme est vrai, alors** je suis moralement responsable d'une action seulement si j'avais pu agir autrement. Le PAP\* est immunisé contre les contre-exemples de Frankfurt, *si ceux-ci présupposent une forme d'indéterminisme*.

## **B. Satisfaction des clauses du PNNP : défense de (2)**

Pour que les contre-exemples de Frankfurt évitent d'enfreindre le PNNP, il reste donc seulement deux issues : la première, c'est évidemment qu'ils ne supposent pas que le déterminisme est faux – nous y consacrerons l'essentiel de ce chapitre; la seconde, c'est qu'ils ne remplissent pas l'une des clauses du PNNP.

Montrons tout de suite que cette dernière avenue est incontestablement un cul-de-sac. D'un côté, il est inutile de s'éterniser sur la première clause puisqu'elle est d'emblée satisfaite : le principe des possibilités alternatives et le déterminisme sont deux prémisses d'un même argument. D'un autre côté, pour éviter d'être soumis au PNNP grâce à sa seconde clause, les contre-exemples de Frankfurt devraient réfuter aussi bien le PAP que la prémisse à laquelle ils s'attaquent implicitement, à savoir P1 (le déterminisme est vrai). De fait, on pourrait très bien envisager qu'un philosophe qui rejette le déterminisme accepte en plus les contre-exemples de Frankfurt de manière à rejeter les deux premières prémisses de l'argument de la conséquence. Ce n'est sans doute pas le cas de Frankfurt, puisque celui-ci soutient que *peu importe si nous vivons dans un monde déterministe, nous disposons de notre responsabilité morale*. Or, Frankfurt défend une vision compatibiliste, pas libertarienne.

Imaginons qu'un libertarien dise :

J'adhère au PNNP et je crois que les contre-exemples de Frankfurt sont efficaces. Ces derniers n'enfreignent pas les conditions du PNNP parce qu'en vertu de la seconde clause, je peux légitimement élaborer un contre-exemple au principe des possibilités alternatives tout en niant le déterminisme puisque je le rejette indépendamment.

Ce serait mal comprendre la seconde clause du PNNP. Celle-ci ne stipule pas que si on souhaite attaquer le PAP, on doit le faire sans supposer que le déterminisme est faux, *à moins qu'on souhaite rejeter le déterminisme...* mais plutôt, *à moins que le contre-exemple avancé réfute la première prémisse du même coup*. Autrement dit, pour qu'un contre-exemple soit soustrait à la sanction du PNNP en ne satisfaisant pas la clause (2), il faut que le contre-exemple *réfute bel et bien* la prémisse qui est niée. Et, bien sûr, si une expérience de pensée controversée met indirectement et involontairement en doute le déterminisme, il serait passablement absurde d'en déduire des raisons de rejeter le déterminisme. Voici un exemple qui illustre bien la pertinence que la clause (2) soit ainsi formulée et appliquée :

IMH : S'il est impossible d'avoir à voler pour survivre et que voler est toujours mal, alors un homme vertueux ne devrait jamais voler  $((p \wedge q) \rightarrow r)$ .

IV : Dans nos sociétés interventionnistes contemporaines, il est impossible d'avoir à voler pour survivre  $(s \rightarrow p)$ .

MV : Voler est toujours mal  $(q)$ .

---

HM : Un homme vertueux ne devrait jamais voler dans nos sociétés interventionnistes contemporaines  $(s \rightarrow r)$ .

Supposons que l'on veuille contester MV par le contre-exemple commenté suivant :

*Aladin est dans une situation où il s'apprête à mourir de faim, dépourvu qu'il est d'un quelconque moyen de se procurer de la nourriture sans voler. Il y a des pommes sur un étalage, qui lui permettront de survivre s'il les dérobe et les mange. Aladin s'exécute.* Notre intuition nous mène à croire qu'Aladin n'a pas



mal agi. Nous en déduisons qu'il est faux de prétendre que voler est toujours mal.

Il semble très intuitif de considérer que ce contre-argument n'est pas une attaque sérieuse à l'argument dans son ensemble. En effet, IV stipule l'impossibilité de ce que le contre-exemple suppose : la possibilité qu'il soit nécessaire de voler pour survivre. La version du PNNP que nous avons proposée est justement transgressée par ce contre-exemple, alors que ne le serait pas une version du PNNP qui aurait pour deuxième clause « à moins qu'on souhaite rejeter la prémisse implicitement niée ». En effet, le contre-exemple n'est pas suffisant pour rejeter l'idée selon laquelle il est impossible de devoir voler pour survivre dans notre société contemporaine interventionniste et pourtant c'est une idée que beaucoup de gens seraient prêts à nier (à moins d'admettre une définition plus restreinte du vol ou de la propriété). C'est dire que la stratégie argumentative employée par le réfuteur de l'argument était mal orientée : il aurait été possible de montrer qu'il existe des situations où voler est nécessaire pour survivre ou encore de faire la démonstration de l'existence de circonstances qui n'ont rien à voir avec la survie et où le vol est tout de même moral. Les intuitions qui ont motivé la formulation du PNNP sont précisément celles qui sont mobilisées ici. Il semble donc qu'un tel exemple discrimine efficacement une interprétation de la seconde clause du PNNP qui aurait pu épargner l'utilisation des contre-exemples de Frankfurt par un libertarien.

### **C. Satisfaction des clauses du PNNP : l'impossibilité de réfuter le déterminisme par un contre-exemple de Frankfurt**

Supposons maintenant que le libertarien réponde par ce que nous nommerons « l'argument à partir des contre-exemples de Frankfurt » :

Si je comprends bien la seconde clause, il faut que les contre-exemples de Frankfurt soient eux-mêmes suffisants pour réfuter le déterminisme pour qu'ils puissent le nier dans leur contre-exemple au PAP? Très bien. *A priori*, j'ignore si le déterminisme est vrai.

Mon intuition morale m'amène à croire que

(a) Les contre-exemples de Frankfurt indéterministes réfutent le PAP et donc l'argument de la conséquence.

Par ailleurs, je sais que

(b) Le PNNP est vrai et que si le PNNP est vrai, les contre-exemples de Frankfurt indéterministes ne peuvent pas réfuter l'argument de la conséquence sans réfuter le déterminisme.

Donc, je sais que

(c) Le déterminisme est faux.<sup>43</sup>

Dans cette situation, les contre-exemples de Frankfurt indéterministes semblent suffisants pour réfuter le déterminisme. Il faut bien sûr y adjoindre le PNNP, mais comme il s'agit d'un principe logico-méthodologique, cette jonction ne change pas grand-chose. En effet, si je soutiens que le fait qu'il neige en avril est suffisant pour infirmer l'énoncé « il ne neige que l'hiver », évidemment, cela nécessite la jonction de principes logiques comme le tiers exclu et pourtant il s'agit d'un exemple banal de suffisance.

Pourtant, force est d'admettre que cette position dialectique est plutôt problématique : par leur nature même, les contre-exemples de Frankfurt sont inadéquats pour réfuter le déterminisme, car ils constituent seulement des outils pour mettre en évidence des intuitions morales, alors que le déterminisme est un énoncé strictement factuel. Il est difficile de voir en quoi une réflexion sur les « ingrédients » de la responsabilité morale – les contre-exemples de Frankfurt visent à nous montrer que la responsabilité morale ne dépend *pas* des possibilités alternatives – pourrait nous aider à comprendre la relation de dépendance entre les événements du monde (le déterminisme).

Pour bien comprendre le genre de problème soulevé par l'argument à partir des contre-exemples de Frankfurt, qui est au demeurant logiquement valide, il peut être utile de regarder un autre argument qui entend déduire un énoncé descriptif sur le monde à partir d'intuitions morales. L'argument à partir de la responsabilité (*the argument from moral responsibility*), que Maier analyse et rejette (2013), s'énonce ainsi :

MR : Il existe des agents qui sont moralement responsables.

PAD : Les agents sont moralement responsables seulement si le déterminisme est faux.

(c) : Le déterminisme est faux.<sup>44</sup>

Maier montre que cet argument ne satisfait pas le critère mooréen (1939) selon lequel les prémisses d'un bon argument doivent être connaissables en même temps. Il serait impossible de savoir que les agents sont responsables *et* qu'ils ne le sont que si le déterminisme est faux sans avoir des raisons *indépendantes* de croire que le déterminisme est faux. C'est ce que l'on nomme « schéma de dépendance de l'information » (*information-dependence template*) (Wright 2003) : l'information sur laquelle repose l'indice (la preuve, l'*evidence*, ce qui nous donne des raisons de croire en la prémisse) de la première prémisse requiert que la conclusion ait un indice indépendant de l'indice et de la prémisse. Autrement dit, pour accepter ce genre d'argument, il faudrait avoir des raisons préalables de croire à sa conclusion. Il y a un consensus sur le fait que tout argument qui correspond à ce schéma échoue à transmettre la justification des prémisses à la conclusion (Moretti et Piazza 2013).

---

<sup>43</sup> Si un libertarien défend un tel argument, c'est que nous aurons plus que réussi à défendre le PNNP!

<sup>44</sup> Lycan (2003), par exemple, avance un argument de ce type qu'il qualifie explicitement de « mooréen ». Il argumente même qu'aucune preuve déductive ne peut entendre faire mieux que nous inviter à comparer la plausibilité des deux « facettes » d'un argument. À propos d'un argument valide, est-ce la vérité de la conclusion qui découle de celle des prémisses ou plutôt la fausseté d'une prémisse qui découle de la fausseté de la conclusion? Les problèmes de transmission de la justification remettent en cause la position de Lycan.

Maier croit qu'il y a trois vertus épistémiques qu'on ne peut exercer simultanément tout en endossant cet argument dans son intégralité<sup>45</sup>. Nous nous contenterons ici d'évoquer ce qu'il nomme « la priorité de l'excuse » et qui constitue pour lui une raison indépendante des trois vertus de rejeter l'argument à partir de la responsabilité morale. La priorité de l'excuse commande que si je sais, par exemple, que quelqu'un ne peut être tenu responsable d'un meurtre s'il a moins de huit ans, alors je ne peux pas savoir que quelqu'un est responsable d'un meurtre *tant et aussi longtemps que je n'ai pas l'assurance qu'il a plus de huit ans*<sup>46</sup>. Dédire de la responsabilité d'un agent que les conditions de possibilité de sa responsabilité

---

<sup>45</sup> Voici les trois vertus en question. La *particularité* suppose que pour savoir qu'il existe des agents qui sont moralement responsables (MR), je dois connaître certains exemples d'agents responsables. Ensuite, l'*autorité* stipule que pour que je sache que certains agents sont responsables, je dois savoir que certains agents (pas nécessairement les mêmes) sont responsables sur une base partiellement non testimoniale. Plus généralement, je ne peux pas avoir de connaissance de certaines données morales sans avoir jamais pu moi-même appliquer mes intuitions morales à des cas particuliers avec lesquels j'étais en contact. À ce point-ci, donc, pour savoir qu'il existe des agents moralement responsables (MR), il faut que je sache sur une base partiellement non testimoniale que certains agents sont, dans certaines situations, responsables. Enfin, on ajoute l'*humilité*, qui stipule que si je sais que les agents sont moralement responsables seulement si le déterminisme est faux (PAD), cela implique que je ne peux pas savoir que certains agents sont responsables dans certaines situations sur une base partiellement non testimoniale sauf si je sais que le déterminisme est faux sur une base indépendante.

<sup>46</sup> Soit l'argument :

S3 : Tuer avec un revolver demande de la préméditation qui est généralement incompatible avec un accident. J'apprends que x a utilisé un revolver pour tuer quelqu'un.

**Ce qui me justifie de croire que**

MR2 : Un individu x est moralement responsable d'un meurtre au revolver.

**Ce qui implique, parce que je sais que IC1 (Si un individu est responsable d'une action grave, alors il a plus de huit ans), que**

CF1 : L'individu x a plus de huit ans.

Si c'est sur la base de S3 que j'accepte MR2, alors je ne peux pas accepter CF1 sur la seule base de 1. (la justification S3 de MR2) et de 2. (le fait que je sais que MR2 implique CF1 grâce à IC1). En effet, pour que j'accepte une thèse à propos d'intuitions morales, je dois supposer soit que les faits du monde qui pourraient mettre à mal ces intuitions ne sont pas des faits de notre monde, soit qu'ils ne sont pas connectés à ces intuitions. Si je suppose qu'ils ne sont pas connectés à ces intuitions, alors cela veut dire que la vérité de mes intuitions morales n'a aucune condition de possibilité factuelle. Je ne peux donc pas accepter l'implication de mes intuitions à la négation d'un fait du monde (la conclusion est donc bloquée). Si je suppose qu'ils ne sont pas des faits de ce monde, comme c'est précisément ce que je tente de démontrer, je suis coupable d'une pétition de principe. Ainsi, dans l'exemple, je déduis du fait qu'une personne est responsable qu'elle a les conditions de possibilité pour être responsable. On déduit de la présomption de la responsabilité les conditions de possibilité de cette responsabilité! Cela paraît sans contredit illégitime.

sont mises en place, c'est procéder à l'envers; il faut constater que les conditions de possibilité de la responsabilité sont mises en place pour *ensuite* l'attribuer à un agent<sup>47</sup>.

Dans l'argument à partir des contre-exemples de Frankfurt, il n'est pas question de priorité de l'excuse, mais comme dans le cas de l'argument à partir de la responsabilité morale, il semble y avoir un problème de *transmission de la justification*. La justification des prémisses d'un argument se transmet dans les conditions suivantes :

1. Le sujet connaissant a une justification de croire  $p$ .
2. Le sujet connaissant sait que  $p \rightarrow q$ .
3. Le sujet connaissant est justifié de croire  $q$  en vertu de 1 et 2<sup>48</sup>.

Supposons que je vois un thermomètre par la fenêtre et qu'il indique  $-4^{\circ}\text{C}$  ( $p$ ). J'ai de bonnes raisons de croire qu'il est fonctionnel et donc de croire qu'il fait effectivement  $-4^{\circ}\text{C}$  (indice que  $p$ ). Je sais aussi que s'il fait  $-4^{\circ}\text{C}$ , l'eau du petit étang dans le parc gèle ( $p \rightarrow q$ ). Je suis donc justifié de croire que  $q$  sur la base de 1. (je suis justifié de croire que  $p$ ) et 2. (je sais que  $p \rightarrow q$ ). La justification se transmet.

Traditionnellement, l'absence de transmission de la justification (bien que cela soit débattu (c.f. Pryor 2004)) est attribuée à l'argument de Moore (1939) et à des arguments de forme similaire :

Mes impressions sensibles me justifient de croire que

(d) J'ai deux mains.

Or, je sais que

(e) Si j'ai deux mains, il existe un monde extérieur à moi.

---

<sup>47</sup> Est-ce à dire que la justification ne se transmet pas dans l'argument suivant?

P : X est moralement responsable d'un meurtre (justification : j'ai lu dans les journaux qu'un juge l'avait jugé coupable et que la responsabilité criminelle correspond assez bien à la responsabilité morale).

Q : Donc, X a plus de huit ans (justification : si je suis moralement responsable, j'ai plus de huit ans).

En fait, le problème d'un tel argument c'est que P est un détour entre sa justification et Q. C'est le *fait* qu'un juge ait jugé responsable X qui me donne une raison adéquate de penser que X avait plus de huit ans et non la responsabilité de X. Je ne déduis donc pas un fait à partir d'un énoncé strictement normatif.

<sup>48</sup> Nous éludons volontairement ici la question de savoir si l'argument à partir des contre-exemples de Frankfurt pourrait *renforcer* la conviction d'un libertarien, conviction déjà appuyée par des indices indépendants.

Je suis donc justifié de croire que  
(f) Un monde extérieur existe.

L'argument de Moore est valide, on peut très bien vouloir accepter (d) et (e), mais la justification ne se transmet pas (Wright 2003). Il est difficile d'admettre que je puisse à la fois être justifié de croire que j'ai deux mains et ignorer *a priori* si le monde extérieur existe. Autrement dit, il s'agit d'une forme de pétition de principe puisque pour accepter (d) (c'est-à-dire pour accepter que mes impressions cognitives me justifient à croire que j'ai deux mains), je dois supposer qu'il existe un monde extérieur (f). Je suis justifié d'accepter que j'ai deux mains, mais *parce que cet énoncé* se justifie en présumant (comme nous le faisons habituellement) qu'il existe un monde extérieur à moi, il ne m'est pas permis d'en déduire qu'il existe en effet un monde extérieur.

Formulons l'argument à partir de la responsabilité morale de manière à pouvoir nous poser la question de la transmission de la justification :

S1 (Indice de MR) Mon intuition morale

***Me justifie de croire que***

MR : Il existe des agents qui sont moralement responsables.

***Ce qui implique, parce que je sais que PAD*** (Les agents sont moralement responsables seulement si le déterminisme est faux), ***que***

(c) : Le déterminisme est faux.

Si la justification se transmet, cela veut dire que 1. S1 me justifie de croire MR; que 2. Je sais que MR implique (c) (parce que je sais que PAD); et que 3. (c) est justifié sur la base de 1. et de 2.

Voici pourquoi la justification ne se transmet pas. Pour être justifié de croire, grâce à mes intuitions morales, que je suis moralement responsable, je dois supposer (naturellement et

implicitement) qu'aucun fait du monde n'élimine ma responsabilité morale<sup>49</sup>. Autrement dit, pour que je sois justifié d'accepter la thèse MR, je dois supposer *soit* que les énoncés qui décrivent des faits du monde qui pourraient éliminer ma responsabilité morale sont faux (y compris le déterminisme), *soit* que ces mêmes faits du monde ne sont pas, en réalité, connectés à ma responsabilité morale (je dois donc endosser  $\neg$  PAD). Bref, pour que je sois justifié d'accepter MR, je dois supposer implicitement soit que le déterminisme est faux (c), soit que l'incompatibilisme est faux ( $\neg$  PAD). En effet, comment être persuadé que nous sommes dotés de responsabilité morale si ce n'est par l'une de ces deux stratégies, à savoir avancer que la responsabilité morale est attribuable indépendamment de faits sur la structure du monde (comme le font la plupart des compatibilistes) ou encore essayer de montrer que le déterminisme est faux ou possiblement faux (comme le font les libertariens)?

Dans le premier cas, c'est-à-dire si je présume que l'incompatibilisme est faux ( $\neg$  PAD) pour appuyer MR, mes deux justifications sont contradictoires : l'acceptation de la première prémisse repose sur le rejet de ce qui permet de relier MR et (c), c'est-à-dire PAD<sup>50</sup>. Cela ressemble à s'y méprendre à une transgression du PNNP : il ne faut pas nier une prémisse pour en soutenir (plutôt qu'attaquer) une autre. Dans le deuxième cas, c'est-à-dire si je présume que le déterminisme est faux (c) pour accepter ma première prémisse, alors je fais une pétition de principe puisque pour justifier ma première prémisse, je suppose implicitement que ma conclusion est vraie.

---

<sup>49</sup> Fait important : cela est vrai peu importe la définition que l'on donne à la responsabilité morale. En effet, si par exemple la responsabilité morale est définie comme une forme d'illusion de contrôle, cela ne change rien au prérequis qu'aucun fait concernant la nature ne doit empêcher la responsabilité morale. Si un fait naturel venait à rompre l'illusion de contrôle, nous ne pourrions pas dire que nous sommes moralement responsables et donc la relation de justification serait brisée.

<sup>50</sup> Ce schéma correspond au schéma de contradiction de 1. et 2. (Moretti et Piazza 2013).

Bref, la transmission de la justification ne s'effectue pas puisque 1. et 2. sont satisfaits, mais pas 3. Je ne peux, en effet, être justifié de croire que le déterminisme est vrai sur la seule base de MR et des deux relations de justification. Personne ne peut donc légitimement croire que le déterminisme est faux sur la seule base de sa croyance en la responsabilité morale et en l'incompatibilisme. Conséquemment, pour qu'un agent soit justifié de croire (c) dans le cadre de cet argument, il doit exister des raisons indépendantes des considérations invoquées par cet argument, ce qui correspond, comme nous l'avons mentionné, au schéma de dépendance de l'information<sup>51 52</sup>.

Il en va de même pour l'argument à partir des contre-exemples de Frankfurt. Il y a un problème de transmission de la justification, c'est-à-dire que je peux avoir des raisons de croire que (a) les contre-exemples de Frankfurt réfutent l'argument de la conséquence; je peux aussi savoir que les contre-exemples de Frankfurt ne peuvent pas réfuter l'argument de la conséquence sans réfuter le déterminisme ((a) → (c)) (l'équivalent logique de ¬ (a) ∨ (c)) et pourtant ne pas être justifié de croire que (c). L'argument est exactement le même pour l'argument à partir des contre-exemples de Frankfurt :

S1 : Mon intuition morale

*Me justifie de croire que*

(a) : Les contre-exemples de Frankfurt fonctionnent.

---

<sup>51</sup> Si on avait plutôt P1, PAD ⊨ ¬ RM, pourrait-on reprocher la même faille à l'argument? Pour être certain de la vérité du déterminisme (P1), il serait pour le moins curieux d'exiger que l'état du monde n'ait aucun impact sur la responsabilité!

<sup>52</sup> Une autre façon de comprendre le problème est à partir de certains principes de limitation dans l'élaboration d'arguments visant à résoudre un doute : si un doute à propos de la conclusion requiert rationnellement l'acceptation d'une proposition contradictoire avec les raisons de croire aux prémisses, la justification ne se transmet pas (Davies 2009). C'est précisément ce qui se passe dans l'exemple : si on ne sait pas si le déterminisme est vrai, alors la seule façon de justifier MR est de postuler que l'incompatibilisme est faux, ce qui empêche d'accepter les deux prémisses. Tout argument qui satisfait le schéma de dépendance à l'information satisfait également l'antécédent de ce principe limitatif. La réciproque est vraie si l'on accepte une forme de conservatisme épistémologique (Davies 2009), c'est-à-dire qui suppose qu'avoir une justification *prima facie* pour croire une proposition requiert d'avoir une justification antécédente de croire que ce qui mine la proposition est faux.



*Qui, je le sais, impliquent, parce que* (b) le PNNP est vrai et parce que s'il est vrai les contre-exemples de Frankfurt peuvent seulement fonctionner s'ils réfutent le déterminisme ( $PNNP \wedge (PNNP \rightarrow (\neg CEP \vee C))$ ), *que*

(c) : Le déterminisme est faux.

Pour que je puisse me fier à mes intuitions morales, je dois soit implicitement faire la supposition que tous les énoncés de faits du monde qui les bloqueraient sont faux, soit implicitement faire la supposition qu'aucun fait du monde ne peut bloquer mes intuitions morales qui appuient les contre-exemples de Frankfurt. Dans le premier cas, c'est-à-dire si je fais implicitement la supposition que tous les faits du monde qui pourraient bloquer mes intuitions morales sont faux, alors j'effectue une pétition de principe. En effet, pour prouver que le déterminisme est faux, je suppose que des faits du monde qui pourraient bloquer ma première prémisse, comme le déterminisme, sont faux. Dans le deuxième cas, c'est-à-dire si je fais implicitement la supposition qu'aucun fait du monde ne peut bloquer mes intuitions morales qui appuient les contre-exemples de Frankfurt, alors j'affirme la compatibilité des contre-exemples de Frankfurt et du déterminisme. Cela contredit (b), qui affirme justement l'incompatibilité des contre-exemples de Frankfurt et du déterminisme. Je ne peux donc pas soutenir (a) et (b) en même temps. Donc, la justification ne se transmet pas.

Même si cela n'est pas nécessaire pour prouver que les contre-exemples de Frankfurt ratent leur cible, il vaut la peine de tirer les conséquences plus générales de notre analyse de la transmission de la justification : notre argument semble valoir pour toute tentative de déduire des énoncés factuels à partir d'intuitions morales. Cette conclusion semble d'autant plus pertinente à exposer ici qu'elle fixe des balises supplémentaires sur l'utilisation de notre intuition morale dans une argumentation (comme le fait le PNNP).

Voici une preuve possible de notre principe général. Par *reductio ad absurdum*, je postule que ce type d'argument peut fonctionner :

S1 : Mon intuition morale

***Me justifie de croire***

(N) : Un énoncé strictement normatif<sup>53</sup>.

***Or, je sais que (N) est incompatible avec***

(D) : Un énoncé strictement descriptif.

***Donc, je suis justifié de croire que***

(F) : L'énoncé (D) est faux.

Pour être justifié de croire (N), grâce à mes intuitions morales (S1), je dois supposer implicitement qu'aucun fait du monde ne bloque (N). Cela peut être vrai pour deux raisons : soit, d'une part, parce que les énoncés de faits (comme (D)) qui pourraient bloquer (N) sont faux dans notre monde; soit, d'autre part, parce qu'aucun énoncé de fait du monde (comme (D)) *ne peut* bloquer (N), (N) étant compatible avec n'importe quel énoncé de fait du monde<sup>54</sup>. Dans le premier cas, il y a pétition de principe. En effet, pour justifier ma première prémisse (N), je suppose implicitement que ma conclusion est vraie (F). Je suppose donc que ma conclusion est vraie pour prouver qu'elle est vraie, c'est-à-dire que je suppose que (D) est faux pour montrer que (D) est faux. Dans le deuxième cas, c'est-à-dire si je suppose que (N) et (D) sont compatibles, alors l'argument ne tient pas parce que ce qui me permet de passer de (N) à (F) est précisément l'incompatibilité de (N) et de (D). Autrement dit, ce qui justifie (N) m'empêche de déduire (F) à partir de (N). Par conséquent, la transmission de la justification ne peut pas s'effectuer puisque 1. et 2. sont satisfaits, mais pas 3. En d'autres termes, je ne

---

<sup>53</sup> L'énoncé « Pierre est moralement responsable de x » n'est pas strictement normatif puisqu'il suppose que Pierre existe.

<sup>54</sup> Dans le débat sur le libre arbitre, la première stratégie est celle du libertarien, la seconde est celle du compatibiliste.

peux pas déduire d'énoncés moraux que ce qui les rend possibles est vrai. C'est procéder à l'envers.

Nous avons donc établi que les contre-exemples de Frankfurt sont soumis au PNNP puisqu'ils répondent à ses deux conditions. Parce qu'ils sont soumis au PNNP, les contre-exemples de Frankfurt au PAP – qui sont utilisés pour bloquer l'argument de la conséquence – ne peuvent pas supposer l'indéterminisme. Il nous reste à faire la démonstration que les contre-exemples de Frankfurt sont impossibles dans un contexte déterministe.

Avant de procéder, nous croyons utile de rappeler que ce résultat a une certaine importance : d'une part, de nombreux philosophes ont formulé des contre-exemples de Frankfurt influents dans des contextes explicitement indéterministes (Pereboom 2003, 2009a, 2001, 2012, Widerker 2006a, 2009, Mele et Robb 1998, Haji et McKenna 2006, Hunt 2007)<sup>55</sup>; d'autre part, plusieurs ont admis qu'il serait dialectiquement peu approprié de formuler des contre-exemples de Frankfurt dans un cadre *déterministe* (Pereboom 2003, Moya 2011,

---

<sup>55</sup> Voici quelques exemples qui entrent en contradiction avec le PNNP :

Widerker :

As things transpire, Jones decides at time *T* to break the promise on his own. [...] Jones's decision not to keep the promise (D(-K)) being neither nomically determined nor caused, *that is, consistent with its being a decision that Jones makes on his own*. [...] These assumptions do not render Jones's actual decision of not keeping the promise (D(-K)) causally determined. (Widerker 2006a, p. 170, nous soulignons)

Al Mele et David Robb :

Our scenario features an agent, Bob, who inhabits a world at which determinism is false [...] (Mele et Robb 1998, p. 101).

Pereboom :

Joe is considering claiming a[n illegal] tax deduction [...]. Suppose he has a strong but not always overriding desire to advance his self-interest [...] even if it involves illegal activity. In fact, it is causally necessary for his failing to decide to evade taxes by the deadline that he imagines [...] being at least fairly severely punished for doing so, a mental state he can produce voluntarily. However, Joe's imagining in this way being punished is not causally sufficient for his failing to choose to evade taxes. If he were to do so, he could, exercising his *libertarian free will*, either decide to evade taxes or refrain from so deciding. (Pereboom 2011, p. 414, nous soulignons)

Le fait que plusieurs auteurs aient conscience des présupposés indéterministes de la situation IRR qu'ils proposent met en relief l'importance de notre défense du PNNP et donc de notre attaque contre la version indéterministe des contre-exemples de Frankfurt.

Robinson 2012, Bates 2007), une idée que nous aborderons dès à présent. Voici le schéma de notre argumentation jusqu'à présent, qui nous renvoie à la formulation de contre-exemples de Frankfurt dans des cadres déterministes.

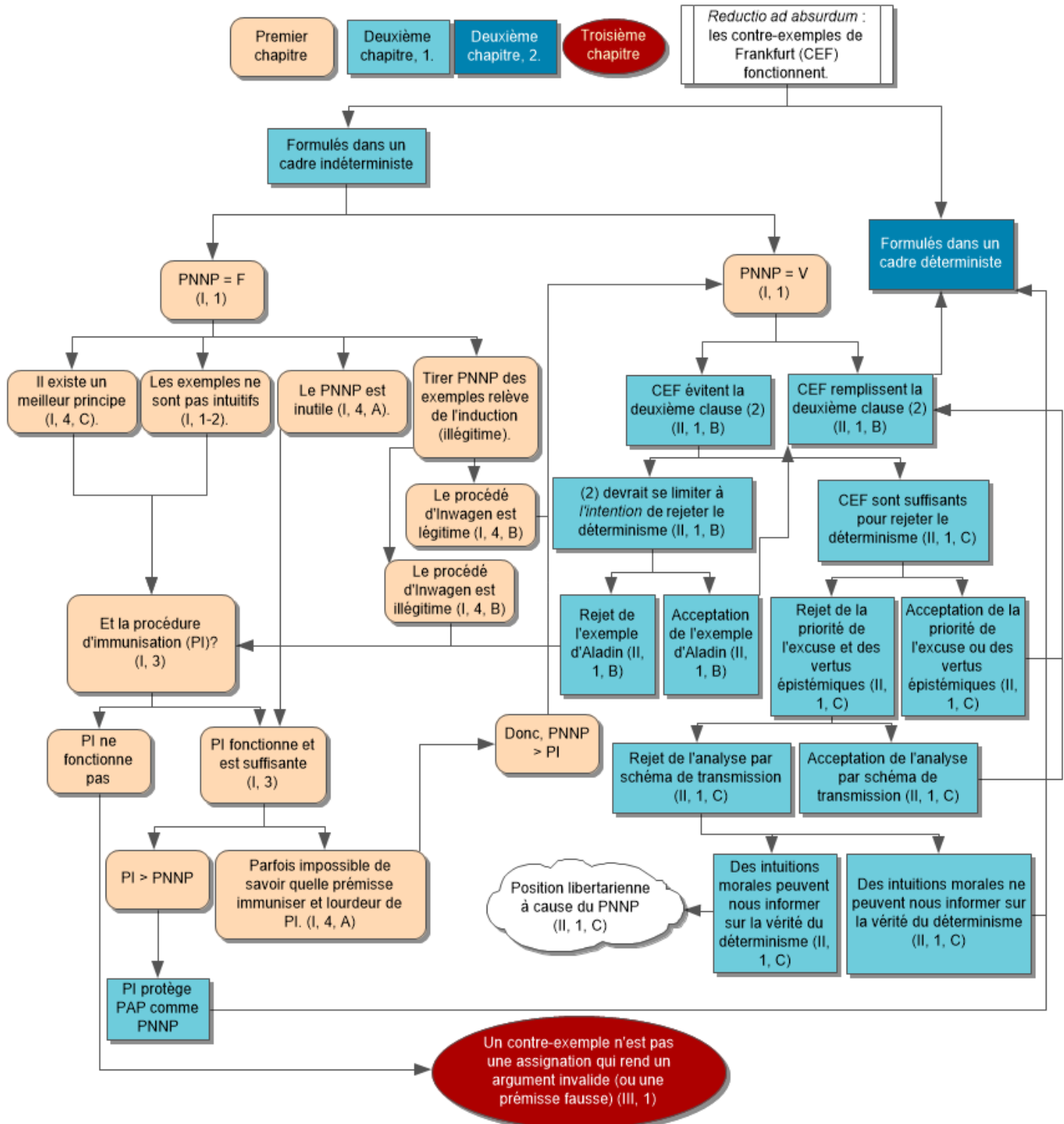


Schéma 1

## 2. L'illégitimité des contre-exemples de Frankfurt déterministes

### A. La défense du dilemme (*dilemma defence*) de Kane-Ginet-Widerker

Kane (1985, 1996), Ginet (1996) et Widerker (2003), présentent un argument sous la forme d'un dilemme, qui est considéré par plusieurs comme la principale menace aux contre-exemples de Frankfurt (Haji et McKenna 2006, Hunt 2007, Fischer 2010, Widerker et Goetz 2013, Garnett 2013). Soit un contre-exemple de Frankfurt tel que nous l'avons exposé précédemment, où Jones se demande s'il va mentir et où Black est un agent contrefactuel qui préférerait ne pas intervenir, mais qui met en place un dispositif qui modifie la décision de Jones si jamais il s'apprête à décider de ne pas mentir. La séquence où Jones effectue son choix sous la supervision de Black est ou bien indéterministe ou bien déterministe. Nous avons démontré dans la première partie du chapitre II que si la séquence est indéterministe, alors elle enfreint le PNNP; Kane, Ginet et Widerker argumentent, quant à eux, dans un tout autre sens. Si, avancent-ils, d'un premier côté, la séquence est *indéterministe*, cela signifie que Jones a, dans les faits, la possibilité d'agir autrement. Jones peut en effet au moins *commencer* à formuler un choix différent de celui pour lequel Black aimerait qu'il opte (s'il ne le peut pas, ce ne peut qu'être en vertu d'une intervention préalable de Black, laquelle retirerait la responsabilité morale de Jones). La situation n'est donc pas un contre-exemple aux possibilités alternatives puisqu'un tel contre-exemple requiert une situation où Jones est moralement responsable de son action *sans avoir de possibilités alternatives*. Cette branche est encore chaudement débattue grâce, notamment aux contre-exemples de Fischer et de Pereboom. Nous laisserons de côté cette discussion puisque notre premier chapitre nous en exempte (c'est transgresser le PNNP que de construire des contre-exemples de Frankfurt indéterministes).

D'un deuxième côté, si la séquence est *déterministe*, alors l'agent n'est pas moralement responsable de son action pour un incompatibiliste. Comme l'impact du déterminisme sur la responsabilité morale d'un agent est précisément la question que sont sensés trancher les contre-exemples de Frankfurt, il semble s'agir d'une pétition de principe. Si Kane, Ginet et Widerker ont raison à l'égard du second terme du dilemme qu'ils exposent, alors notre travail se termine ici puisque nous avons montré que l'autre branche du dilemme était déjà fermée dans la première partie du chapitre. Celui qui accepte notre démonstration que seul PAP\* (si le déterminisme est vrai, alors si nous sommes responsables de  $x$ , nous avons des possibilités alternatives à  $x$ ) est nécessaire pour l'argument de la conséquence ou celui qui consent à ce qu'un contre-exemple de Frankfurt ne puisse pas légitimement supposer le déterminisme faute de quoi il enfreint le PNNP peut déclarer d'emblée (s'il accepte l'argument de Kane-Ginet-Widerker) que les contre-exemples de Frankfurt sont inoffensifs contre l'argument de la conséquence. Ce résultat est étonnant : dans le cadre de l'argument de la conséquence, le PAP ne pourrait souffrir aucun contre-exemple pour des raisons strictement logiques et méthodologiques!

Ce serait cependant aller trop vite en affaires que d'écarter du revers de la main le débat entourant la question de la pétition de principe dans les contre-exemples de Frankfurt. Reprenons donc cette accusation dans le détail, laquelle concerne non seulement les contre-exemples de Frankfurt, mais tout contre-exemple possible au principe des possibilités alternatives :

P13 : Un contre-exemple au PAP doit identifier une situation où

- Un agent est moralement responsable de ses actions;
- Un agent n'avait pas de possibilités alternatives.

P14 : Pour accepter qu'un agent soit moralement responsable dans un monde déterministe, il faut implicitement nier le PAP.

P15 : Faire la démonstration qu'un principe est faux sur la base de sa négation implicite est une pétition de principe<sup>56</sup>.

P16 : Les contre-exemples de Frankfurt font la démonstration que le PAP est faux.

C4 : Si un contre-exemple de Frankfurt au PAP est déterministe, alors il s'agit d'une pétition de principe.

Celui qui défend P14, c'est-à-dire celui qui soutient que pour accepter qu'un agent soit moralement responsable dans un monde déterministe, il faut implicitement nier le PAP, s'appuie sur l'argument suivant. *Comme le déterminisme supprime les possibilités alternatives*, pour accepter qu'un agent soit moralement responsable dans un monde déterministe, il faut accepter que les possibilités alternatives ne soient pas nécessaires à la responsabilité morale. En effet, dans une situation déterministe, la personne qui n'est pas déjà convaincue des conclusions de Frankfurt ne peut pas admettre que Jones est responsable de son action. Bien sûr, comme la construction de contre-exemples de Frankfurt nécessite une situation un agent est *bel et bien* responsable moralement (P13), il faut, pour construire un contre-exemple de Frankfurt déterministe, supposer que le PAP est faux. Et, cela constitue une pétition de principe (P15-16).

## **B. Deux réponses à l'accusation de pétition de principe**

Deux lignes argumentatives similaires ont été formulées et débattues autour de la question de la pétition de principe dans les contre-exemples de Frankfurt déterministes et chacun des débats qui en ont résulté nous paraît apporter des nuances légitimes à l'accusation initiale.

Dans les deux cas, c'est P14 qui est explicitement attaquée.

---

<sup>56</sup> Faire la démonstration qu'un principe est faux sur la base de sa négation implicite est une pétition de principe tout court. Faire cette même démonstration sur la base de positions que n'acceptent pas préalablement la personne que l'on veut convaincre constitue une pétition de principe contre cette même personne.

La première réponse à cet argument vient du côté de McKenna et Haji (Haji et McKenna 2006, Haji 2006). Pour eux, la construction d'un contre-exemple de Frankfurt déterministe ne suppose pas *que les possibilités alternatives ne sont pas nécessaires à la responsabilité morale* ( $\neg$  PAP), mais seulement que *le déterminisme n'implique pas forcément l'absence de possibilités alternatives*. La première question est celle des ingrédients de la responsabilité morale; la seconde est celle des conséquences directes du déterminisme. Autrement dit, Haji et McKenna rejettent P14 parce que ceux qui adhèrent naturellement au PAP n'adhèrent pas automatiquement au fait que le déterminisme supprime les possibilités alternatives. La question qui est sur la table n'est *pas* celle du lien entre le déterminisme et les possibilités alternatives, mais bien celle du lien entre les possibilités alternatives et la responsabilité morale. L'interlocuteur qui n'est pas *a priori* compatibiliste ou incompatibiliste ne devrait avoir aucune difficulté à régler les deux questions séparément. Bref, seul l'incompatibiliste de la marge de manœuvre acceptera P14 et non tout défenseur du PAP. L'argument présenté précédemment en faveur de P14 ne tient en fait que pour lui.

C'est dire qu'un contre-exemple de Frankfurt déterministe ne peut convaincre un incompatibiliste de la marge de manœuvre et constitue une pétition de principe *contre lui*<sup>57</sup>, comme le souligne d'ailleurs Pereboom (2003). Il serait impossible de déterminer qui, de l'incompatibiliste de la marge de manœuvre, de l'incompatibiliste de la source ou du compatibiliste a la réaction intuitive la plus appropriée face aux contre-exemples de Frankfurt.

---

<sup>57</sup> Sur ce point, Fischer s'éloigne de Haji et McKenna puisqu'il pense non seulement que son argument peut convaincre des incompatibilistes de la marge de manœuvre, mais qu'il fonctionne même en faisant la supposition que le déterminisme est incompatible avec les possibilités alternatives. Il croit que cette dernière hypothèse ne fournit pas de quoi l'accuser d'une pétition de principe puisqu'il ne l'utilise pas dans son argument (Fischer 2010).



Pour cette raison, Pereboom croit qu'il reste préférable de construire des contre-exemples de Frankfurt indéterministes, une emprise que nous avons rejetée (II, 1)<sup>58</sup>.

Haji et McKenna (2006) sont d'accord avec le premier constat de Pereboom : ils acceptent que les contre-exemples de Frankfurt constituent une pétition de principe dans un contexte dialectique large (« *broad* »), un contexte où l'on suppose que le déterminisme éradique les possibilités alternatives ou, plus simplement, un contexte où l'on tente de convaincre l'incompatibiliste de la marge de manœuvre. Par ailleurs, les contre-exemples de Frankfurt ne constituent pas une pétition de principe dans un contexte dialectique étroit (« *narrow* »), c'est-à-dire dans un contexte où l'auditeur ne se prononce pas sur le lien entre déterminisme et possibilités alternatives. Pourtant, contrairement à Pereboom, Haji et McKenna sont convaincus qu'un contexte large est en fait peu recommandable, puisqu'il faut débattre de la question du PAP en « terrain neutre », eu égard à la question de la compatibilité du déterminisme et du libre arbitre. Frankfurt aurait, selon eux, ouvert ce terrain neutre qui aurait permis de mettre (temporairement) un terme à l'impasse dialectique entre compatibilistes et incompatibilistes. Ils notent d'ailleurs que les contre-exemples de Frankfurt peuvent convaincre les compatibilistes classiques parce que selon l'interprétation

---

<sup>58</sup> Pereboom pense tout de même que les contre-exemples de Frankfurt ravivent la possibilité que des faits à propos de l'histoire causale d'une action soient les pivots de l'explication de la responsabilité morale d'un agent. Si nous partageons son analyse, il faudrait voir les contre-exemples de Frankfurt comme une invitation à inspecter une théorie concurrente à celle qui fonde la responsabilité morale sur les possibilités alternatives sans pourtant l'invalidier. Si tel est le cas, les contre-exemples de Frankfurt ne sont donc pas des contre-exemples, mais des exemples éclairants de l'application d'une théorie concurrente à l'incompatibilisme de la marge de manœuvre. Le fait que le PAP ne puisse souffrir de contre-exemple dans le cadre de l'argument de la conséquence montre que l'incompatibilisme de la marge de manœuvre est une théorie cohérente, qui peut analyser de nombreuses situations différentes tout en demeurant intuitive. Par conséquent, un cadre alternatif ne devrait pas tant tenter de s'imposer en tentant d'identifier des contre-exemples au PAP, mais plutôt en avançant des raisons indépendantes de préférer l'un ou l'autre des cadres. Par exemple, David Copp (2006, 2008) a montré que des considérations sur l'équité pouvaient nous conduire à endosser le PAP. Dans la section suivante, nous donnerons justement des raisons métaphysiques de penser que le cadre d'analyse qui fait fi des possibilités alternatives suppose une curieuse conception de la relation « rendre inévitable ».

conditionnelle des possibilités alternatives qu'ils adoptent (je peux faire autrement que  $x$  si et seulement si, *si j'avais voulu faire autrement que  $x$ , j'aurais fait autrement que  $x$* ), l'agent (Jones) n'a pas de possibilités alternatives et est moralement responsable de son action. Ils y voient une raison *prima facie* de croire que les contre-exemples de Frankfurt déterministes ne sont pas coupables d'une pétition de principe.

Une deuxième réponse à l'accusation de pétition de principe a été défendue par Fischer (1999, 2005a, 2010, 2013)<sup>59</sup>, lequel a tenté d'établir que les contre-exemples déterministes de Frankfurt, même s'ils ne permettent pas d'établir d'un coup que le PAP est faux, constituent en réalité une forme d'argument en deux étapes. Il utilise, pour ce faire, le contre-exemple de Frankfurt suivant :

So we assume causal determinism obtains and [...] we make no assumption about the relationship between causal determinism and access to alternative possibilities. Black checks and sees the “prior sign” at  $t_1$  that is associated with a subsequent vote for the Democrat — say, the furrowed left brow. Given that Black knows that causal determinism obtains, he can now relax, as it were; under these circumstances, Black knows that Jones in fact will subsequently choose to vote for Obama and carry out that choice. It is also true, given Black's device and dispositions, that if Jones were to show the sign at  $t_1$  associated with voting for a Republican at  $t_2$  (appropriately enough, the furrowed right brow), Black's device would swing into action and stimulate Jones's brain so as to ensure that he chooses at  $t_2$  to vote for Obama and does so vote at  $t_3$ . I claim that this additional fact, *when added to the assumption of causal determinism* and the fact that Black can *thus* be sure that Jones's showing the prior sign at  $t_1$  will in fact be followed by his choosing accordingly at  $t_2$ , renders it true that Jones cannot at  $t_2$  choose to vote for McCain (or subsequently vote for McCain). These two facts together make it the case that Jones cannot at  $t_2$  choose to vote for McCain or carry out such a choice (Fischer 2010, p. 326, nous soulignons).

---

<sup>59</sup> Le débat entre Fischer, Widerker et Goetz s'étend sur plusieurs articles et années. Nous avons retenu le meilleur de l'argumentation de Fischer.

Au premier chef, le contre-exemple précédent permet de montrer que le *simple fait* de la présence de possibilités alternatives n'est pas explicatif (« *is explanatorily irrelevant* ») de l'attribution de la responsabilité morale. Pour le prouver, Fischer doit montrer que certains faits suppriment les possibilités alternatives en laissant la responsabilité morale intacte, et ce, sans faire appel au fait que le déterminisme retire les possibilités alternatives. Cette précision est primordiale, sans quoi il s'agira d'une pétition de principe *et* contre l'incompatibiliste de la marge de manœuvre *et* contre l'interlocuteur indécis. En effet, celui qui se demande si le déterminisme menace notre responsabilité morale ne pourra pas être convaincu par une simple situation où l'on suppose que cette menace échoue : on ne peut montrer qu'un énoncé est faux en demandant d'imaginer qu'il l'est. Il faut un scénario où ce qui menace potentiellement la responsabilité ne joue aucun rôle causal déterminant, mais où pourtant il n'y a pas de possibilités alternatives; bref, un scénario où ce n'est pas le déterminisme qui supprime directement les possibilités alternatives.

Au deuxième chef, une fois acceptée l'idée selon laquelle *le simple fait* des possibilités alternatives ne joue pas un rôle dans la responsabilité morale, il faut vérifier si la présence du déterminisme viendrait, pour une autre raison que l'absence de possibilités alternatives, miner la responsabilité morale. Voyant que ce n'est pas le cas, nous serions alors amenés à rejeter le principe des possibilités alternatives.

Fischer focalise toute son attention sur la première étape de son argumentation : comme Haji et McKenna, il doit prouver que Jones n'a pas de possibilités alternatives à  $t_2$  sans faire appel au fait que le déterminisme supprime les possibilités alternatives. Ce sont trois faits qui jouent ce rôle pour Fischer :

- (F1) si Jones avait choisi à t2 de voter pour McCain, son sourcil se serait préalablement levé, à t1;
- (F2) si le sourcil de Jones s'était levé à t1, la machine de Jones aurait fait en sorte que Jones vote pour Obama à t2;
- (F3) le déterminisme est vrai (sans postuler de lien entre le déterminisme et les possibilités alternatives).

Fischer admet que le déterminisme est en effet nécessaire au contrefactuel (F1), lequel doit être fiable pour Black<sup>60</sup>. Or, ces faits (F1, F2, F3) ne changent rien à l'attribution de la responsabilité morale de Jones; ils la laissent intacte puisque Jones ne lève pas son sourcil à t1 et que nous ne postulons aucune influence du déterminisme sur les possibilités alternatives. Fischer utilise alors l'argumentation suivante : si  $p$  n'est pas explicatif pour  $r$  et que  $p$  est une des façons de faire en sorte que  $q$  (*make it the case that q*), le *simple fait*<sup>61</sup> que  $q$  (peu importe ce qui a fait en sorte que  $q$  est vrai) ne peut pas être explicatif de  $r$ . Conséquemment, si la conjonction de (F1), (F2) et (F3) ( $p$ ) n'est pas explicative de l'attribution de la responsabilité morale ( $r$ ) et que la conjonction de (F1), (F2) et (F3) ( $p$ ) élimine les possibilités alternatives ( $q$ ), alors le *simple fait* de la présence de possibilités alternatives ( $q$ ) n'est pas explicatif de l'attribution de la responsabilité morale ( $r$ ).

Bref, en postulant un agnosticisme concernant les liens entre le déterminisme et les possibilités alternatives, Fischer avance la même idée que McKenna et Haji : il n'y a pas lieu, dialectiquement, de supposer d'avance que le déterminisme supprime bel et bien les possibilités alternatives et en s'abstenant de se prononcer sur la question, on réalise que cet

---

<sup>60</sup> Il n'est pas si clair que le déterminisme puisse sécuriser quoi que ce soit sans utiliser sa propriété de retirer les possibilités alternatives. Le cas échéant, Fischer est coupable de la même pétition de principe qu'il dit parvenir à éviter.

<sup>61</sup> Il est très important de spécifier que  $p$  fait en sorte que  $q$  et qu'il ne s'agit que du *simple fait que q*, sinon des contre-exemples peuvent être formulés à cette forme de raisonnement (Widerker et Goetz 2013).

effet supposé du déterminisme (supprimer les possibilités alternatives) n'est pas du tout nécessaire pour montrer que Jones n'a pas de possibilités alternatives.

### **C. Analyse critique des solutions de Fischer et de Haji-McKenna**

Le débat ne s'est pas arrêté à cette double réponse de Fischer et de Haji-McKenna : la réplique presque unanime de leurs opposants (Goetz 2005, Palmer 2005, Palmer 2006, Widerker et Goetz 2013) a consisté à montrer qu'il est impossible de ne pas postuler que le déterminisme supprime les possibilités alternatives sans faire échouer les contre-exemples de Frankfurt. En effet, si on s'abstient d'affirmer que le déterminisme supprime les possibilités alternatives, Jones a toujours bel et bien des possibilités alternatives à sa disposition. Bien sûr, le cas échéant, le scénario ne prouve aucunement que l'absence de possibilités alternatives est compatible avec la responsabilité morale. Il est donc nécessaire de postuler que le déterminisme supprime les possibilités alternatives, ce qui nous ramène à l'accusation initiale de pétition de principe.

C'est d'abord Palmer (2005, 2006) qui a tenté de montrer qu'il était nécessaire de supposer que le déterminisme supprimait les possibilités alternatives pour que Jones soit dépourvu de telles possibilités. Palmer s'attaque à des contre-exemples de Frankfurt « traditionnels », c'est-à-dire où le signe qu'observe l'agent contrefactuel Black n'est non pas *nécessaire*, mais *suffisant* pour effectuer l'action en question. Si le déterminisme ne supprime pas les possibilités alternatives, cela signifie que Jones peut bel et bien *commencer à décider* d'exécuter l'action que Black veut l'empêcher d'accomplir. En effet, à tout moment, Jones peut *créer une nouvelle chaîne causale*, ce qui signifie que Jones a réellement des possibilités alternatives. Black ne peut rectifier le tir *que subséquemment*.

Haji et McKenna répondent que la faille identifiée par Palmer n'est pas applicable dans un cas où le signe observé par Black est *nécessaire* à l'action (2006) plutôt que *suffisant*. Il est nécessaire à Jones de lever son sourcil droit avant de commencer à décider de voter pour McCain et donc Black peut intervenir avant ce début de décision, c'est-à-dire au moment du sourcillement. Pour Haji et McKenna, le déterminisme ne joue aucun rôle dans un contre-exemple de Frankfurt déterministe sinon celui de surdéterminer l'action de Jones. Il n'est même pas nécessaire de postuler le déterminisme dans la situation, ce qui n'est pas très différent, au final, de construire un contre-exemple de Frankfurt indéterministe, remarque avec justesse Garnett (2013)<sup>62</sup>.

La position de Haji et McKenna n'a pas fait l'unanimité. En fait, deux positions concurrentes s'en distinguent et c'est en les exposant systématiquement que nous pourrions découvrir s'il y a finalement ou non une pétition de principe dans les contre-exemples de Frankfurt déterministes. Dans un contre-exemple de Frankfurt déterministe, pour Haji-McKenna, on n'a pas besoin de postuler que le déterminisme est vrai ni de postuler qu'il retire les possibilités alternatives; pour Fischer, il semble nécessaire de postuler que le déterminisme est vrai, mais non que le déterminisme enlève les possibilités alternatives; pour Goetz-Widerker (et pour Palmer), il faut postuler aussi bien que le déterminisme est vrai qu'il enlève les possibilités alternatives.

### **A) La position de Haji-McKenna**

Haji et McKenna soutiennent l'argument suivant :

---

<sup>62</sup> Garnett prétend que construire un contre-exemple de Frankfurt déterministe sans utiliser le déterminisme pour faire la démonstration de l'absence des possibilités alternatives ne constitue pas un gain dialectique par rapport à la formulation d'un contre-exemple de Frankfurt indéterministe. Sur ce point, nous ne pouvons qu'être en

(F0) À t2, soit Jones vote pour McCain, soit Jones vote pour Obama (disjonction exclusive);

(F1) si Jones avait choisi à t2 de voter pour McCain, alors à t1 son sourcil se serait levé;

(F2) si le sourcil de Jones s'était levé à t1, la machine de Jones aurait fait en sorte que Jones vote pour Obama à t2;

*Intermédiaire : (F5) Jones ne choisit pas à t2 de voter pour McCain;*

(F4) Donc, Jones n'a pas la possibilité de voter pour McCain à t2 (son vote pour Obama est inévitable).

Il y a deux problèmes fondamentaux dans cet argument. Le premier, c'est que, comme l'a montré Fischer (2010, 2013), si le déterminisme est faux, alors nous n'avons aucune garantie de (F1), et, par la force des choses, Black n'en a pas non plus. Soutenir (F1) implique d'affirmer que la séquence des événements est déterministe. En effet, sans la vérité du déterminisme, Jones a bel et bien des possibilités alternatives puisqu'il peut commencer à décider de voter pour McCain. Bref, Black et sa machinerie sont en eux-mêmes insuffisants pour saper complètement les possibilités alternatives de Jones<sup>63</sup>. C'est là une critique souvent adressée aux versions indéterministes des contre-exemples de Frankfurt. Ce problème peut être évité en postulant le déterminisme sans postuler qu'il supprime les possibilités alternatives

---

désaccord puisqu'un contre-exemple de Frankfurt déterministe n'enfreint pas le PNNP contrairement à un contre-exemple de Frankfurt indéterministe.

<sup>63</sup> Par ailleurs, il y a peut-être une autre raison pour laquelle ces exemples échouent : en délayant le rôle de l'agent contrefactuel à un dilemme qui n'a jamais lieu au final dans l'esprit de Jones, Fischer (et Pereboom (c.f. 2011)) le rend responsable non plus d'avoir évité de payer ses taxes, mais d'avoir évité de se mettre dans les dispositions appropriées pour ce faire (cf. Otsuka 1998). La richesse du débat qui s'ensuit nous oblige à délaissier cette proposition, d'ailleurs très controversée. Il peut très bien se faire que nous ne puissions être rendus coupables de nous mettre dans des dispositions nécessaires à la prise de la bonne décision pour la simple et bonne raison que ces conditions peuvent être impossibles à identifier et leurs conséquences peuvent être imprévisibles (Pereboom 2009a). Pourtant, si Jones est un agent autonome capable d'introspection, il devrait peut-être mieux connaître les conditions d'une délibération adéquate. Il n'a pas besoin de savoir qu'en atteignant cet état, peu importe le résultat de sa réflexion, il sera dépourvu de sa responsabilité morale par l'intervention de l'agent contrefactuel. Tout ce qu'il doit savoir, c'est que cette réflexion est une condition *nécessaire* pour éviter de frauder. En ne se plaçant pas dans cette disposition une seule fois avant de frauder, il refuse continuellement de se placer dans une situation où il pourra choisir convenablement. Remettre cette capacité en question pourrait remettre en cause que l'agent est « reason responsive » (Moya 2011) et du même coup sa responsabilité. Reste que Pereboom a peut-être raison de soutenir que c'est l'adhésion à l'incompatibilisme de la marge de manœuvre

(Funkhouser 2006). Le deuxième problème, souligné par Goetz-Widerker et par Palmer, c'est que (F0)-(F2) ne permettent que de déduire (F4), mais seulement (F5). La notion de possibilité émerge dans la conclusion sans apparaître dans les prémisses.

### **B) La position de Fischer**

Nous considérons à présent une méthode qui permet de régler le premier problème, c'est-à-dire de garantir (F1). C'est la position que Fischer soutient par l'argument suivant :

(F0) À t2, soit Jones vote pour McCain, soit Jones vote pour Obama (disjonction exclusive);

(F1) si Jones avait choisi à t2 de voter pour McCain, alors à t1 son sourcil se serait levé;

(F2) si le sourcil de Jones s'était levé à t1, la machine de Jones aurait fait en sorte que Jones vote pour Obama à t2;

(F3) le déterminisme est vrai (sans postuler de lien entre le déterminisme et les possibilités alternatives);

*Intermédiaire : (F5) Jones ne choisit pas à t2 de voter pour McCain;*

(F4) Donc, Jones n'a pas la possibilité de voter pour McCain à t2 (son vote pour Obama est inévitable).

Comme Fischer le souhaite, (F1) est maintenant garantie : « The role of determinism is to ensure the truth of this counterfactual » (Fischer 2013). Le second problème demeure pourtant présent : dans la conclusion (F4), la notion de possibilités surgit sans avoir été même mentionnée dans (F0-F3). Il reste alors à comparer la position de Fischer à celle de Goetz-Widerker.

### **C) La position de Goetz-Widerker**

Goetz et Widerker prétendent tout simplement que le passage de (F1, F2, F3) à (F4) échoue. Il ne peut, encore une fois qu'arriver à l'intermédiaire. Ils écrivent :

---

qui mène à appliquer le schéma de la responsabilité délayée pour contrer son argument, rendant cette position



But one lesson we have learned from arguments employed by incompatibilists for the incompatibility of determinism and one's freedom to avoid acting as one does is that these entailments [F1-F2] together with

(2b) [Jones has furrowed his left eyebrow at t1]

and

L [the laws of nature]

do not suffice to infer that

[F4]/(4a) It is unavoidable (relative to times later than t1) that

Jones will choose at t2 to vote for Obama.

[...]

For this, the incompatibilist needs to appeal to stronger assumptions, namely, that it is not within Jones's (or anyone's) power (at  $t1 + \Delta$ ) to bring about the non-obtaining of (hard) facts about the past, and specifically a (hard) fact about the past such as (2b); that it is not within Jones's (or anyone's) power to falsify the laws of nature; and also to the assumption that certain modal principles such as van Inwagen's Rule alpha and Rule beta are valid. These assumptions, however, jointly imply that causal determinism rules out alternative possibilities. (Widerker et Goetz 2013, p. 290-291)

Goetz et Widerker ne rendent peut-être pas assez compte du fait que Fischer utilise (F3), mais leur critique ne concerne pas le fait que (F1) n'est pas « sécurisé ». Pourtant, Fischer ne leur donne pas une réponse satisfaisante :

[I]t should be obvious that the argumentation offered by Widerker and Goetz to show that (2a)-(2c) [F1-F3] do not rule out (3) [F4] is quite beside the point; it doesn't touch my argument. And there is no straightforward way that it could be extended to apply to my argument, since my argument explicitly employs the conjunction of the presence of Black's device and dispositions and causal determinism. [...] Allow me here to reiterate my explanation of how the argument is supposed to work. Black checks and sees the prior sign at t1 (the raised left eyebrow) that is associated with a subsequent vote for Obama. Thus Black knows that Jones in fact will subsequently choose to vote for Obama; but this of course in itself does not imply that Jones cannot just prior to or at t2 choose otherwise. Given the assumption of causal determinism, we know that the following counterfactual is true: 'Had Jones chosen at t2 to vote for McCain, then his right brow would have been raised at t1.' *The role of the assumption of causal determinism is to ensure the truth of this counterfactual;*

---

coupable d'une pétition de principe. (Pereboom 2012)

*Black thus doesn't have to concern himself with the possibility that, although Jones furrowed his left brow at t1, he would nevertheless (begin to) choose to vote for McCain at t2. And it is also true, given the set-up of the situation, that if Jones were to show the sign at t1 associated with choosing to vote for McCain at t1, Black's device would swing into action and stimulate Jones's brain so as to ensure that he chooses at t2 to vote for Obama. These facts together make it the case that Jones cannot at t2 choose to vote for McCain (or subsequently vote for McCain). (Fischer 2013, p. 492-493, nous soulignons)*

Il semble un peu étrange que Fischer fasse si peu de cas du passage du registre des contrefactuels au registre des possibilités, alors que Widerker et Goetz l'ont justement mis en garde sur ce point. En effet, Goetz et Widerker ont bien pris la peine d'exposer que deux conditions étaient nécessaires pour qu'un agent n'ait pas de possibilités alternatives dans un contre-exemple de Frankfurt : que le déterminisme soit vrai et qu'il supprime les possibilités alternatives de l'agent.

Le dialogue entre Fischer et Goetz-Widerker procède donc d'une façon désespérante, c'est-à-dire qu'à chaque phase du dialogue s'ajoute un nouveau malentendu : Fischer dit d'abord : « je peux faire ma déduction en supposant le déterminisme sans supposer de ses liens avec les possibilités alternatives »; Widerker et Goetz répondent : « il est impossible de faire votre déduction, *comme vous le faites, sans utiliser le déterminisme* et sans supposer qu'il enlève les possibilités alternatives<sup>64</sup> » et Fischer de répondre : « vous m'avez mal compris, j'utilise le déterminisme »!

Il nous reste alors à trancher par nous-mêmes la question : (F0-F3) sont-ils suffisants pour en arriver à déduire (F4)? Autrement dit, il faut voir si dans l'argument de Fischer, Widerker et Goetz ont raison de réclamer non seulement l'utilisation du déterminisme, mais *celle de la thèse selon laquelle il élimine les possibilités alternatives.*

Pour en arriver à l'introduction de la notion de *possibilités* (ou d'inévitabilité) dans l'argument, c'est-à-dire pour passer du registre de (F0-F3) au registre de (F4) il faut ou bien la définir dans les termes des prémisses précédentes, ou bien admettre un ou plusieurs principes qui fassent le pont entre les contrefactuels décrits et les possibilités de Jones (comme ceux énumérés par Widerker et Goetz), mais qui (contrairement aux principes énumérés par Widerker et Goetz) n'obligent pas à rejeter le PAP pour en prouver la fausseté. Dans les deux cas, ce que nous recherchons, c'est une façon non controversée de comprendre la notion de possibilité qui fasse en sorte que (F0-F3) soit *suffisante* pour accepter (F4). Le hic, c'est que toute conception de la notion de possibilité *est* controversée.

Nous proposons alors deux solutions pour sortir Fischer de son embarras. La première est de démontrer que, sous chacune des façons paradigmatiques de comprendre les possibilités, Jones n'a pas de possibilités alternatives. En d'autres termes, il nous faut montrer que peu importe notre façon de comprendre les possibilités, on passe de (F0-F3) à (F4). Nous aurons alors montré qu'il n'est point besoin de postuler que le déterminisme supprime les possibilités alternatives pour obtenir que Jones n'en a aucune. Inspectons trois conceptions paradigmatiques. La première, celle de l'incompatibiliste (soit la conception « étroite » (Taylor et Dennett 2011)), est satisfaite. Pour l'incompatibiliste, Jones a la possibilité de voter pour McCain si et seulement s'il existe un monde possible dont le passé est identique au monde réel et où Jones vote pour McCain. Cette éventualité est bloquée par (F0-F3) : un tel monde n'existe pas. La seconde, la conception conditionnelle de la possibilité (Hume 2000/1748), stipule que Jones a la possibilité de voter pour McCain si et seulement si, s'il avait voulu voter pour McCain, rien ne l'en aurait empêché. Encore une fois, selon cette

---

<sup>64</sup> Goetz et Widerker ne sont pas très clairs sur ce point, mais c'est l'impression qu'ils donnent en page 288.

conception de la possibilité, F0-F3 suffisent à éliminer les possibilités alternatives de Jones à t2. La troisième, la conception « large » de la possibilité, (Dennett 1984b, Taylor et Dennett 2011) stipule que Jones a la possibilité de voter pour McCain à t2 si et seulement s'il existe une part non négligeable de mondes possibles dans lesquels le passé est très similaire à celui du monde réel et où Jones vote pour McCain. Supposer que Black aurait échoué à activer sa machine serait postuler un passé trop différent de celui du monde réel, aussi (F0-F3) éliminent encore une fois les possibilités alternatives de Jones sous cette conception de la notion de possibilité. Bien sûr, un travail similaire resterait à effectuer pour plusieurs autres conceptions des possibilités, mais on peut présager de la bonne marche de l'opération.

Cette solution n'est pourtant pas complètement satisfaisante puisqu'elle nous conduit dans les bras d'un nouveau dilemme : nous aurons eu besoin soit d'une conception de la possibilité qui implique que le déterminisme les élimine toutes et il s'agira d'une pétition de principe en général<sup>65</sup>, soit nous aurons utilisé une conception de la possibilité qui n'est pas admise par un incompatibiliste, et alors nous aurons été coupable d'une pétition de principe contre lui. Il reste que cette solution permet de réaffirmer ce que Haji et McKenna ont défendu précédemment : les contre-exemples de Frankfurt dans des contextes déterministes ne constituent pas une pétition de principe pour des interlocuteurs neutres sur la question de la conception de la notion de la possibilité.

Il reste une seconde solution pour essayer de montrer que construire un contre-exemple de Frankfurt déterministe ne constitue pas une pétition de principe *contre l'incompatibiliste* : trouver une caractérisation des possibilités qui soit *nécessaire* autant pour le compatibiliste

---

<sup>65</sup> Il n'est pas si clair que cette hypothèse amène réellement à une pétition de principe parce que même s'il faut supposer que le déterminisme supprime les possibilités alternatives, il ne semble pas obligatoire d'*utiliser* cette supposition pour montrer que Jones n'a pas de possibilités alternatives. Cela nous amène alors à la section E.

que pour l'incompatibiliste. Nous avons donc besoin d'un principe *plus faible* qui fera le travail qui convient à montrer que, dans la situation, Jones n'a pas de possibilités alternatives. La notion conditionnelle des possibilités semble être une candidate adéquate. En effet, est-ce qu'un incompatibiliste refusera qu'il est *nécessaire* pour qu'un agent soit libre qu'il soit vrai que « s'il avait voulu faire x, il aurait fait x »? Autrement dit, n'est-il pas vrai que l'incompatibiliste considère que le fait que *la volonté soit en accord avec les actions* est une *condition nécessaire* de la liberté? Après tout, ce que l'incompatibiliste reproche à la position humienne (conditionnelle) sur la question de la liberté, c'est qu'elle est *insuffisante*, mais pas qu'elle est *non-nécessaire*<sup>66</sup>. C'est une question riche que nous devons laisser en plan, mais il faut bien comprendre que si nous y répondons par l'affirmative, nous devons accepter que la démarche que propose Fischer n'est *pas* (pour l'instant) bloquée par la réponse de Goetz-Widerker.

#### **D. Constats et nuances sur la question de la pétition de principe**

Voici les constats préliminaires que nous force à tirer notre analyse de la question de la pétition de principe dans la branche déterministe de la défense du dilemme. Un premier constat est que l'accusation de pétition de principe tient toujours si le contexte dialectique est celui où le tenant des contre-exemples de Frankfurt tente de convaincre l'incompatibiliste de la marge de manœuvre qu'il fait fausse route, sauf si la liberté conditionnelle est nécessaire à la conception incompatibiliste des possibilités alternatives. Le cas échéant, l'incompatibiliste doit accepter avec Fischer que (F0-F3) sont suffisants pour en arriver à (F4); qu'il n'est donc pas nécessaire de postuler que le déterminisme supprime les possibilités alternatives pour faire

---

<sup>66</sup> Cela semble s'appliquer autant à un tenant d'une conception des possibilités à la Dennett, c'est-à-dire selon laquelle un agent a une possibilité de faire x s'il existe des mondes possibles similaires où l'agent fait x.

la démonstration que Jones n'a pas de possibilités alternatives dans sa situation. Comme nous entendons montrer que la construction des contre-exemples de Frankfurt échoue, il nous faudra faire la démonstration que même si (F0-F3) est suffisant pour déduire (F4), il y a bel et bien pétition de principe.

Un deuxième constat est qu'il est *a priori* dialectiquement légitime de convaincre un agnostique sur la question de la compatibilité entre le déterminisme et les possibilités alternatives à l'aide des contre-exemples de Frankfurt déterministes, *du moins si l'une de nos deux stratégies argumentatives* (la nécessité de la notion conditionnelle de la possibilité pour l'incompatibiliste ou le fait que sous toutes les conceptions de la possibilité F0-F3 est suffisant à déduire F4) *fonctionne*. Ce constat mérite pourtant d'être nuancé.

Il n'est pas si clair que de s'adresser à l'agnostique sur le lien entre les possibilités alternatives et le déterminisme soit la meilleure posture dialectique. Ce que proposent Fischer, Haji et McKenna – se poser la question du lien entre responsabilité morale et possibilités alternatives indépendamment de la question du lien entre déterminisme et possibilités alternatives – comporte au moins trois risques, risques qui apportent un bémol important à notre second constat. Primo, il est possible que ce soit *la conjonction du déterminisme et de l'absence des possibilités alternatives* qui retire la responsabilité morale (Goetz 2005). C'est la vérité du déterminisme qui expliquerait que l'absence de possibilités alternatives supprime la responsabilité morale<sup>67</sup>. Si tel est le cas, la méthode en deux temps élaborée par Fischer est vouée à l'échec puisqu'elle propose de séparer l'analyse du déterminisme et des possibilités alternatives. Si Goetz a raison sur ce point, la méthode de Fischer ressemble un peu à celle de

---

<sup>67</sup> Goetz se demande même si les possibilités alternatives ne sont pas simplement une façon d'affirmer l'indéterminisme.

celui qui, pour vérifier s'il est dangereux de fumer près d'une station-service, étudie indépendamment les dangers du tabagisme et ceux de la fréquentation d'une station-service sans cigarette. La thèse de Goetz est d'autant plus intéressante pour nous que PAP\* (si le déterminisme est vrai, alors le PAP est vrai) peut justement rendre compte de cette intuition en mentionnant le déterminisme dans la nouvelle formulation du PAP. Nous profitons de cette occasion pour mentionner qu'une telle situation bloque un contre-argument possible à l'application du PNNP aux contre-exemples de Frankfurt que nous n'avons pas mentionné jusqu'à présent. Pour certains lecteurs, il a pu en effet sembler saugrenu que nous ayons formulé l'exigence que les contre-exemples de Frankfurt soient construits dans des environnements déterministes, en II, 1. La raison en est qu'il aurait pu sembler légitime d'endosser le raisonnement suivant :

Il n'y aucune raison pour laquelle un principe moral de liaison comme le PAP soit vrai dans un monde et faux dans un autre. Pourquoi les conditions de la responsabilité morale varieraient? Ce qui peut varier, c'est la responsabilité morale elle-même, non pas ses ingrédients nécessaires. Conséquemment, si le PAP est vrai, il est vrai autant dans les mondes déterministes qu'indéterministes. Si on peut le réfuter dans un monde déterministe, on peut le réfuter également dans un monde indéterministe. PAP\*, qui essaie de restreindre l'affirmation du PAP à un monde déterministe est artificiel et les contre-exemples de Frankfurt n'ont pas à avoir lieu dans un monde particulier.

Outre le fait qu'il semble approprié de demander

s'il est vrai que le PAP est faux dans un monde indéterministe, pourquoi ne pourriez-vous pas le montrer dans un monde déterministe et si vous ne le pouvez pas, ne croyez-vous pas que cela manifeste peut-être une erreur dans le raisonnement précédent?

il est raisonnable de croire que le PAP est vrai *en vertu de la vérité du déterminisme*<sup>68</sup>. Il est donc erroné de croire automatiquement que si le PAP est vrai, il l'est autant dans un monde déterministe qu'indéterministe. Goetz donne précisément des raisons de croire que le PAP est vrai précisément parce que *c'est le déterminisme qui supprime nos possibilités alternatives* (cela ne veut pas dire qu'aucune autre forme de suppression de nos possibilités alternatives ne pourrait supprimer notre responsabilité morale)<sup>69</sup>.

Deuxio, le second constat (le fait qu'il n'y a pas de pétition de principe contre un public indécis sur la question du lien entre déterminisme et possibilités alternatives) doit être nuancé par le fait qu'il n'est pas si clair qu'il soit dialectiquement approprié d'accepter les contre-exemples de Frankfurt sur la base d'un agnosticisme sur la question du rapport entre déterminisme et possibilités alternatives et d'ensuite accepter que le déterminisme prive effectivement les agents des possibilités alternatives (mais pas de leur responsabilité morale, comme nous aurions alors accepté les contre-exemples de Frankfurt). Imaginons que nous voulions régler le débat entre la conception aristotélicienne et moderne de la physique sur la question de la chute des corps. Imaginons qu'on nous demande, pour ce faire, d'ignorer les débats sur l'existence du vide, un débat trop controversé pour être mêlé à celui qui nous concerne. Supposons que l'efficacité empirique des prédictions aristotéliciennes, par exemple sur le fait que les corps lourds chutent plus vite (parce qu'ils sont attirés vers le sol), nous convainque et que nous optons pour la physique aristotélicienne pour décrire la chute des

---

<sup>68</sup> Cette option semble d'autant plus pertinente que si nous acceptons des analyses de l'action comme celle d'Alvarez (2009), qui considère que les moyens de contrôle employés par Black ne peuvent pas résulter en une action de Jones, on comprend que le problème pour la liberté entraîné par le déterminisme peut être plus que l'élimination de possibilités alternatives, mais bien l'élimination des actions elles-mêmes (par l'élimination des possibilités alternatives).

<sup>69</sup> À celui qui voudrait tenter de défendre une version plus restreinte du PNNP, où seules certaines prémisses sont immunisées, c'est-à-dire celles qui donnent le contexte et le sens aux contre-exemples, on pourrait répondre



corps. Si, par la suite, nous admettons l'existence du vide et le distinguons d'un environnement riche en azote et en dioxygène, il serait curieux de ne pas reconsidérer notre position sur la chute des corps. En effet, sachant que l'air offre une résistance aux objets qui chutent, nous sommes à mêmes de rendre compte de la différence de vitesse entre une plume et une pièce de monnaie qui tombent sans faire appel à la différence de leur masse. De la même manière, si nous acceptons les contre-exemples de Frankfurt sur la base de l'argument de Fischer, ou d'Haji-McKenna, mais qu'après un travail fin sur la notion de possibilité, nous réalisons par la suite que nous adhérons bel et bien à l'idée selon laquelle le déterminisme retire les possibilités alternatives, il serait imprudent de ne pas reconsidérer notre adhésion à l'argument de Fischer ou d'Haji-McKenna. Fischer répond à l'avance à un tel argument en prétendant que l'agnosticisme sur le lien entre les possibilités alternatives et le déterminisme fait partie de la « doctrine officielle du semicompatibilisme », la théorie qu'il défend (Fischer 2010). Cette réponse ne peut pourtant qu'être insatisfaisante dans la mesure il n'y a pas de réelle controverse sur la question à laquelle ne répond pas le semicompatibilisme :

The problem is with [Fischer's heuristic agnosticism on the issue whether determinism prevents one from doing otherwise]. Any potential audience for this argument must at least understand what the doctrine of determinism is. They will, therefore, understand that determinism rules out alternative possibilities in some senses and not in others, and they will also understand, presumably, just which sorts of alternatives it rules out and which it does not. After all, this is not an issue that is typically treated as controversial. Nearly everyone agrees, for instance, that determinism does not take away merely conditional possibilities (e.g. of the form "I would if I tried"), or counterfactual possibilities where the past and the laws are not fixed, while it does take away counterfactual possibilities where the past and the laws are fixed, as well as any kind of physical possibility. What is controversial, and forms the major focus of the dispute, is which (if any) of these types of alternative are relevant to

---

qu'en endossant les vues de Goetz, on en arrive à la conclusion que le déterminisme fait partie de ces prémisses dans lesquelles le contexte d'application de l'argument tout entier est précisé.

moral responsibility: when philosophers argue over whether determinism rules out alternatives, they are arguing over whether determinism rules out alternatives in the sense required for moral responsibility. (Garnett 2013)

C'est dire que la position dialectique de Haji, McKenna et Fischer n'est pas aussi sûre qu'ils pourraient le croire.

Tertio, il est possible que le fait que l'agent contrefactuel rende inévitable l'action de Jones dans un monde déterministe ne soit qu'une illusion (Goetz 2005). Autrement dit, peut-être qu'au fond la présence de l'agent contrefactuel n'explique pas l'absence de possibilités alternatives (Garnett 2013), même si (F0-F3) est suffisant pour avoir (F4) et qu'au fond c'est la vérité du déterminisme (F3) qui fait tout le travail. On peut le voir en se demandant pourquoi l'agent contrefactuel n'a pas de rôle explicatif pour l'attribution de la responsabilité morale dans les contre-exemples de Frankfurt. Il est tout à fait possible de concevoir que la raison pour laquelle l'agent contrefactuel n'a pas d'importance morale, ce n'est pas tant parce qu'il n'affecte pas les raisons pour lesquelles l'agent agit, mais c'est tout simplement parce que les agents contrefactuels ne *font jamais en sorte que des actions soient inévitables dans un monde déterministe*. L'élucidation de la relation « rendre inévitable » et du rôle des agents contrefactuels dans un monde déterministe nous semble, dès à présent, inévitable.

Avant de passer à la prochaine section, il convient toutefois d'effectuer une précision capitale. Il faut en effet nous demander si le prérequis pour construire un contre-exemple de Frankfurt déterministe sans commettre de pétition de principe est un prérequis « d'utilisation » ou de « postulation ». Autrement dit, si la construction d'un contre-exemple de Frankfurt déterministe oblige à postuler que le déterminisme supprime les possibilités alternatives, mais sans utiliser cet énoncé dans le passage de F0-F3 à F4, s'agit-il d'une pétition de principe? Si oui, alors la construction de contre-exemples de Frankfurt déterministes est bloquée. En effet,

dans un contre-exemple de Frankfurt, soit le déterminisme supprime les possibilités alternatives, soit il ne les supprime pas. D'une part, s'il ne les supprime pas, alors le contre-exemple enfreint le PNNP puisque le fait que le déterminisme supprime les possibilités alternatives est l'équivalent de P2, une autre prémisse de l'argument de la conséquence (P2 : si mes actions sont impliquées par une chose sur laquelle je n'ai aucun pouvoir, alors je n'aurais pas pu agir autrement). D'autre part, si le déterminisme supprime les possibilités alternatives, alors comme le simple fait de postuler qu'il les supprime entraîne une pétition de principe, le contre-exemple de Frankfurt entraîne, de fait, une pétition de principe.

Si, au contraire, postuler que le déterminisme supprime les possibilités alternatives dans les contre-exemples de Frankfurt n'entraîne pas automatiquement une pétition de principe, alors il faut tout de même montrer que l'agent contrefactuel a vraiment un rôle dans un monde déterministe (et pas simplement qu'il est conceptuellement possible qu'il ait un rôle). Le problème, c'est que tant et aussi longtemps que nous sommes obligés de postuler que le déterminisme supprime les possibilités alternatives, il n'est pas si clair que nous aurons identifié un cas où un agent contrefactuel supprime les possibilités alternatives.

### **E. La relation « rendre inévitable » et les agents contrefactuels dans un monde déterministe**

Nous présenterons dans cette section un second argument pour montrer qu'il est illégitime de construire des contre-exemples de Frankfurt dans un monde déterministe. Pour ce faire, nous procéderons en deux temps. En premier lieu, nous retracerons la version de Widerker de la défense du dilemme (2003). En deuxième lieu, nous ferons la démonstration que cette analyse suppose une façon particulière de comprendre le rôle des agents contrefactuels (défendue entre autres par Garnett (2013)) dans des situations déterministes et nous argumenterons en faveur

de cette compréhension en l'opposant à plusieurs autres compréhensions concurrentes. Nous montrerons que cette analyse est incompatible avec la construction de contre-exemples de Frankfurt dans un contexte déterministe, y compris des contre-exemples qui entendent précisément répondre à l'argument de Widerker.

Ce processus argumentatif semble d'autant plus pertinent que ses conséquences sont grandes, qu'il soit accepté ou rejeté : d'un côté, celui qui accepte notre analyse des agents contrefactuels nous accorde que nous avons posé un jalon de plus sur la voie du rejet des contre-exemples de Frankfurt; d'un autre côté, celui qui refuse notre analyse des agents contrefactuels ne peut en même temps accepter la défense du dilemme présentée par Widerker, un argument considéré par plusieurs comme la plus sérieuse menace (ou la plus centrale) aux contre-exemples de Frankfurt (Haji et McKenna 2006, Hunt 2007, Fischer 2010, Widerker et Goetz 2013, Garnett 2013). C'est dire que l'analyse qui suit devrait clarifier les positions de chacun.

### **1) L'argument de Widerker**

Si on en croit Frankfurt (1969, 2003) et Widerker les contre-exemples de Frankfurt dépendent du postulat IRR (les agents contrefactuels sont causalement « *irrelevant* ») :

(IRR) Il peut y avoir des circonstances qui n'entraînent en aucun cas qu'une personne exécute une certaine action; néanmoins, ces circonstances précises rendent impossible pour cette personne d'éviter d'exécuter cette action.  
(Widerker 2006a, p. 175)

Imaginer une situation IRR implique la négation du déterminisme. Il est difficile de le voir sans inspecter l'argument proposé par Widerker (1995a, 2003) pour montrer que les situations IRR sont impossibles. Voici le premier exemple qu'il analyse :

Jones doit décider s'il mentira ou non à son oncle. Un agent contrefactuel, Black, tient absolument à ce que Jones mente, mais désire n'intervenir que si cela est nécessaire. Il influencera Jones uniquement si celui-ci ne montre pas de signe (S1) signifiant qu'il brisera sa promesse de ne pas mentir. Jones ment à son oncle pour ses propres motifs et Black n'a pas à intervenir.

La situation ainsi décrite, croit Frankfurt, montre qu'il y a des circonstances (la présence de Black) qui n'entraînent pas qu'une personne exécute une certaine action (c'est Jones qui lui-même a décidé de mentir), mais qui rendent la décision inévitable<sup>70</sup> (Black interviendra pour corriger la situation si Jones ne manifeste pas qu'il s'apprête à décider de ne pas mentir).

Widerker avance dans son article de 2003 non seulement qu'il ne s'agit pas d'une IRR, mais que celles-ci sont impossibles (pour les cas de décision). Son argument se décline de la façon suivante : soit le signe S1 est causalement suffisant pour que Jones prenne la décision au moment  $t$  de briser sa promesse soit il ne l'est pas. D'un côté, si S1 n'est pas causalement suffisant, alors ce n'est qu'un indicateur moyennement fiable et donc la décision de Jones aurait pu être évitée. En effet, si Black décidait de ne pas intervenir, Jones aurait encore le choix de décider de ne pas mentir à son oncle. La situation brise alors l'une des deux conditions d'une IRR, c'est-à-dire le fait que l'action doive être inévitable. D'un autre côté, si S1 est causalement suffisant, alors l'exemple décrit n'est pas une situation IRR puisque les circonstances précises qui rendent l'action inévitable causent l'action. Ce qui rend la situation inévitable, ce sont les opérations mentales de Jones. *La présence de Black n'a plus de rôle à*

---

<sup>70</sup> Il est assez clair que la question des possibilités alternatives est différente s'il est question d'action ou de prise de décision. En effet, Widerker soutient, avec Frankfurt, que dans le cas d'actions, si S1 n'est pas causalement suffisant, alors il est faux que l'action de Jones aurait pu être évitée. Cela ne veut pas dire que des situations IRR sont possibles puisqu'il est tout à fait cohérent de supposer que Jones ne serait pas moralement responsable d'une action pour laquelle il n'a pas de possibilités alternatives, mais seulement de la décision qui mène à cette action (Brown 2006, pp. 277-278). Cependant, la distinction entre les actes mentaux et physiques ne change rien à la question qui nous occupe plus particulièrement : qu'il s'agisse d'action ou de décision, la présence d'un agent contrefactuel ne nous paraît pas réellement participer à rendre inévitable l'action ou la décision de Jones si le déterminisme est supposé.

*jouer, ce qui veut dire que ce n'est pas cette circonstance précise qui rend l'action inévitable.*

Dans ce cas, Black ne rendrait pas l'action inévitable; sa présence ne ferait que *surdéterminer* le caractère inévitable de l'action de Jones.

La deuxième possibilité, celle où le signe S1 est causalement suffisant pour provoquer l'action de Jones, est analogue à une tentative de construction de situation IRR dans un monde déterministe. Si l'analyse de Widerker est exacte, il faut croire que les contre-exemples de Frankfurt sont impossibles dans un monde déterministe. En effet, si le déterminisme est vrai, l'état mental de Jones et les lois de la nature sont toujours suffisants à rendre inévitable son mensonge. Si cela fait que Black ne rend plus inévitable le mensonge de Jones, alors il n'y a effectivement plus de situation IRR. Une situation IRR suppose donc un contexte indéterministe (si elle est possible). Il faut noter que nonobstant la question de savoir si les contre-exemples de Frankfurt nécessitent réellement une situation IRR, si le fait que Jones et Black se retrouvent dans un monde déterministe fait en sorte que Black n'a aucun rôle, alors des arguments comme celui présenté par Fischer précédemment tombent à l'eau.

Widerker a-t-il raison de soutenir que Black ne joue plus aucun rôle dans le cas où le signe est causalement suffisant à l'exécution de l'action? Ce qui est certain, c'est qu'il répond trop rapidement à cette question (tout comme Funkhouser, Haji et Fischer qui commettent le vice inverse de postuler sans argument que Black *joue un rôle* dans un monde déterministe (Funkhouser 2006, Haji 2006, Haji et McKenna 2006, Fischer 2010, 2013)). Bien sûr, Black surdétermine l'action au sens où il n'est pas une condition nécessaire au mensonge de Jones, mais il pourrait (comme le croient Funkhouser, Haji, McKenna et Fischer) être malgré tout une « condition qui rend le mensonge de Jones inévitable sans pour autant le causer ». Voilà qui pose à nouveau la question : dans un contexte où la décision de Jones est inévitable en

fonction de ses dispositions mentales (par exemple, si le déterminisme est vrai), la présence de l'agent contrefactuel Black change-t-elle les possibilités alternatives de Jones? L'argument que nous allons présenter pour défendre le caractère inoffensif des contre-exemples de Frankfurt à l'égard de PAP\* nous semble avoir un caractère suffisamment général pour qu'il soit indépendant de tentatives futures d'identification de situations IRR.

En effet, notre argument, même s'il est issu de l'analyse de Widerker, ne dépend pas d'elle. Le fait que, dans un monde déterministe, Black ne joue d'emblée aucun rôle puisqu'il est impossible que Jones ne décide pas de mentir par lui-même, est indépendant de l'analyse de « signes ». Cette précision est d'autant plus importante que Widerker lui-même a changé son fusil d'épaule à propos de la possibilité de situations IRR (Widerker 2006a) et en a lui-même proposé quelques-unes sans utiliser de signe comme S1. Notre analyse des agents contrefactuels et de la relation « rendre inévitable » vaut donc autant pour les agents contrefactuels dans les scénarios analysés par Widerker que pour ceux présentés par Fischer, Haji et McKenna.

## **2) Le rôle des agents contrefactuels et l'impossibilité de construire des situations IRR déterministes**

Voici donc notre argument qui montre que les contre-exemples de Frankfurt supposent soit l'indéterminisme, soit des présupposés compatibilistes. Nous disions donc que pour invalider le PAP Frankfurt part du fait qu'intuitivement l'on considère que dans certaines situations certaines personnes sont responsables et que si un agent contrefactuel retire à un individu la possibilité de faire autrement sans jamais intervenir *dans les faits*, alors il ne devrait rien changer à la responsabilité de l'individu. Frankfurt recherche donc une situation IRR, laquelle comporte deux conditions : une action  $y$  est inévitable à cause de  $x$ ;  $x$  ne cause pas  $y$ .

D'emblée, il est possible de soutenir qu'une situation IRR est impossible dans un monde déterministe. On peut croire que *si le déterminisme est vrai, alors ce qui rend inévitable un événement est toujours sa cause et uniquement sa cause* (DIC). Si l'on adhère à DIC, alors une situation IRR est impossible. En effet, si un état  $E$  du monde  $M$  correspond à une situation IRR, alors il existe dans  $M$  des circonstances précises et une action telles que ces circonstances ne causent pas l'exécution de l'action et telles que ces circonstances rendent inévitable l'action. Donc, si un état  $E$  du monde  $M$  correspond à une situation IRR, alors il existe dans  $M$  des actions non causées, ce qui est absurde si le déterminisme est vrai. Conséquemment, une situation IRR est impossible dans un monde déterministe.

DIC est motivée par l'idée selon laquelle dans un monde où le déterminisme est vrai chaque action est bel et bien inévitable<sup>71</sup>, parce qu'elle est le résultat d'une chaîne causale (infinie ou qui remonte aux débuts des temps) qui ne peut être rompue (les actions humaines étant elles-mêmes des éléments de la chaîne causale). Tout autre élément du monde ne peut qu'au mieux surdéterminer la relation  $x/y$  (où  $x$ , la cause, rend  $y$  inévitable). Les agents « contrefactuels » n'auraient donc aucun rôle particulier dans un tel environnement puisqu'il serait d'emblée déterminé qu'ils n'auraient jamais à intervenir. Les expériences de pensée courantes pour faire comprendre le déterminisme comme le fait que si nous reculions dans le temps, les mêmes événements adviendraient peuvent être utiles pour visualiser l'absence de rôle des agents contrefactuels. Prenons le cas où Jones décide de mentir à son oncle sans qu'un agent contrefactuel Black ait à intervenir. Si le déterminisme est vrai, alors Jones n'a pas accès à un monde dans lequel Black aurait à intervenir. En fait, un monde où Black interviendrait

---

<sup>71</sup> Cela ne veut pas dire qu'elle est « nécessaire ». Si elle l'était, alors peu importe les lois de la nature et l'état du monde initial, elle adviendrait, ce qui n'est pas le cas. Quand on dit que chaque action est inévitable, on veut dire



dans ce cas précis ne serait pas « connecté » au monde actuel. Donc, même si on avait à recommencer cette situation un millier de fois, jamais Black ne jouerait un rôle. Ce point de vue se fonde sur l'idée que, comme pour la relation causale, la relation « rendre inévitable » est affectée par un critère comme la notion de « contact ». Cela veut dire que si deux événements,  $x$  et  $y$ , qui se produisent sont chacun suffisants pour rendre un événement  $z$  inévitable, il y en a un des deux qui est vraiment celui qui rend inévitable  $z$ . Et ce serait, dans un monde déterministe, toujours l'événement qui *cause*  $z$ . Le cas échéant, une situation IRR serait impossible dans un monde déterministe. Notons qu'une thèse est implicite et beaucoup plus faible dans cette explication : dans un monde déterministe, si  $x$  cause  $z$ , alors  $x$  contribue d'une manière ou d'une autre à faire que  $z$  est inévitable.

On peut également refuser le principe stipulant que si le déterminisme est vrai, alors ce qui rend inévitable un événement est toujours sa cause et uniquement sa cause (DIC). On suppose alors qu'il est possible que dans un cas où un agent contrefactuel est prêt à intervenir, même si le déterminisme est vrai, il rend l'action d'un sujet inévitable, même s'il n'intervient jamais. On repousse alors l'intuition selon laquelle l'agent contrefactuel dans un monde déterministe n'a aucun rôle parce que dans tous les mondes possibles accessibles au monde actuel, il n'intervient jamais. L'agent contrefactuel aurait justement un rôle dans la relation « rendre inévitable », mais pas dans la relation causale. De quel rôle s'agit-il? On peut imaginer deux réponses. L'agent contrefactuel pourrait rendre inévitable l'action à titre de membre d'une disjonction ou d'une conjonction.

---

qu'un agent donné n'aurait pas pu éviter d'accomplir cette action, mais pas qu'il l'aurait exécutée dans tous les mondes possibles.

Afin de mieux illustrer ce que cela signifie, prenons le cas de Susie et Billy, deux enfants qui lancent une pierre sur une bouteille de verre avec précision. Susie touche la cible avant Billy : la pierre de Billy ne fait donc qu'atteindre des éclats de verre. La question classique concernant cet exemple concerne la causalité : les deux tirs sont suffisants pour causer le bris de la bouteille, mais sont-ils tous les deux la cause de ce bris? La meilleure réponse à cette question semble être tout simplement que c'est le tir de Susie qui a causé le bris (Taylor et Dennett 2011). Cela pourrait se justifier en vertu du critère de priorité temporelle (sa pierre touchant la bouteille de verre en premier, c'est elle qui briserait cette bouteille) ou en vertu du critère de « contact » : sa pierre étant celle qui entre en contact avec la bouteille, c'est elle qui cause le bris. Si, par ailleurs, les deux pierres avaient percuté la bouteille en même temps et avec la même force, la réponse à la question de savoir quelle pierre a causé le bris serait très différente, les critères de contact et de priorité temporelle ne pouvant pas s'appliquer. La réponse qui nous semble la plus adéquate serait que c'est la disjonction des deux tirs qui cause l'éclat de la bouteille. Pourquoi? Tout simplement parce que pour que l'énoncé « le tir *A* ou le tir *B* a causé l'éclat de la bouteille » soit vrai, il faut soit que le tir *A*, soit que le tir *B*, soit que les deux tirs aient été effectués. La disjonction des deux lancers est une condition nécessaire au bris de la bouteille. En termes de contrefactuels, s'il n'y avait eu ni tir *A* ni tir *B*  $\neg (A \vee B)$ , alors la bouteille n'aurait pas éclaté. Ce n'est pourtant *que parce que les critères temporels et de contact ne s'appliquent pas* que, dans cette dernière situation, nous serions enclin à penser que c'est la disjonction des deux tirs qui est la cause du bris de la bouteille plutôt que le tir de Susie. En effet, si Susie avait lancé sa pierre en premier, jamais nous n'aurions voulu dire que c'était la disjonction des lancers de Susie et de Billy qui était la cause de l'éclat de la bouteille.

Reprenons le premier exemple (celui où Susie lance la première), mais plutôt que de nous poser la question de la relation causale, inspectons la relation « rendre inévitable ». Postulons aussi que la situation se produit dans un monde déterministe : cela veut dire que le bris de la bouteille est inévitable. Quel événement précis rend ce bris inévitable?

Un premier choix consisterait à dire qu'en fait les circonstances qui rendent le bris de la bouteille inévitable sont tout simplement la conjonction des lois de la nature et de l'état du monde à un temps antérieur au moment du bris de la bouteille. Cette analyse rend laborieuse l'identification d'une situation IRR : la conjonction des lois de la nature et de l'état du monde à un temps antérieur cause-t-elle les événements ultérieurs? Si oui, alors les situations IRR sont bel et bien impossibles. Sinon, alors toutes les situations répondent aux deux conditions des situations IRR : il existe quelque chose, la conjonction des lois de la nature et de l'état initial du monde, qui rend inévitable tous les événements sans les causer. Dans tous les cas, le fait que les lois de la nature et les événements du passé participent, même indirectement, aux événements postérieurs n'en font pas des agents contrefactuels adéquats pour les contre-exemples de Frankfurt.

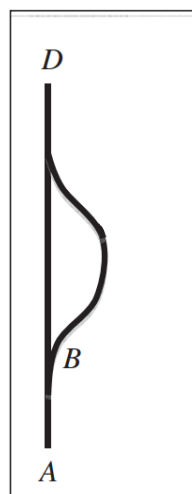
Un second choix, que nous avons anticipé précédemment, consisterait à dire que c'est uniquement le lancer de Susie qui rendrait l'éclat de la bouteille inévitable<sup>72</sup>. On aurait à ce moment une analyse de la relation « rendre inévitable » similaire à notre analyse de la relation causale, c'est-à-dire que le critère de « contact » pointerait vers un événement particulier qui serait « celui qui rend réellement inévitable » (dans ce cas, le lancer de Susie). Si c'est la meilleure analyse, alors une situation IRR est impossible dans un monde déterministe, comme

---

<sup>72</sup> Chacun des choix 2-4 pourraient aussi être enrichis par l'ajout de « ou la conjonction des lois de la nature et d'un état du monde précédent ». Le cas échéant, une situation IRR serait encore une fois, soit impossible soit présente partout.

nous l'avons vu. Accepter cette analyse nous amène donc à soutenir DIC. C'est le choix que semble également nous recommander Garnett (2013), choix qu'il motive par l'analyse suivante :

Suppose that a main track runs from *A* to *D*, with another track branching off at *B*. The set of points that would switch a train onto the branch-line at *B* are rusted shut. So a train starting from *A* has no choice but to continue straight on to *D*. This, I take it, is analogous to how an incompatibilist pictures a normal agent, such as Jones, facing an 'alternative' under determinism: determinism, like the rusted points, blocks off any genuine possibility of Jones doing otherwise. Suppose next that the branch-line that diverges at *B* in fact simply loops round and rejoins the main track, ending also at *D*. So it is now the case that, even were a train to switch over at *B* (despite the rusted points), it would still end up at *D* [See Image 1]. The looping branch-line therefore plays a role with respect to the train similar to the role that Black plays with respect to Jones (though this case is not, and is not intended to be, an exact analogy to a Frankfurt case). That is, even were Jones to begin to choose otherwise (despite determinism), Black would ensure that Jones would still end up back performing the action he is in fact determined to perform. This is then a genuine case of overdetermination. We have two perfectly good explanations of why the train will end up at *D*: because the points are rusted shut and because the branch-line loops round<sup>73</sup>. The same is true of why Jones will perform his action: because determinism is true and because of Black's presence. Were this the whole of the situation, Fischer would have his premise (1) [F4]. Yet matters are more complex. [...]



<sup>73</sup> Ceci correspond à la quatrième explication... Image 1 ... du rôle des agents contrefactuels dans un monde déterministe.

So suppose then that there is a secondary branch-line branching off from the first branch-line at *C* (see Figure 1 [image 2]). This secondary branch-line does not loop back round to *D*. Yet, as with the turning at *B*, the points at *C* are rusted shut. So a train taking the first branch-line will, as before, end up at *D*, but this is now the case only because the points at *C* are rusted shut, and not because of the nature of the branch-line itself. In sum: a train, starting from *A*, will end up at *D*; it will do so, first, because the points at *B* are rusted shut and so it cannot turn off from the main-line; second, because even if it were to turn off at *B*, it would proceed back round to *D*, due to the nature of the branch-line together with the fact that the points at *C* are also rusted shut. Say that we now ask the question: why does a train, starting from *A*, have no alternative but to end up at *D*? Is it (i) because the points are rusted shut at *B* and *C* or (ii) because of the partially looping first branch-line? I take it that the answer is clearly (i). (Garnett 2013, p. 390-392)

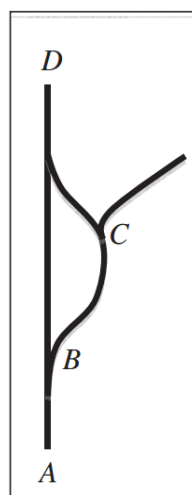


Image 2

La réponse que propose Garnett nous incite à rejeter le quatrième modèle de l'analyse des agents contrefactuels (ci-dessous). On voit très clairement que c'est le déterminisme qui élimine les possibilités alternatives et non l'agent contrefactuel. Cette analyse nous semble correspondre *grosso modo* au deuxième modèle.

Un troisième choix consisterait à dire que la priorité temporelle ou la notion de contact ne sont pas des propriétés pertinentes pour la relation « rendre inévitable » et que ce serait la disjonction des deux lancers qui rendrait inévitable l'éclat de la bouteille. Cette analyse est

motivée par le fait qu'il est *possible* que chacun des lancers soit une condition suffisante, mais non nécessaire, pour que l'événement en question soit inévitable. De plus, elle rend compte de l'intuition selon laquelle il serait étrange que plusieurs événements particuliers rendent chacun un autre événement inévitable. Bref, nous sommes face à une justification similaire à celle de la cause disjunctive. Encore une fois, cette analyse montre qu'une situation IRR est impossible dans un monde déterministe. Car l'action  $x$  serait rendue inévitable par  $s \vee b$  (le lancer de Susie ou le lancer de Billy). Pour que la deuxième condition d'une situation IRR soit satisfaite, il faudrait que  $s \vee b$  ne cause pas  $x$ . Pour que  $s \vee b$  ne cause pas  $x$ , il faudrait que ni  $s$  ni  $b$  ne cause  $x$ . Bien sûr, cela n'est pas le cas (puisque  $s$  cause le bris). Conséquemment, les situations IRR semblent une fois de plus impossibles dans un monde déterministe. En affirmant que ce qui rend  $x$  inévitable est  $s \vee b$ , on ne se prononce pas sur le fait  $s$  seul ou  $b$  seul rend inévitable  $x$ . Ce faisant, un argument comme celui de Fischer est bloqué puisqu'il ne peut pas réussir à montrer que F4 est rendu inévitable par (F0-F3) dans un monde déterministe. En effet, c'est seulement  $(F0-F3) \vee$  (le fait que le déterminisme enlève directement les possibilités alternatives) qu'on déclare être ce qui rend inévitable l'action de Jones. Le modèle ne permet pas à Fischer de prétendre n'être pas coupable d'une pétition de principe.

Un quatrième et dernier choix reviendrait à affirmer que la relation « rendre inévitable » s'applique à chacun des deux lancers par rapport au bris de la bouteille. Le lancer de Susie rendrait le bris de la bouteille inévitable; le lancer de Billy aussi. Ce choix ferait alors abstraction des intuitions précédentes. À ce moment, le lancer de Billy jouerait réellement le rôle d'un agent contrefactuel. Mais quelle intuition, au juste, motive cette interprétation? Que veut-on dire lorsqu'on affirme que le lancer de Billy rend le bris de la bouteille inévitable? Il ne nous semble pas possible de donner un sens à cet énoncé sans avoir en tête une situation

hypothétique où le lancer de Suzy rate sa cible (et où le passé et les lois de la nature sont identiques), ce qui contredit l'hypothèse déterministe. Il faut toujours se rappeler que pour le déterministe incompatibiliste, les mondes possibles différents du nôtre sont tout simplement inaccessibles. Même là, pourrait-on parler d'une situation IRR? Il faut faire la distinction entre deux interprétations d'une situation IRR. La première, c'est qu'il faut tout simplement qu'il existe un événement qui entretienne la relation de rendre inévitable  $x$  sans entretenir celle de causer  $x$ . Selon cette acception, le cas de Billy et de Susie est bel et bien une situation IRR. Mais si une situation IRR exige qu'un événement soit *ce* (l'événement le plus important ou le seul) qui rend  $x$  inévitable et qui ne cause pas  $x$ , alors l'enjeu est très différent. Le lancer de Billy ne peut être vu ni comme le seul ni comme le principal événement qui rend  $x$  inévitable. La première notion d'une situation IRR est-elle suffisante pour un contre-exemple de Frankfurt? Garnett (2013) affirme que oui sans nous procurer de raisons pour le croire. Son analyse nous permet tout de même de rejeter celle de Fischer puisqu'elle considère que l'agent contrefactuel n'a *aucun* rôle à jouer dans la relation de « rendre inévitable ».

Au final, ce que montre notre analyse c'est que le rôle d'un agent contrefactuel dans un monde déterministe est extrêmement limité. Nos deux premières propositions affirment que ce rôle est nul : l'agent contrefactuel n'enlève aucune possibilité, il ne barre aucune avenue menant à différents mondes possibles à un autre agent puisqu'un monde déterministe n'a qu'une voie possible. Nos troisième et quatrième propositions concèdent que le monde n'est pas d'emblée limité et que l'agent contrefactuel *peut participer* (troisième proposition) ou *participe* (quatrième proposition) à la limitation des possibilités de l'autre agent; mais il reste clair que l'agent contrefactuel n'est pas la seule ou la principale chose qui rend l'action d'un autre inévitable. Au contraire, le fait d'être dans un monde déterministe fait qu'à même la

psyché de tout agent est inscrite l'impossibilité d'agir autrement. Et cela rend illégitime la construction d'un contre-exemple de Frankfurt dans un cadre déterministe.

Nous sommes ramené à l'accusation de pétition de principe contre la branche déterministe de l'argument du dilemme de Kane-Ginet-Widerker. Notre analyse des agents contrefactuels nous montre que même si (F0-F3) est suffisant pour déduire (F4) – et cela n'est vrai que si l'on accepte que sous toutes les conceptions paradigmatiques de la notion de possibilité Jones n'a pas de possibilités alternatives si (F0-F3) sont vrais ou si l'on admet que la conception conditionnelle de la notion de possibilité est nécessaire à ces conceptions –, les agents contrefactuels n'ont concrètement aucun rôle dans un monde déterministe ou ils ont, du moins, un rôle absolument insuffisant pour ce que Fischer tente de montrer. Cela nous semble vrai à cause de la supériorité du premier, du deuxième et du troisième modèle de compréhension de la relation « rendre inévitable » et du rôle des agents contrefactuels par rapport au quatrième. Si, donc, c'est réellement le déterminisme qui supprime les possibilités alternatives de Jones, alors nous pouvons accepter P14 : pour accepter qu'un agent soit moralement responsable dans un monde déterministe, il faut implicitement nier le PAP. En effet, une fois que nous acceptons que le déterminisme est vrai et qu'il supprime les possibilités alternatives, nous ne pouvons admettre qu'un agent soit moralement responsable sans supposer qu'il est possible qu'un agent sans possibilités alternatives soit moralement responsable, autrement dit sans supposer que le PAP est faux.

Les contre-exemples de Frankfurt ne constituent-t-ils pas, alors, non seulement une pétition de principe contre l'incompatibiliste, mais *une pétition de principe tout court*, c'est-à-dire une pétition de principe dans le cadre dialectique « étroit » au sens où l'entendent Haji-



McKenna? Tout dépend de la réception de notre façon de concevoir le rôle des agents contrefactuels. Il est possible que notre analyse soit surtout intuitive pour un incompatibiliste. Le cas échéant, on peut tout à fait admettre que nous avons seulement prouvé que les contre-exemples de Frankfurt déterministes forment des pétitions de principe contre l'incompatibiliste.

### **3) Une conception alternative des agents contrefactuels**

La compréhension de la possibilité défendue par Dennett et Taylor (Taylor et Dennett 2011) contredit justement notre analyse des agents contrefactuels. Comme on l'a vu, pour eux, avoir la possibilité de faire une action  $x$  signifie que dans des mondes similaires au nôtre, j'ai réussi à exécuter cette action  $x$ . Cela veut dire que même dans un monde déterministe, j'aurais des possibilités alternatives. Cela veut également dire que l'ajout d'un intervenant contrefactuel a un rôle plus significatif dans les possibilités de l'agent. Prenons un exemple où Susie est seule et lance une pierre sur une bouteille de verre dans un monde déterministe. Selon notre analyse précédente, Susie n'a pas dans cette situation la possibilité de conserver la bouteille pour y mettre du lait; selon celle de Taylor et Dennett, Susie a en effet cette possibilité puisque dans un monde similaire, Susie aurait décidé de ne pas lancer la pierre ou alors elle aurait raté sa cible et aurait conservé sa bouteille. En reprenant la situation précédente, où Billy lance aussi une pierre, il est toujours aussi clair que Susie n'a pas la possibilité de conserver sa bouteille selon notre analyse. Or, dans le cas de Dennett et de Taylor, il faudrait croire que Susie n'a pas non plus la possibilité de conserver la bouteille au moment précis où elle aurait lancé la pierre, et ce, à cause de l'agent contrefactuel Billy (on peut même supposer que Billy est un robot pour faciliter les choses). C'est dire que si l'on adopte l'analyse de Dennett et de Taylor de la notion de possibilité, un agent contrefactuel fait réellement une différence, même dans un

monde déterministe. De plus, la situation de Susie et de Billy est, selon cette analyse, une situation IRR : c'est seulement l'agent contrefactuel qui rend la situation inévitable, même dans un monde déterministe.

Le problème principal de cette réponse possible est qu'elle s'éloigne foncièrement du projet de Frankfurt et de Fischer. Frankfurt comptait séparer le débat concernant le lien entre la responsabilité morale et les possibilités alternatives de celui concernant le lien entre le déterminisme et les possibilités alternatives. Il comptait du même coup permettre à celui qui trouverait ses contre-exemples intuitifs de préserver la notion incompatibiliste de possibilité tout en rejetant le PAP. Fischer, pour sa part, avec sa doctrine du semi-compatibilisme, comptait pouvoir rejeter l'incompatibilité du déterminisme et de la responsabilité morale sans avoir à se prononcer sur la compatibilité entre le déterminisme et les possibilités alternatives. Or, ce que Dennett et Taylor proposent n'a rien à voir avec ces voies mitoyennes : il s'agit pour eux de rejeter la majorité, voire l'intégralité, du cadre conceptuel incompatibiliste. Accepter leur analyse des agents contrefactuels pour sécuriser l'argument en deux étapes de Fischer revient à nouveau à effectuer une pétition de principe, mais cette fois seulement contre l'incompatibiliste (et non contre l'indécis), mais aussi contre le compatibiliste qui accepte que le déterminisme supprime les possibilités alternatives sans pour autant mettre en danger la responsabilité morale (ce qui semble parfois être la position de Fischer (2013)).

En fait, accepter l'analyse de Dennett et de Taylor implique d'abandonner une autre prémisse de l'argument de la conséquence, soit P2 (si mes actions sont impliquées par une chose sur laquelle je n'ai aucun pouvoir, alors je n'aurais pas pu agir autrement). En effet, P2 est la prémisse qui montre que si le déterminisme est vrai, nous n'avons pas de possibilités alternatives. Ce n'est donc pas une façon satisfaisante d'attaquer le PAP, en vertu du PNNP.

En effet, nous nous rappelons que selon le PNNP, il est proscrit de s'attaquer à une prémisse d'un argument en niant implicitement une autre prémisse de ce même argument sauf si l'attaque permet de rejeter les deux prémisses en question. C'est là une confirmation qu'utiliser la compréhension des agents contrefactuels de Dennett et de Taylor pour remettre en question le PAP sans attaquer du même coup le lien entre le déterminisme et les possibilités alternatives, autrement dit attaquer P2, – ce que tentent d'éviter Frankfurt et Fischer – est un procédé dialectiquement illégitime. Il faut donc ou bien rejeter *et le PAP, et P2* (ce que nous rejeterons en III, 2-3), ou bien accepter les deux.

Cela nous ramène à ce que nous disions à la suite des propos de Pereboom (2003) : si nous analysons convenablement les contre-exemples de Frankfurt, nous réalisons qu'il ne s'agit pas de réels contre-exemples au PAP, mais plutôt d'illustrations d'un autre cadre théorique de la compréhension de la responsabilité morale, cadre qui par ailleurs s'engage sans le vouloir sur la voie d'un compatibilisme entre possibilités alternatives et déterminisme.

Après avoir montré par de multiples chemins que les contre-exemples de Frankfurt indéterministes (voir schéma 1) et déterministes échouaient (voir schéma 2), nous pouvons conclure que les contre-exemples de Frankfurt ne peuvent ni se dérouler dans des univers déterministes ni dans des univers indéterministes. Les abandonner nous semble alors la seule solution<sup>74</sup>.

---

<sup>74</sup> Nous éviterons de proclamer, à l'instar de van Inwagen face à un dilemme ayant la même structure logique, que nous sommes face à un « mystère » (van Inwagen 2000).

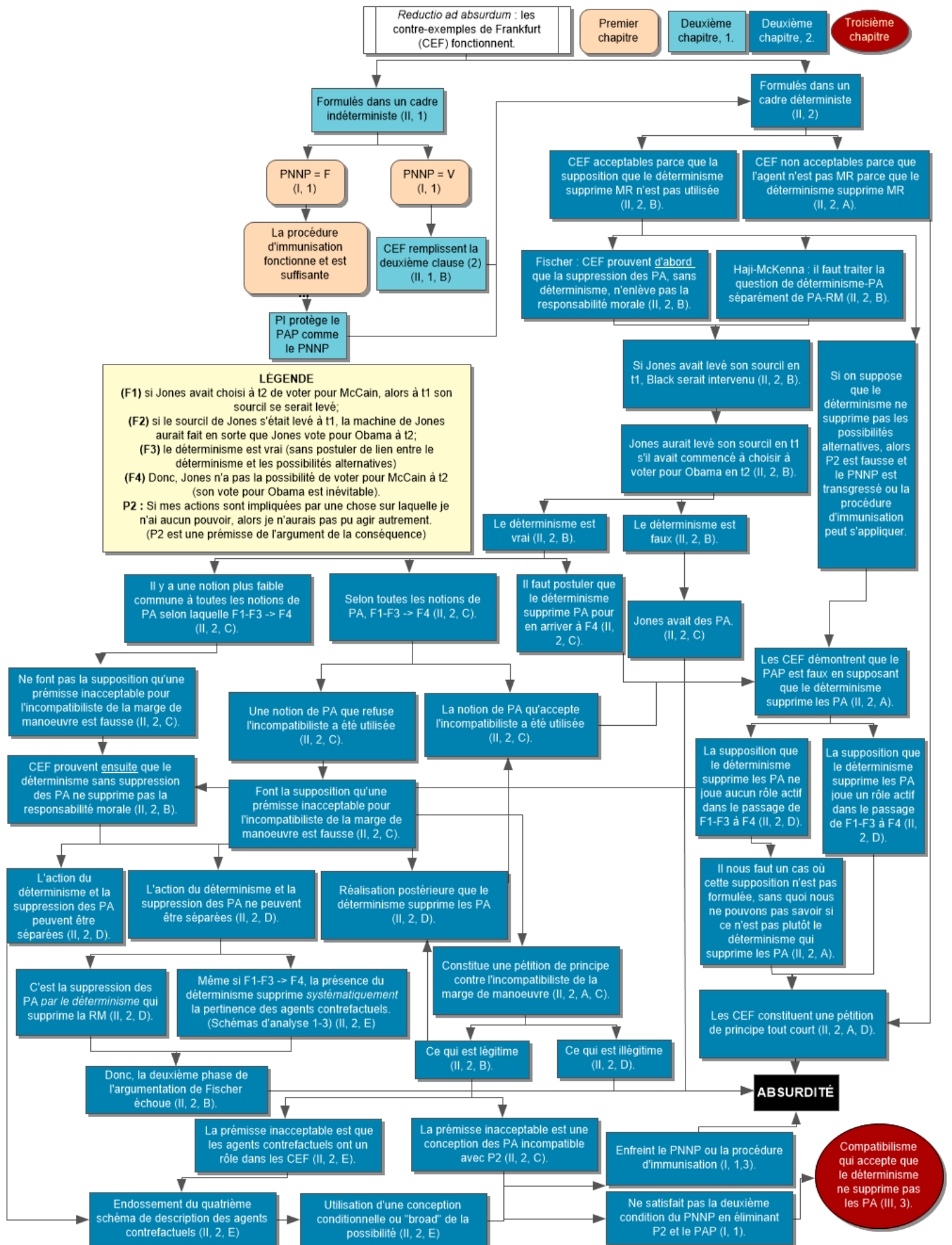


Schéma 2

# Chapitre III : réponse aux répliques de Frankfurt, et de Dennett et Taylor

## 1. La réponse de Frankfurt

La réponse que Frankfurt donne à Widerker sur l'impossibilité de construire une situation IRR dans *Some Thoughts Concerning PAP* (Frankfurt 2003) doit être examinée puisque si elle est acceptable, elle mine tous nos efforts précédents. En fait, Frankfurt diminue l'importance de l'identification réelle d'une situation IRR dans son article, en réponse à l'article de Widerker, qui tentait de faire la démonstration que les situations IRR étaient métaphysiquement impossibles (Widerker 2003)<sup>75</sup>. Même si ses contre-exemples sont métaphysiquement impossibles, croit Frankfurt, ils devraient néanmoins nous amener à dissocier les relations « rendre une action inévitable » et « causer une action ».

The [IRR] examples effectively undermine the appeal of PAP even if it is true that circumstances that do not bring an action about invariably leave open the possibility that the action might not be performed. What the examples are essentially intended to accomplish is to call attention to an important conceptual distinction. They are designed to show that making an action unavoidable *is not the same thing* as bringing it about that the action is performed. [...] Appreciating this distinction tends to liberate us from the natural but nonetheless erroneous supposition that it is proper to regard people as morally responsible for what they have done only if they could have done otherwise. [...] Clarifying the difference between making something unavoidable and making it happen facilitates a recognition that someone who could not have avoided performing a certain action may nonetheless be morally responsible for having performed it because of how he was in fact led to do what he did. (Frankfurt 2003, pp. 339-340)

---

<sup>75</sup> À vrai dire, dans son article, Widerker n'a pas prétendu avoir définitivement écarté la possibilité des situations IRR, se contentant de rejeter plusieurs tentatives individuelles d'en construire.

Nous pourrions adapter la réponse de Frankfurt à notre cas et dire que même si les situations IRR sont indéterministes en vertu d'une nécessité métaphysique, cela n'empêche pas qu'il est conceptuellement possible d'imaginer une situation IRR dans un monde déterministe et que conséquemment notre objection tombe. Tout obstacle à la possibilité métaphysique d'une situation IRR ou à la possibilité métaphysique d'une situation IRR déterministe n'empêcherait pas son rôle : faciliter la distinction entre rendre une action inévitable et causer une action et donc invalider le PAP.

Nous pourrions lui répondre que même en opérant cette distinction (distinction dont nous rendons compte dans notre argument en distinguant  $xIy$  et  $xCy$ ), si, dans notre monde, le fait d'être responsable d'une action s'accompagne toujours (en vertu de propriétés métaphysiques) du fait que nous aurions pu l'éviter, alors le PAP reste vrai. Le principe des possibilités alternatives ne spécifie pas que c'est le fait de n'avoir pas pu faire autre chose que l'action  $x$  qui *cause* l'absence de responsabilité morale pour  $x$ , mais simplement que le fait de n'avoir pas pu faire autrement est une *condition suffisante* à la non-responsabilité morale. Le PAP n'est qu'une implication logique et non une relation de dépendance métaphysique (comme une relation causale) : la responsabilité morale implique la capacité de faire autrement, mais ce n'est pas forcément l'absence de capacité de faire autrement qui engendre la non-responsabilité (même si logiquement elle l'implique). L'absence de la possibilité de faire autrement pourrait impliquer l'absence de responsabilité morale sans pour autant que ce soit elle qui la rende vraie. De la même manière, dans une petite pièce fermée où il n'y a qu'une bougie et deux couteaux, le fait qu'un couteau réfléchisse la lumière implique que l'autre le fasse aussi sans qu'il y ait pour autant de relation causale entre les deux états de choses.

Autrement dit, s'il n'existe pas de contre-exemple au PAP, le PAP est vrai, malgré ce qu'en dit Frankfurt. Si à la question « est-il métaphysiquement possible qu'un individu soit responsable d'une action sans avoir pu faire autrement », la réponse est « non », alors il s'impose à l'esprit que le PAP est vrai. Frankfurt ne fait, au mieux, que montrer que la fausseté du PAP est une possibilité conceptuelle, ce qui n'a rien à voir avec le fait de nous montrer que le PAP est réellement faux. Si les situations IRR sont impossibles, les contre-exemples de Frankfurt ne font que créer le mirage d'une réfutation du PAP; si les situations IRR sont nécessairement indéterministes, elles ne font qu'*illusoirement* miner l'argument incompatibiliste. Il semble donc que la réponse de Frankfurt à Widerker ne le protège pas plus de la tentative de ce dernier de réfuter la possibilité d'une situation IRR que de notre démonstration de l'incompatibilité métaphysique<sup>76</sup> des situations IRR avec le déterminisme.

En somme, la réponse de Frankfurt est assez étrange : il accepte que le PAP puisse être faux même s'il n'a pas de contre-exemple. Or, la plus élémentaire leçon de logique nous apprend que  $A \rightarrow B$  est vrai dès que  $A$  est faux ou  $B$  est vrai. Le PAP a justement la forme  $A \rightarrow B$  et un contre-exemple serait précisément une situation où  $A$  serait vrai et  $B$  serait faux. Dans l'argument de la conséquence, il n'importe absolument pas de savoir si c'est le fait de n'avoir pas pu faire autrement qui élimine la responsabilité morale. Il suffit de savoir que dès que  $x$  est responsable de  $y$ , il aurait pu éviter  $y$ . La réponse de Frankfurt à Widerker ne porte donc pas atteinte à notre thèse, à savoir que les contre-exemples de Frankfurt sont inopérants dans un contexte déterministe. Peut-être avons-nous mal su distinguer une nouvelle conception

---

<sup>76</sup> Dans notre argument précédent, DIC (si toutes les actions du monde  $M$  sont causées et inévitables, alors pour toute action dans  $M$ , les circonstances précises qui causent l'action la rendent inévitable) énonce une vérité métaphysique et non conceptuelle, puisqu'il est conceptuellement possible d'imaginer un monde où tout est causé et inévitable, mais où ce qui cause certains événements n'est pas ce qui les rend inévitables.

de ce qu'est un contre-exemple. Le cas échéant, nous laisserons le soin d'innover sur ce plan à nos contradicteurs.

## **2. La critique de Dennett et de Taylor sur la relation entre déterminisme et causalité**

Daniel Dennett et Christopher Taylor, dans « Who's Still Afraid of Determinism » (2011), s'attaquent à l'argument de la conséquence, dont ils croient qu'il repose essentiellement sur deux prémisses :

- (1) Dans un monde déterministe, quelqu'un ne peut jamais dire avec raison « j'aurais pu faire autrement ».
- (2) Dans un tel univers, quelqu'un ne pourrait jamais être blâmé ou félicité pour avoir causé un événement parce que tous les événements sont déterminés par des conditions qui étaient réalisées lors de la naissance de l'univers.

La première affirmation, croient Taylor et Dennett, repose sur une mauvaise compréhension du rapport entre le déterminisme et les possibilités; la seconde témoigne d'une analyse erronée du rapport entre la causalité et le déterminisme. Commençons par nous pencher sur ce deuxième rapport, qui met en doute nos propres conclusions.

Taylor et Dennett croient que la vérité du déterminisme n'enlève rien à notre rôle causal en tant qu'agents. Ainsi, le fait que Kennedy meure en 1963 n'est pas causé par la conjonction des lois de la nature ( $\lambda$ ) et de l'état initial du monde ( $\sigma_0$ ). La raison en est que  $\lambda \wedge \sigma_0$  est une condition suffisante, mais pas nécessaire à la mort de Kennedy, ce dernier ayant pu mourir en 1963 même si les lois de la nature avaient été différentes. Comme la nécessité est l'élément le plus important de la relation causale selon Taylor et Dennett, on ne peut pas dire que  $\lambda \wedge \sigma_0$  retire à Lee Harvey Oswald son pouvoir causal. Cela fait dire aux deux auteurs que le déterminisme est une thèse concernant la suffisance et non la nécessité et que



conséquemment « puisque la causalité présuppose généralement la nécessité, la vérité du déterminisme devrait n'avoir que peu à faire avec la validité de nos jugements causaux » (Taylor et Dennett 2011, p. 235, traduction libre).

Nous avons précédemment avancé des raisons de croire que si le déterminisme est vrai, alors pour tout événement, ce qui le rend inévitable est sa cause. Si cette solution particulière (nous en avons proposé d'autres) est adéquate, alors Taylor et Dennett ont tort : la vérité du déterminisme lie le caractère inévitable des événements aux événements précédents qui les causent. Le déterminisme a donc un impact sur nos jugements causaux. On pourrait même ajouter : si le déterminisme est vrai, le caractère inévitable de nos actions nous est transmis par le caractère inévitable de ce qui cause nos actions (ce qui se rapproche de DIC).

Taylor et Dennett ont probablement raison lorsqu'ils affirment que la mort de Kennedy en 1963 n'est pas *causée* par  $\lambda \wedge \sigma_0$ ; cela n'enlève pas le fait que pour le déterministe, même si un agent cause certains événements, sa force causale a un sens différent précisément parce qu'elle est le résultat d'une chaîne qui remonte à  $\lambda \wedge \sigma_0$  et qui ne dépend pas de lui. C'est soit en vertu du caractère inévitable de son action, soit en vertu du fait qu'il n'est pas la source ultime de ses actions que le déterministe refuse de croire qu'être une « condition nécessaire » à la production de son action ne signifie rien de plus qu'être un rouage nécessaire au fonctionnement de la chaîne causale. L'idée que sans notre présence le monde aurait été différent (Taylor et Dennett 2011) est donc insuffisante. En prenant un exemple d'action blâmable, la réflexion de Dennett et de Taylor paraît d'ailleurs plutôt contre-intuitive : sachant que le déterminisme est vrai, après avoir commis un meurtre, nous contenterions-nous de savoir que « sans notre présence, l'univers aurait pris une tournure significativement

différente » pour accepter notre culpabilité (en supposant que nous n'avons pas été contraints)?

Au final, nous croyons que Taylor et Dennett font fausse route lorsqu'ils dénoncent (2) – dans un univers déterministe, quelqu'un ne pourrait jamais être blâmé ou félicité pour avoir causé un événement parce que tous les événements sont déterminés par des conditions lors de la naissance de l'univers – parce qu'ils jugent à tort que pour adhérer à (2) il faut croire premièrement que  $\lambda \wedge \sigma_0$  cause les événements actuels et deuxièmement que nous ne sommes pas la cause de nos actions.

### **3. La critique de Dennett et de Taylor sur la relation entre déterminisme et la possibilité**

La justification que nous venons d'esquisser de (2) mobilise d'une façon ou d'une autre (1) – dans un monde déterministe, quelqu'un ne peut jamais dire avec raison « j'aurais pu faire autrement ». Taylor et Dennett auraient tort de croire que (1) et (2) sont « profondément confuses » (2011, p. 211). L'argument pourrait être formulé de la façon suivante<sup>77</sup> :

P17 : C'est parce que le déterminisme est vrai (tous les événements sont déterminés par des conditions lors de la naissance de l'univers) que je n'ai accès qu'à un seul futur possible.

P18 : C'est parce que je n'ai accès qu'à un seul futur que je n'ai pas la possibilité de faire autrement.

---

(1)\* C'est parce que le déterminisme est vrai que je ne peux jamais dire avec raison « j'aurais pu faire autrement ».

(1)\* C'est parce que le déterminisme est vrai que je ne peux jamais dire avec raison « j'aurais pu faire autrement ».

PAP : Je suis moralement responsable d'une action seulement si j'avais pu agir autrement.

---

---

<sup>77</sup> (1) y a été remplacé par (1)\* pour plus de précision.

P19 : Je ne peux être blâmé ou félicité que pour les événements ou actions dont je suis responsable.

---

(2) Dans un univers déterministe, quelqu'un ne pourrait jamais être blâmé ou félicité pour avoir causé un événement parce que tous les événements sont déterminés par des conditions lors de la naissance de l'univers.

Si Taylor et Dennett rejettent (1) (et (1)\*), c'est principalement parce qu'ils refusent P17. Il faut, pour comprendre ce rejet, examiner leur compréhension de la possibilité et de sa relation avec le déterminisme.

Taylor et Dennett distinguent deux façons d'envisager les possibilités : une méthode « étroite » (« *narrow method* ») et ce que nous appellerons une méthode « large ». Soit la proposition  $\phi$  : « Austin compte un but ». Austin rate son coup et le déterminisme est vrai. La méthode étroite suppose que pour savoir si Austin avait la possibilité de  $\phi$ , il faut voir si l'ensemble  $X$  de mondes identiques au monde actuel à un moment antérieur au tir contient un monde où  $\phi$  est vrai. Or,  $X$  ne contient qu'un seul monde, le monde actuel. Donc,  $\phi$  est impossible. La méthode large suppose que pour savoir si Austin avait la possibilité de  $\phi$ , il faut voir si l'ensemble  $Y$  des mondes similaires au monde actuel contient des mondes où  $\phi$  sont vrais. Ainsi, la méthode large permet de dire que dans une partie d'échec où, au septième coup, l'ordinateur effectue un roque, il aurait pu également déplacer son roi d'une case (à moins que le coup en question dépasse sa capacité ordinaire d'analyse). La méthode étroite serait donc inutile pour expliquer des capacités ou des « *patterns* ». La méthode large, au contraire, permettrait d'apporter une compréhension ainsi que des éléments moralement pertinents supplémentaires sur la situation. Taylor et Dennett en déduisent que le fardeau de la preuve à l'égard de l'utilisation de la méthode étroite repose sur les épaules de l'incompatibiliste.

Il nous semble que Taylor et Dennett présentent d'excellents arguments pour croire que le cadre déterministe n'est pas adéquat dans la pratique, mais non pour croire qu'un agent peut avoir des possibilités alternatives même dans un monde déterministe. En effet, le fait que la méthode large soit plus informative et plus intéressante pour prédire certaines actions n'entraîne pas que les agents ont un accès aux possibilités inspectées par cette méthode. En effet, comment un agent peut-il avoir accès à un monde dont le passé est différent que celui dans lequel il se trouve? Au contraire, il semble toujours très intuitif que si un agent est dans un monde indéterministe, il aura accès à plusieurs mondes dont le passé est identique à celui dans lequel il se trouve; s'il est dans un monde déterministe, il n'aura accès qu'à un seul monde.

Par ailleurs, Taylor et Dennett affirment qu'entre un monde déterministe et indéterministe, il n'y a aucune différence de possibilités, de compétences ou d'opportunités qui en valent la peine puisque les conséquences sont les mêmes et parce que nous ne pouvons pas distinguer l'un de l'autre (Taylor et Dennett 2011, p. 234). C'est, nous croyons, aller trop loin : dans un monde où je suis un agent indéterministe, j'ai accès à davantage de mondes possibles que si je suis un agent d'un monde similaire, mais déterministe. L'illusion de possibilités alternatives ne saurait correspondre à de réelles possibilités alternatives! Il est bien possible qu'au final l'existence de possibilités alternatives ne soit pas particulièrement souhaitable. Mais cette raison, que Taylor et Dennett utilisent pour soutenir leur conception de la possibilité, n'est en réalité qu'un motif de défendre une forme d'illusionnisme et non une forme de compatibilisme.

La doctrine déterministe entraîne vraisemblablement que le fait que je ne m'envole pas en battant des bras ou que je ne me sois pas prévalu de mon droit de vote hier sont des faits

aussi inévitables l'un que l'autre. L'idée même de capacité est altérée dans un monde déterministe puisque, par exemple, la capacité de nager ne donne accès à un monde dans lequel je nage que si la conjonction des lois de la nature et de l'état initial du monde l'admet. Rejeter l'incompatibilisme parce que la conception de la possibilité qu'elle utilise est peu utile ou trop éloignée du sens commun est problématique : l'adhésion au déterminisme a souvent été explicitement opposée au sens commun, associant certains présupposés sur notre liberté à notre ignorance (Spinoza 1965/1677). C'est aussi une doctrine qui n'entend pas réformer notre façon de prévoir les événements probables, mais bien de jeter un éclairage sur notre passé pour nous orienter dans nos jugements moraux.

Nous concluons que les raisons pour lesquelles Taylor et Dennett rejettent l'incompatibilisme sont des raisons pratiques : la méthode large pour examiner les possibilités est beaucoup plus utile que ne l'est la méthode étroite puisqu'elle permet raisonnablement de prédire certains comportements. Elle permet en effet d'attribuer des capacités à certaines entités et d'en tirer des prédictions. En revanche, l'idée selon laquelle avoir des possibilités au sens large change quelque chose du point de vue moral nous paraît plutôt affirmée que défendue dans le texte de Dennett et de Taylor.

Nous croyons alors pouvoir continuer à défendre DIC : si toutes les actions du monde M sont causées et inévitables (déterminisme), alors pour toute action dans M, les circonstances précises qui causent n'importe quelle action sont ce qui la rend inévitable. La méthode « large » pour concevoir les possibilités n'aurait pu admettre comme « inévitables » que certains événements, liés aux capacités des agents (par exemple, il aurait été inévitable que je ne m'envole pas en battant des mains).

Taylor et Dennett croient pouvoir démonter l'argument de la conséquence à partir de leur analyse, laquelle leur permet d'attaquer Beta, une prémisse similaire à P2 : « Si  $A$  a le pouvoir de causer  $\alpha$  et que  $\alpha \rightarrow \beta$  est vrai dans tous les mondes possibles, alors  $A$  a le pouvoir de causer  $\beta$  ». Autrement dit, Beta stipule que pour que  $A$  ait le pouvoir de causer  $\beta$ , il suffit que  $A$  ait le pouvoir de causer une condition suffisante à  $\beta$ . Cette idée ne capture pas, selon Taylor et Dennett, l'idée selon laquelle la nécessité est l'élément le plus important de la relation causale.

L'idée de Taylor et Dennett peut très bien être interprétée dans le cadre de la version de l'argument de la conséquence sur laquelle nous nous sommes précédemment attardé :

P1 : Le déterminisme est vrai, ce qui signifie que la conjonction des lois de la nature et de l'état du monde à un temps  $t$  **implique** l'état du monde à un temps  $t_{+/-n}$ , incluant chacune de mes actions.

P2 : Si mes actions **sont impliquées** par une chose sur laquelle je n'ai aucun pouvoir, alors je n'aurais pas pu agir autrement.

P3 : Je n'ai de pouvoir ni sur un état du monde précédant ma naissance ni sur les lois de la nature.

PAP : Je suis moralement responsable d'une action seulement si j'avais pu agir autrement.

---

C1 : Je ne suis responsable d'aucune de mes actions.

L'analyse de Dennett et de Taylor contraint à clarifier la relation « implique » (ou, dans d'autres versions de l'argument, la relation « être la conséquence de »). Il est évident qu'il doit s'agir de la même notion dans P1 et dans P2 sans quoi l'argument est invalide. De deux choses l'une : d'un côté, si « impliquer » est une relation strictement logique, alors Taylor et Dennett acceptent P1, mais refusent P2. Pour eux, que mes actions soient les conséquences logiques des événements du passé ne change rien à mes possibilités. D'autre part, si « impliquer » est une relation strictement causale, alors Taylor et Dennett acceptent P2, mais refusent P1. Ils

prétendent en effet que les événements d'un passé lointain ne sont pas la cause de mes actions et que ce sont les relations causales qui limitent mes possibilités.

Fischer a tenté de répondre à l'attaque de Dennett et Taylor de l'argument de la conséquence (Fischer 2005a) en en proposant une seconde version, fondée sur le principe de fixité du passé et des lois. Taylor et Dennett y répondent (Taylor et Dennett 2011, p. 240, note 30) en disant que ce principe n'est que la réaffirmation de la méthode étroite de compréhension des possibilités. Ayant argumenté contre leur conception large de la possibilité, nous croyons pouvoir défendre une version de l'argument de la conséquence similaire à celle de Fischer :

P20 : Dans le monde actuel, le déterminisme est vrai : la conjonction de l'état du monde à un temps  $t$  ( $\sigma_t$ ) et des lois de la nature  $\lambda$  implique l'état du monde à un temps  $t_n$  ( $\sigma_{t_n}$ ).  $(\sigma_t \wedge \lambda) \rightarrow \sigma_{t_n}$

P21 : Si  $A$  a le pouvoir de causer  $\neg \phi$  au temps  $t_n$ , alors il existe un monde, identique au monde actuel jusqu'immédiatement avant le temps  $t_n$ , où  $\neg \phi$  est vrai.

P22 : Si  $\sigma_t \wedge \lambda = T$  dans ce monde et dans tous les mondes identiques jusqu'au temps  $t_n$ , alors  $\phi = T$  dans ce monde et dans tous les mondes identiques au monde actuel au temps  $t_n$ <sup>78</sup>.

C5 : Donc,  $A$  n'a pas le pouvoir de causer  $\neg \phi$ .

Bien sûr, P21 est l'incarnation de la méthode étroite. Il s'agit en fait d'inspecter uniquement les mondes auxquels l'agent a accès au moment de poser son action. Cet ensemble de mondes n'en contient qu'un seul si le déterminisme est vrai : le nôtre. On comprend difficilement comment des mondes similaires, mais non identiques au monde actuel pourraient être accessibles à l'agent, puisque celui-ci ne peut pas changer le passé.

---

<sup>78</sup> Il est très clair que c'est la conceptualisation de la notion de possibilité qui constitue le cœur de l'argument. C'est ce qui ressort du débat entre Tomberlin et Lycan sur la possibilité d'un argument qui montrerait l'incompatibilité *logique* entre le déterminisme et la liberté (Tomberlin 1988, Lycan 2003).

# Conclusion

## 1. Récapitulatif

Nous n'avons pas cru faire fausse route en affirmant qu'il est illégitime de construire un contre-exemple à une prémisse en niant pour ce faire une autre prémisse de l'argument principal (sauf quand le contre-exemple réfute les deux prémisses). Ce principe, le principe de non-négation des prémisses, nous semble assez exigeant pour faire office de garde-fou contre de nombreuses dérives dans le cheminement argumentatif sans pour autant obstruer la vue de celui qui le poursuit. Pour celui qui accepte que le PNNP est vrai (I) et qu'il s'applique aux contre-exemples de Frankfurt (II, 1), le choix de la plupart des pourfendeurs du PAP (Pereboom 2003, 2009a, 2001, 2012, Widerker 2006a, 2009, Mele et Robb 1998, Haji et McKenna 2006, Hunt 2007), celui de formuler des contre-exemples de Frankfurt indéterministes, pourra sembler être un véritable cul-de-sac. Nous avons montré qu'il n'est pas possible de prétendre le contraire dès lors qu'on est confronté à la version réajustée (à la suite de manipulations triviales) de l'argument, qui exige, pour invalider PAP\*, que le déterminisme soit vrai. Cela a quelque chose d'étonnant : comment se fait-il qu'autant de philosophes chevronnés aient commis une erreur de la même nature que celle du second électeur dans l'exemple d'Hochelaga ou de Geneviève dans celui du compost? C'est que la transgression du PNNP peut être commise de toutes sortes de manières, de la plus évidente et absurde à la plus subtile et insidieuse. Cela n'empêche pas que c'est la même erreur dans les deux cas, un peu comme les arguments ontologiques et cosmologiques de l'existence de Dieu sont les facettes plus ou moins naturelles d'une même médaille de faussaire :



C'était une démarche totalement contre nature [...] que de vouloir tirer d'une idée forgée de façon entièrement arbitraire l'existence de l'objet correspondant à cette idée [par la preuve ontologique]. [...] [Quant à elle,] la *preuve cosmologique* [...] replace [...] tout sur la voie d'un type de raisonnement dont j'ignore s'il est rationnel ou sophistique, mais qui en tout cas a quelque chose de naturel et qui apporte avec lui, non seulement pour l'entendement commun, mais aussi pour l'entendement spéculatif, la plus grande persuasion [...]. [E]lle fait apparaître comme nouveau un vieil argument sous une forme déguisée et en appelle à l'accord de deux témoins, à savoir le témoin que constitue la raison pure et un autre dont le témoignage est de source empirique, alors que c'est en fait uniquement le premier qui change simplement son costume et sa voix afin d'être considéré comme un second témoin. [...] C'est donc à proprement parler seulement la preuve ontologique procédant à partir de simples concepts qui contient toute la force démonstrative de la prétendue preuve cosmologique; et l'expérience que l'on met en avant ne sert rigoureusement à rien [...]. (Kant 2001/1781/1787, pp. 536-539)

Si nous avons eu raison sur ce point, alors seuls les contre-exemples de Frankfurt *déterministes* sont dialectiquement acceptables. Mais le sont-ils vraiment? Nous avons apporté une réponse négative à cette question en retraçant le débat sur l'accusation de pétition de principe dans les contre-exemples de Frankfurt déterministes. La complexité des débats nous a amené à la poursuite de réponses à deux questions.

D'abord, nous nous sommes demandé si postuler la responsabilité morale d'un agent déterminé constituait une pétition de principe *contre l'incompatibiliste de la marge de manœuvre*. Nous avons vu que la seule façon de l'éviter était de montrer que l'agent dans la situation avait perdu ses possibilités alternatives sans qu'il soit nécessaire de postuler que le déterminisme les supprimait (et d'utiliser cette supposition). En revanche, pour en arriver à un tel résultat, il fallait soit faire la preuve que, sous toutes les conceptions classiques de la possibilité, l'agent était dépourvu d'alternative ou encore montrer que les possibilités conditionnelles étaient un ingrédient nécessaire aux autres conceptions de la possibilité. Mais

chacune de ces branches étaient vouées à l'échec : la première menait à une nouvelle disjonction problématique – d'une part, utiliser une conception alternative des possibilités constituait une pétition de principe contre l'incompatibiliste; d'autre part, user de la conception incompatibiliste des possibilités aurait mené à faire la supposition que le déterminisme supprime les possibilités alternatives (ce qui nous aurait ramené à l'accusation de départ) – la seconde nous forçait au final à admettre que les agents contrefactuels participaient à rendre des événements inévitables même dans un monde déterministe. Ce rôle des agents contrefactuels, nous avons montré que ni un incompatibiliste ni un compatibiliste agnostique sur le rapport entre déterminisme et possibilités alternatives ne pouvaient l'admettre.

Ensuite, il fallait savoir si les contre-exemples de Frankfurt impliquaient une pétition de principe *contre un interlocuteur agnostique sur la notion de possibilité et sur la question de la compatibilité*. Nous avons soulevé le danger qu'un tel interlocuteur puisse par la suite se positionner sur le débat et soit amené à rejeter totalement les contre-exemples de Frankfurt, mais nous nous sommes surtout consacré à montrer que la voie était bloquée à cause du rôle des agents contrefactuels dans un monde déterministe. En effet, accepter que les agents contrefactuels aient un rôle dans ces mondes revenait à endosser la conception de Taylor et Dennett de la possibilité, ce qui allait à l'encontre du PNNP pour celui qui désirait uniquement s'attaquer au PAP (leur conception des possibilités amène à rejeter P2). Autrement dit, la tentative de Frankfurt de détourner le débat de l'éternelle opposition entre compatibilistes et incompatibilistes en l'orientant sur le PAP semble échouer. Il est loin d'être sûr que le rejet du PAP et d'une certaine façon d'envisager les possibilités (P2) soit intuitif; mais il semble

carrément incorrect de ne rejeter que le PAP. C'est dire que le PNNP peut s'avérer être un véritable couteau suisse.

Widerker et Fischer rappelaient, respectivement :

Opponents of PAP have tried to come up with other, more sophisticated, examples of IRR-scenarios. These examples have again been criticized by PAP-defenders, and then have again been improved upon or defended by opponents of PAP, and so on. (Widerker 2006a, p. 165)

The jury is still out with respect to the indeterministic horn of the Dilemma defence. (Fischer 2010, p. 323)

Peut-être est-il temps de faire revenir le jury, non pour qu'il rende son jugement, mais pour que nous l'investissions d'une autre cause. Peut-être est-il aussi temps de cesser d'ériger de nouveaux contre-exemples au PAP comme autant de châteaux de sables le long des plages de l'intuition, prêts à être emportés par la marée de l'analyse.

Par ailleurs, notre défense du PAP nous a amené à nous avancer sur le rapport entre le déterminisme, la causalité et les possibilités. Il nous a alors été nécessaire de confronter notre vision à celle de Dennett et Taylor. Convaincu que le déterminisme change quelque chose aux mondes auxquels nous avons accès et à la question de la causalité, nous avons repoussé la manière limitée dont Taylor et Dennett comprennent le déterminisme : il s'agit de bien plus qu'une doctrine qui se borne à la suffisance. Leur préconisation de la méthode « large » pour comprendre les possibilités nous a quant à elle paru insatisfaisante parce qu'elle reposait sur leur valeur pratique et prédictive plutôt que sur leur plausibilité métaphysique. Du même coup, nous avons pu défendre à nouveau l'argument de la conséquence avant d'en proposer une nouvelle version.

## **2. Les conséquences de notre analyse sur l'incompatibilisme de la source**

La vaste influence de Frankfurt nous proscrit l'exhaustivité dans l'énumération des conséquences qu'il faudrait tirer de notre analyse. Nous nous contenterons de quelques remarques sur l'incompatibilisme de la source. En effet, si notre analyse a jusqu'ici montré que les contre-exemples de Frankfurt ne réfutaient pas l'argument de la conséquence, force est pourtant de constater que peu importe la valeur réelle de ces contre-exemples, de nombreuses façons intéressantes d'articuler une vision incompatibiliste sans faire appel aux possibilités alternatives ont été récemment proposées (entre autres, Kane 1989, Pereboom 2003, 2011, 2012, Shabo 2010, Yang 2012). Ces théories peuvent toutes être regroupées sous l'épithète « de la source » et même si dans les faits ces incompatibilismes trouvent la plupart du temps leur origine dans le rejet du principe des possibilités alternatives, nous faisons nôtre la catégorisation de Tognazzini (2011), lequel propose de les définir comme suit : l'incompatibilisme est vrai parce que le déterminisme élimine le fait d'être la source ultime, ce qui est requis pour la responsabilité morale. Ce genre de catégorisation permet en réalité à un incompatibiliste d'adhérer aussi bien à la thèse selon laquelle la responsabilité morale nécessite d'être la source ultime de nos actions qu'au PAP.

C'est le cas, notamment de Robert Kane (1989, 1996, 2001), qui prétend que le principe des possibilités alternatives est vrai parce qu'être la source ultime de nos actions nécessite ultimement des possibilités alternatives. Autrement dit, pour lui, la raison pour laquelle le déterminisme supprime notre responsabilité morale est que, s'il est vrai, nous ne sommes pas être la source ultime de nos actions et cela s'explique entre autres parce que pour être la source ultime, ou initiale, de nos actions, il faut bien qu'il y ait au moins un

moment où nous ayons des possibilités alternatives. Des débats récemment initiés sur cette question sont en cours (Campbell 2006, Timpe 2007, Hermes et Campbell 2012, Yang 2012). Bien que plusieurs questions soient abordées, les éternels contre-exemples de Frankfurt constituent encore une fois la motivation principale de ceux qui défendent le « *direct source incompatibilism* », l'incompatibilisme qui soutient que les possibilités alternatives ne sont jamais nécessaires, mais que le déterminisme est tout de même faux. Cela est tellement vrai que l'adhésion aux contre-exemples de Frankfurt est considérée comme faisant partie du canon de cette théorie (Campbell 2006). Évidemment, notre analyse permet de trancher le débat puisqu'elle illustre l'inefficacité des contre-exemples de Frankfurt.

La seconde remarque concerne la question de la pétition de principe pour un incompatibiliste. Il est vrai que notre propos a été la plupart du temps défensif : nous avons montré que certaines attaques formulées contre l'argument de la conséquence ne tenaient pas la route, mais nous n'avons pas particulièrement argumenté en faveur de la plausibilité des prémisses qu'il emploie (sauf par moment pour le PAP et pour P2). Notre argumentation par le PNNP, par exemple, a consisté à montrer que certains attaquaient à tort le PAP en isolation. Comme nous l'avons mentionné précédemment, la question du fardeau de la preuve est trop difficile à résoudre pour présumer d'emblée qu'il repose sur les épaules de nos adversaires. Dans ce contexte, l'incompatibiliste de la source est l'interlocuteur idéal : même s'il rejette le PAP, il est prêt à accepter certaines idées, notamment que c'est l'influence du déterminisme qui nous prive de notre responsabilité morale. Pour cette raison, il est plus enclin à remettre en cause la légitimité dialectique de procédés comme celui de Fischer et va, dans plusieurs cas, préférer construire des contre-exemples de Frankfurt indéterministes (Pereboom 2003). Qui plus est, l'incompatibiliste de la source ne risque pas d'accepter une notion de la possibilité sur

la base d'arguments comme ceux mobilisés par Dennett et Taylor puisqu'il croit qu'il y a une différence fondamentale entre la liberté dont nous disposons dans un monde déterministe et dans un monde indéterministe, différence dont ne peut pas du tout rendre compte cette dernière conception de la possibilité. Il a donc encore plus de raisons qu'un Frankfurt ou qu'un Fischer de rejeter le quatrième modèle du rôle des agents contrefactuels dans un monde déterministe. Cela le confine à un exercice qui ne peut être légitime puisqu'il brave le PNNP : construire des contre-exemples de Frankfurt indéterministes. Cet exercice peut sembler encore plus illégitime pour un déterministe dur comme Pereboom : comme ce dernier adhère au déterminisme, il est étrange de le voir naviguer dans des mondes indéterministes pour y chercher ses intuitions morales. Comme nous l'avons vu, il y a plusieurs raisons qui nous portent à croire que le PAP puisse être vrai dans un monde déterministe et faux dans un monde indéterministe, ce qui rend inutile pour un partisan du déterminisme la recherche de contre-exemples de Frankfurt indéterministes.

Enfin, et plus généralement, notre argument récuse que le choix de l'incompatibilisme de la source au détriment de l'incompatibilisme de la marge de manœuvre (MM) s'effectue en vertu des contre-exemples de Frankfurt. Si Pereboom a raison d'affirmer que les versions contemporaines des positions de la source sont typiquement motivées par les contre-exemples de Frankfurt (Pereboom 2012) », alors cette remarque touche d'autres théories incompatibilistes. Certaines attaques de Pereboom contre différentes versions de l'incompatibilisme MM (par exemple, le libertarisme de Kane (Pereboom 2009c, p. 27)) doivent être revues. Bref, il est nécessaire de comparer entre elles les valeurs respectives des incompatibilismes MM et de la source sur une base différente des contre-exemples de Frankfurt. Il est bien possible, après tout, que le principe des possibilités alternatives soit vrai,

mais que *ce en vertu de quoi un agent est moralement responsable* ne soit pas ses possibilités alternatives. La dichotomie entre incompatibilisme de la source et incompatibilisme de la marge de manœuvre est peut-être appelée à tomber.

C'est ici que prend fin notre analyse critique des contre-exemples de Frankfurt. Même si nous les rejetons, nous n'en croyons pas moins qu'il s'agit d'expériences de pensée remarquables : ces scénarios forcent les incompatibilistes à aller au fond de leur pensée, à inspecter de nouveaux territoires comme le rôle des agents contrefactuels, à approfondir leur compréhension des ingrédients de la responsabilité morale (comme le fait d'être la source de nos actions), à clarifier leur compréhension de la notion de la possibilité et enfin à mieux saisir ce que signifie la construction d'un contre-exemple et les limites inhérentes aux constructions de ce genre. S'il est une leçon à tirer de toute cette entreprise, c'est sans doute qu'il est toujours plus difficile qu'on pourrait le croire de concocter des théories intermédiaires capables de réchauffer une place aux intuitions de ceux qui croient que le déterminisme nous transforme presque en machines, mais en même temps de ceux qui s'accommodent le plus confortablement du monde de la pensée que si nous avions eu d'autres désirs, nous aurions agi autrement. N'est-il pas enfin remarquable que la logique formelle, que l'on a souvent associée à l'atomisme du début du 20<sup>e</sup> siècle et à ses défauts, nous guide grâce à ses théorèmes les plus simples, vers une façon partiellement holistique de considérer les arguments?

## Bibliographie

- Alvarez, Maria. (2009) "Actions, Thought-Experiments and the 'Principle of Alternate Possibilities'." *Australasian Journal of Philosophy*, vol. 87, no. 1, pp. 61 – 81.
- Ayer, A. J. (1954) "Freedom and Necessity." in Cahn, ed. *Philosophical Essays* (St).
- Bates, Tomas. (2007) "An Argument against the Principle of Alternative Possibilities." *British Journal of Undergraduate Philosophy*, vol. 2, p. 1.
- Berofsky, Bernard. (2003) "Classical Compatibilism: Not Dead Yet." in McKenna et Widerker, eds. *Moral Responsibility and Alternative Possibilities* (Ashgate), p 107.
- Bobzien, Susanne. (1997) "Stoic Conceptions of Freedom and Their Relation to Ethics." *Bulletin of the Institute of Classical Studies*, vol. 41, no. S68, pp. 71-89.
- Bonicalzi, Sofia. (2013) "Moral Responsibility Beyond Classical Compatibilist and Incompatibilist Accounts." *Prolegomena: časopis za filozofiju*, vol. 12, no. 1, pp. 21-41.
- Bourget, David et David J. Chalmers. (2013) "What Do Philosophers Believe?" *Philosophical Studies*, pp. 1-36.
- Brown, Vivienne. (2006) "Choice, Moral Responsibility and Alternative Possibilities." *Ethical Theory and Moral Practice*, vol. 9, no. 3, pp. 265-288.
- Campbell, Joseph Keim. (2006) "Farewell to Direct Source Incompatibilism." *Acta Analytica*, vol. 21, no. 4, pp. 36 - 49.
- Carroll, John W et Ned Markosian. (2010) *An Introduction to Metaphysics*. (Cambridge University Press).
- Clarke, Randolph. (2008) "Incompatibilist (Nondeterministic) Theories of Free Will."



- Copp, David. (2006) "Moral Responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities." in Widerker et McKenna, eds. (Ashgate Publishing, Ltd.), pp 265-299.
- Davies, Martin. (2009) "Two Purposes of Arguing and Two Epistemic Projects." *Minds, Ethics, and Conditionals: Themes from the Philosophy of Frank Jackson*, p. 337.
- Deery, Oisín, Taylor Davis et Jasmine Carey. (2014) "The Free-Will Intuitions Scale and the Question of Natural Compatibilism." *Philosophical Psychology*, pp. 1-26.
- Dennett, Daniel C. (1984a) *Elbow Room: The Varieties of Free Will Worth Wanting*. Vol. 36. (MIT Press).
- . (1984b) "I Could Not Have Done Otherwise--So What?" *Journal of Philosophy*, vol. 81, no. 10, pp. 553-565.
- Elzein, Nadine. (2013) "Basic Desert, Conceptual Revision, and Moral Justification." *Philosophical Explorations*, vol. 16, no. 2, pp. 212-225.
- Fischer, J. M. (2013) "The Deterministic Horn of the Dilemma Defence: A Reply to Widerker and Goetz." *Analysis*, vol. 73, no. 3, pp. 489-496.
- Fischer, John Martin. (2005a) "Dennett on the Basic Argument." *Metaphilosophy*, vol. 36, no. 4, pp. 427-435.
- . (2002) "Frankfurt-Type Examples and Semi-Compatibilism." in Kane, ed. *The Oxford Handbook of Free Will* (Oxford University Press).
- . (2010) "The Frankfurt Cases: The Moral of the Stories." *Philosophical Review*, vol. 119, no. 3, pp. 315-336.
- Fischer, John Martin. (1999) "Recent Work on Moral Responsibility." *Ethics*, vol. 110, no. 1, pp. 93-139.

- Frankfurt, Harry. (1969) "Alternate Possibilities and Moral Responsibility." *The Journal of Philosophy*, vol. 66, no. 23, pp. 829-839.
- . (1971) "Freedom of the Will and the Concept of a Person." *Journal of Philosophy*, vol. 68, no. 1, pp. 5-20.
- . (2003) "Some Thoughts Concerning Pap." *Moral responsibility and alternative possibilities. Essays on the importance of alternative possibilities*, pp. 339-345.
- Funkhouser, Eric. (2006) "Comments on David Palmer's "Moral Responsibility, Alternative Possibilities, and Determinism." *Southwest Philosophy Review*, vol. 22, no. 2, pp. 91-93.
- Garnett, Michael. (2013) "Fischer-Style Compatibilism." *Analysis*, vol. 73, no. 2, pp. 387-397.
- Ginet, Carl. (1996) "In Defense of the Principle of Alternative Possibilities: Why I Don't Find Frankfurt's Argument Convincing." *Philosophical perspectives*, vol. 10, pp. 403-17.
- . (1966) "Might We Have No Choice?" in Lehrer, ed. *Freedom and Determinism* (Random House), pp 87-104.
- . (1990) *On Action*. (Cambridge University Press).
- Glock, Hans-Johann. (2008) *What Is Analytic Philosophy?* (Cambridge University Press).
- Goetz, Stewart. (2005) "Frankfurt - Style Counterexamples and Begging the Question." *Midwest Studies in Philosophy*, vol. 29, no. 1, pp. 83-105.
- Haji, Ishtiyaque. (2006) "The Principle of Alternate Possibilities and a Defeated Dilemma." *Philosophical Explorations*, vol. 9, no. 2, pp. 179 – 201.
- Haji, Ishtiyaque et Michael McKenna. (2006) "Defending Frankfurt's Argument in Deterministic Contexts: A Reply to Palmer." *The Journal of Philosophy*, pp. 363-372.

- Hermes, Charles et Joe Campbell. (2012) "More Trouble for Direct Source Incompatibilism: Reply to Yang." *Acta Analytica*, vol. 27, no. 3, pp. 335-344.
- Hoefer, Carl. (2010) "Causal Determinism." *The Stanford Encyclopedia of Philosophy*.
- Hume, David. (2000/1748) *An Enquiry Concerning Human Understanding: A Critical Edition*. Vol. 3. (Oxford University Press).
- Hunt, David P. (2007) "Black the Libertarian." *Acta Analytica*, vol. 22, no. 1, pp. 3-15.
- Kane, Robert. (1985) *Free Will and Values: Adaptive Mechanisms and Strategies of Prey and Predators*. (State University of New York Press).
- Kane, Robert H. (2001) *Free Will*. Vol. 81. (Blackwell).
- . (1996) *The Significance of Free Will*. Vol. 110. (Oxford University Press).
- . (1989) "Two Kinds of Incompatibilism." *Philosophy and Phenomenological Research*, vol. 69, no. December, pp. 219-54.
- Kant, Emmanuel. (2001/1781/1787) "Critique De La Raison Pure (1781, 1787), Trad." *A. Renaut, Paris: Flammarion*.
- Kant, Immanuel / Pluhar, Werner S. (2002) *Critique of Practical Reason*. (Hackett Publishing).
- Langsam, Harold. (2000) "Kant's Compatibilism and His Two Conceptions of Truth." *Pacific philosophical quarterly*, vol. 81, no. 2, pp. 164-188.
- Lycan, William G. (2003) "Free Will and the Burden of Proof." in O'Hear, ed. *Minds and Persons* (Cambridge University Press), pp 107-122.
- Mach, E. (1919) *The Science of Mechanics*. Translated by McCormack. (Cambridge University Press).

- Mach, Ernst. (1889) *Die Mechanik in Ihrer Entwicklung*. Vol. 59. (FA Brockhaus).
- Maier, John. (2013) "The Argument from Moral Responsibility." *Australasian Journal of Philosophy*, pp. 1-19.
- McKenna, Michael. (2011) "Contemporary Compatibilism: Mesh Theories and Reasons-Responsive Theories." *the Oxford handbook of free will*, pp. 175-198.
- . (2008) "Saying Good-Bye to the Direct Argument the Right Way." *Philosophical Review*, vol. 117, no. 3, pp. 349-383.
- Mele, Alfred R. et David Robb. (1998) "Rescuing Frankfurt-Style Cases." *Philosophical Review*, vol. 107, no. 1, pp. 97-112.
- . (1998) "Rescuing Frankfurt-Style Cases." *The Philosophical Review*, vol. 107, no. 1, pp. 97-112.
- Moore, George Edward. (1939) *Proof of an External World*. Vol. 25. (H. Milford).
- Moretti, Luca et Tommaso Piazza. (2013) "Transmission of Justification and Warrant." *The Stanford Encyclopedia of Philosophy*.
- Moya, Carlos J. (2011) "On the Very Idea of Robust Alternative." *Crítica: Revista Hispanoamericana de Filosofía*, vol. 43, no. 128, pp. 3-26.
- Nahmias, Eddy, Stephen Morris, Thomas Nadelhoffer et Jason Turner 1. (2005) "Surveying Freedom: Folk Intuitions About Free Will and Moral Responsibility." *Philosophical Psychology*, vol. 18, no. 5, pp. 561-584.
- Otsuka, Michael. (1998) "Incompatibilism and the Avoidability of Blame." *Ethics*, vol. 108, no. 4, pp. 685-701.
- Palmer, David. (2006) "Moral Responsibility, Alternative Possibilities and Determinism: Begging the Question in the Frankfurt Cases." *Southwest Philosophy Review*, vol. 22, no. 1, pp. 79-86.

- . (2005) "New Distinctions, Same Troubles: A Reply to Haji and Mckenna." *The Journal of Philosophy*, vol. 102, no. 9, pp. 474-482.
- Peijnenburg, Jeanne et David Atkinson. (2003) "When Are Thought Experiments Poor Ones?" *Journal for General Philosophy of Science / Zeitschrift für allgemeine Wissenschaftstheorie*, vol. 34, no. 2, pp. 305-322.
- Pereboom, Derk. (2012) "Frankfurt Examples, Derivative Responsibility, and the Timing Objection1." *Philosophical Issues*, vol. 22, no. 1, pp. 298-315.
- . (2011) "Free Will Skepticism and Meaning in Life." in Kane, ed. *The Oxford Handbook of Free Will* (Oxford University Press).
- . (2009) "Free Will, Love, and Anger." *Ideas y Valores*, vol. 141, pp. 5-25.
- . (2009a) "Further Thoughts About a Frankfurt-Style Argument." *Philosophical Explorations*, vol. 12, no. 2, pp. 109 – 118.
- . (2009c) "Hard Incompatibilism and Its Rivals." *Philosophical Studies*, vol. 144, no. 1, pp. 21 - 33.
- . (2001) *Living without Free Will*. (Cambridge University Press).
- . (2003) "Source Incompatibilism and Alternative Possibilities." in McKenna et Widerker, eds. *Freedom, Responsibility, and Agency: Essays on the Importance of Alternative Possibilities* (Ashgate).
- Pryor, James. (2004) "What's Wrong with Moore's Argument?" *Philosophical Issues*, vol. 14, no. 1, pp. 349-378.
- Robinson, Michael. (2012) "Modified Frankfurt-Type Counterexamples and Flickers of Freedom." *Philosophical Studies*, vol. 157, no. 2, pp. 177-194.
- Schlick, Moritz. (1966) "When Is a Man Responsible?" *Problems of ethics*, pp. 143-56.

- Schnall, Ira M. et David Widerker. (2012) "The Direct Argument and the Burden of Proof." *Analysis*, vol. 72, no. 1, pp. 25-36.
- Shabo, Seth. (2010) "Uncompromising Source Incompatibilism." *Philosophy and Phenomenological Research*, vol. 80, no. 2, pp. 349-383.
- Smilansky, Saul. (2002) "Free Will, Fundamental Dualism, and the Centrality of Illusion." *the Oxford handbook of free will*, pp. 489-505.
- Spencer, Jack. (2013) "Able to Do the Impossible." *Unpublished*.
- Spinoza, Baruch. (1965/1677) *Ethique, Traduction De Charles Appuhn*, Paris, Gf-Flammarion.
- Strawson, P. F. (1974) *Freedom and Resentment and Other Essays*. (Routledge).
- Taylor, Christopher et Daniel Dennett. (2011) "Who's Still Afraid of Determinism? Rethinking Causes and Possibilities." *The Oxford Handbook of Free Will*. Oxford University Press, New York (2011, forthcoming).
- Timpe, Kevin. (2007) "Source Incompatibilism and Its Alternatives." *American Philosophical Quarterly*, pp. 143-155.
- Tognazzini, Neal A. (2011) "Understanding Source Incompatibilism." *The Modern Schoolman*, vol. 88, no. 1/2, pp. 73-88.
- Tomberlin, James E. (1988) "Whither Compatibilism: A Query for Lycan." *Philosophical Papers*, vol. 17, no. August, pp. 127-131.
- van Inwagen, Peter. (1983) "An Essay on Free Will." (Oxford: Clarendon Press).
- . (2000) "Free Will Remains a Mystery." *Philosophical perspectives*, vol. 14, pp. 1-20.
- . (2008) "How to Think About the Problem of Free Will." *The Journal of ethics*, vol. 12, no. 3-4, pp. 327-341.

- . (1975) "The Incompatibility of Free Will and Determinism." *Philosophical Studies*, vol. 27, no. 3, pp. 185-199.
- Widerker, David. (2003) "Blameworthiness and Frankfurt's Argument against the Principle of Alternative Possibilities." *Moral Responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities*, pp. 53-73.
- . (2009) "A Defense of Frankfurt-Friendly Libertarianism." *Philosophical Explorations*, vol. 12, no. 2, pp. 87 – 108.
- . (1995a) "Libertarianism and Frankfurt's Attack on the Principle of Alternative Possibilities." *The Philosophical Review*, vol. 104, no. 2, pp. 247-261.
- . (2006a) "Libertarianism and the Philosophical Significance of Frankfurt Scenarios." *The Journal of Philosophy*, vol. 103, no. 4, pp. 163-187.
- Widerker, David et Stewart Goetz. (2013) "Fischer against the Dilemma Defence: The Defence Prevails." *Analysis*, vol. 73, no. 2, pp. 283-295.
- Widerker, David et Michael McKenna. (2006b) *Moral Responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities*. (Ashgate Publishing, Ltd.).
- Wilkes, Kathleen V. (1988) *Real People: Personal Identity without Thought Experiments*. Vol. 39. (Oxford University Press).
- Wright, Crispin. (2003) "Some Reflections on the Acquisition of Warrant by Inference." *New essays on semantic externalism and self-knowledge*, pp. 57-77.
- Yang, Eric. (2012) "Defending Direct Source Incompatibilism." *Acta Analytica*, vol. 27, no. 3, pp. 325-333.