

Zeitschrift: Revue de Théologie et de Philosophie

Band: 51 (2001)

Heft: 2

Artikel: Société romande de philosophie : séance du 6 mai 2000 : rationalité sans moralité

Autor: Haller, Markus

DOI: <https://doi.org/10.5169/seals-381618>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. [Siehe Rechtliche Hinweise.](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. [Voir Informations légales.](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. [See Legal notice.](#)

Download PDF: 13.10.2024

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

Société romande de philosophie

Séance du 6 mai 2000

RATIONALITÉ SANS MORALITÉ *

MARKUS HALLER

Résumé

Selon David Hume, la raison ne peut pas nous indiquer les fins que nous devrions poursuivre. Les philosophes «anti-humiens» cherchent à réfuter l'argument de Hume et à proposer une conception de la rationalité qui soit basée sur des valeurs pratiques que l'on ne peut pas violer sans contradiction. Les philosophes «néo-humiens» cherchent à compléter la conception de la rationalité de Hume pour qu'elle soit compatible avec nos intuitions. Je défendrai la position néo-humienne en montrant, d'abord, que les deux conceptions anti-humiennes les plus importantes ne nous permettent pas de distinguer les standards de la rationalité de ceux de la moralité. Ensuite, j'expliquerai pourquoi et comment le critère de la consistance des préférences complète la conception de Hume dans la direction voulue et permet le développement d'une conception de la rationalité sans moralité.

Dans la philosophie contemporaine, nous distinguons normalement entre *rationalité théorique* et *rationalité pratique*. La rationalité théorique est un standard évaluatif qui s'applique aux *croyances*. On peut distinguer deux rôles de la rationalité théorique. Elle indique, premièrement, les conditions qui doivent être satisfaites pour qu'une croyance (ou un système de croyances) soit justifiée. C'est son *rôle critique*. Elle exige, deuxièmement, que nos croyances soient justifiées ou au moins justifiables. C'est son *rôle normatif*. Je suppose qu'il est facile d'admettre que nous ne pouvons exiger les uns des autres d'avoir des croyances rationnelles (ou du moins de nous débarrasser de croyances irrationnelles) sans connaître les conditions que nos croyances doivent satisfaire pour être justifiées. Ainsi, le rôle normatif de la rationalité théorique présuppose son rôle critique.

* Ce travail a pu être achevé avec le soutien du Fonds National Suisse (Projet 11-55845.98).

Mais que faut-il entendre par l'expression «être justifiée» ? Je considère comme allant de soi que nous nous référons avec cette expression à une valeur *intellectuelle*, à savoir la valeur de posséder des raisons qui démontrent ou qui suggèrent de manière satisfaisante la *vérité* de ce que l'on croit. Cela explique pourquoi il est naturel de dire qu'il est *bon* d'avoir des croyances rationnelles et *mauvais* de ne pas en avoir. Nous approuvons naturellement l'effort de croire ce qui est vrai et nous blâmons celui qui s'en moque.

La rationalité pratique est un standard évaluatif qui s'applique aux *actions*. Nous pouvons également distinguer son rôle critériel et son rôle normatif. La rationalité pratique indique, premièrement, les conditions qui doivent être satisfaites pour qu'une action (ou un plan, ou une règle d'action) soit justifiée et exige, deuxièmement, que nos actions soient justifiées ou au moins justifiables. Mais il est douteux que nous nous référions dans ce cas, avec l'expression «être justifié», à une valeur *intellectuelle*. Supposons que je poursuive des plaisirs sensuels tout en reconnaissant que je ruine ainsi ma santé ; ou que ma poursuite du prestige et du statut social me fasse renoncer au développement de mes talents et de ma personnalité, bien que je sache que cela me rendra malheureux. Je crois que ces deux exemples illustrent ce que nous appellerions intuitivement des actions irrationnelles. Mais qu'est-ce qui est problématique avec *ma manière de raisonner* dans ces exemples ? Rien du tout, aurait répondu David Hume ; on peut délibérément ruiner sa vie sans mépriser les valeurs intellectuelles. Voici ce qu'il écrit dans son *Traité de la nature humaine* en 1739 :

Il n'est pas contraire à la raison que je choisisse de me ruiner complètement [...] Il est aussi peu contraire à la raison de préférer à mon plus grand bien un bien reconnu moindre et d'aimer plus ardemment celui-ci que celui-là.¹

Hume explique qu'il suffit de distinguer clairement les *croyances* et les *émotions* (ou passions) pour comprendre que la raison est (et doit être) seulement l'esclave des émotions. Son argument peut être résumé comme suit. *Les croyances* représentent des états de choses, elles ont un contenu propositionnel. Par conséquent, les contenus de nos croyances – c'est-à-dire, les propositions tenues pour vraies – peuvent se contredire mutuellement, et quand tel est le cas, nous pouvons dire que certaines de nos croyances sont logiquement incompatibles les unes avec les autres et manquent ainsi de justification. Or, si l'on peut montrer, à l'aide d'arguments, premièrement, qu'il est très probable que les contenus de nos croyances représentent ce qui existe et, deuxièmement, qu'ils ne se contredisent pas logiquement, alors nos croyances sont justifiées. *Les actions* résultent du fait qu'une personne éprouve des *émotions* (passions)

¹ HUME 1983, p. 525-526. «'Tis not contrary to reason for me to chuse my total ruin [...] 'Tis as little contrary to reason to prefer even my own acknowledg'd lesser good to my greater, and have a more ardent affection for the former than the latter.» Hume 1739, p. 416.

qui la poussent à réaliser une fin. Ce sont elles qui motivent les actions, qui indiquent les fins à poursuivre. Or, les émotions ne représentent rien. Bien qu'il puisse y avoir des tensions entre nos différentes émotions, les émotions ne peuvent pas se contredire logiquement. Par conséquent, elles ne peuvent pas correspondre à quelque chose qui les rendrait vraies. Il n'y a donc pas non plus d'arguments qui pourraient démontrer ou au moins suggérer leur vérité. Et puisqu'il n'y a pas – hormis les émotions et les motivations qu'elles engendrent – un autre élément constitutif des actions qui nous permettrait de distinguer entre actions rationnelles et actions irrationnelles, cet argument nous amène à la conclusion qu'il n'y a pas de rationalité pratique ².

Cette conclusion ne nous empêche pas, selon Hume, d'appeler sans absurdité certaines actions irrationnelles, tant que nous n'oublions pas que ce n'est qu'une manière de parler. Car, en réfléchissant correctement, nous découvri- rons toujours que nous nous référons finalement à une croyance fautive ou douteuse, et que cette croyance concerne soit l'existence de l'objet de l'émotion qui nous pousse à agir – par exemple l'existence d'un geste d'invitation ou de menace –, soit la disponibilité et l'effectivité des moyens choisis pour réaliser nos fins – par exemple s'il vaut mieux riposter ou fuir en réaction à une provocation violente. Comme le dit Hume :

Bref, une passion doit s'accompagner de quelque faux jugement pour être déraisonnable ; même alors ce n'est pas, à proprement parler, la passion qui est déraisonnable, c'est le jugement.³

Peu de philosophes contemporains ont voulu se rallier à Hume, mais tous reconnaissent que la défense de la rationalité pratique implique soit une réfutation de la conclusion humienne selon laquelle on ne peut pas rationnellement évaluer nos actions, soit une élaboration et une extension des critères dont l'observation nous permet d'éviter des choix que nous voudrions appeler «irrationnels». Les philosophes qui cherchent à réfuter Hume défendent la position selon laquelle il existe des valeurs morales que l'on ne peut pas mépriser sans violer en même temps la rationalité théorique. Appelons-les des philosophes anti-humiens. Selon eux, l'expression «être justifié» se réfère, dans le contexte de la rationalité pratique, à la valeur de posséder des raisons qui démontrent ou qui suggèrent de manière satisfaisante la *bonté* de nos actions. Les philosophes qui cherchent à développer la position de Hume peuvent être appelés néo-humiens. Selon eux, la position de Hume est correcte, mais incomplète.

Je voudrais présenter par la suite un argument indirect en faveur de la position néo-humienne. Je procéderai en deux étapes. Dans la première partie

² HUME 1983, livre II, partie III, ch. 3.

³ HUME 1983, p. 526. «A passion must be accompany'd with some false judgment, in order to its being unreasonable; and even then 'tis not the passion, properly speaking, which is unreasonable, but the judgment.» Hume 1739, p. 416.

je soutiens que les conceptions anti-humiennes de la rationalité pratique les plus courantes sont inadéquates si elles ne permettent pas de distinguer les standards de la rationalité de ceux de la moralité. Dans la deuxième partie, j'expliquerai d'abord le contenu et l'importance d'un critère supplémentaire de la rationalité pratique, qui est la consistance des préférences. À partir de quelques résultats de recherche en sciences humaines, je montrerai ensuite que les stratégies qui nous rendent capables de satisfaire ce critère supplémentaire nous permettent en même temps d'éviter des fins indésirables. L'adoption de ce critère est une des extensions possibles de la conception de Hume qui nous permet d'élaborer une conception néo-humienne de la rationalité pratique, c'est-à-dire une conception de la rationalité sans moralité.

1. *Deux conceptions anti-humiennes de la rationalité pratique*

La plupart des philosophes anti-humiens défendent leur position en s'inspirant de la philosophie pratique de Kant ou d'Aristote. Je voudrais brièvement caractériser ces deux familles en soulignant leurs éléments pertinents quant à la question du rapport entre rationalité et moralité.

Les *kantiens* pensent que *l'autonomie* – dans le sens de l'autolégislation de la raison – est *une valeur à la fois intellectuelle et pratique*. D'après eux, nos croyances sont justifiées si et seulement si elles le sont d'un *point de vue impersonnel* ; les contingences constitutives d'un point de vue personnel – qu'elles soient d'ordre biologique, culturel ou biographique – ne contribuent en rien à la justification de nos croyances. De manière analogue, nos actions sont justifiées si et seulement si elles le sont d'un point de vue impersonnel, et *l'impératif catégorique* formule cette exigence de justification impersonnelle. Nos raisons démontrent (ou suggèrent de manière satisfaisante) la vérité de ce que nous croyons ou la bonté de ce que nous faisons, à condition que ces raisons soient impersonnelles, c'est-à-dire dépourvues de toute référence à une expérience, une norme ou une capacité qui ne serait pas accessible à tout un chacun. Selon les kantiens, il serait irrationnel de ruiner délibérément sa propre vie ou de promouvoir consciemment son propre malheur, non pas parce que l'émotion qui pousse une personne à agir de la sorte serait «fausse» ou «logiquement incohérente» avec ses autres émotions, mais parce qu'un tel choix ne peut pas résulter d'une motivation autonome ou impersonnelle. Une telle motivation consisterait en effet dans l'observation de la loi morale qui exige, selon Kant, de soutenir l'humanité dans la personne de chacun, en promouvant sa propre perfection et le bonheur d'autrui. L'observation de cette loi ne signifie donc pas seulement que l'on agit rationnellement, c'est grâce à elle que nos actions sont bonnes. Même si les kantiens contemporains⁴ défendent souvent des variantes bien plus sophistiquées de l'argument

⁴ Par exemple, NAGEL, 1970; HAMPTON, 1998; KORSGAARD, 1996.

kantien que mon esquisse ne le laisse penser, ils souscriraient tous à la proposition suivante :

- *Proposition kantienne* : nos actions sont rationnelles si et seulement si elles se laissent justifier d'un point de vue impersonnel qui garantit également la bonté de nos actions.

Tandis que les kantien croient que la rationalité de nos actions constitue la bonté de nos fins, les *aristotéliens* pensent que nos actions sont rationnelles quand elles réalisent des fins qui sont bonnes. Et nos fins sont bonnes dans la mesure où elles contribuent à notre bon fonctionnement en tant qu'êtres humains. Aristote croyait apparemment que l'on reste au niveau de la description empirique en parlant de «bon fonctionnement», comme quand on parle plus spécifiquement de la santé des gens. Les aristotéliens contemporains préfèrent dire que le bon fonctionnement exprime une valeur – à savoir, la valeur d'une vie consacrée à l'exercice des vertus ou des excellences de raisonnement et de caractère. Mener une telle vie requiert la capacité de spécifier cas par cas ce qu'une conduite vertueuse exige dans des circonstances particulières, et cette capacité est, selon Aristote, avant tout un résultat de l'éducation et de l'entraînement. Quelle que soit la description des vertus dont l'exercice assurerait le bon fonctionnement en tant qu'être humain, il serait irrationnel du point de vue aristotélien de ruiner délibérément sa propre vie ou de promouvoir consciemment son propre malheur parce que ces fins sont contraires à l'idée même d'un bon fonctionnement. Une fois que nous connaissons les vertus ou les excellences dont l'exercice assure le bon fonctionnement d'un être humain, nous avons en effet un standard de valeurs qui est indépendant des motivations individuelles et qui nous permet de dire, au moins après une période d'apprentissage, si une action est rationnelle, c'est-à-dire s'il y a de bonnes raisons de croire qu'elle contribue à notre bon fonctionnement ou non. Bien que les aristotéliens contemporains⁵ défendent des théories de la valeur différentes, ils souscriraient tous à la proposition suivante :

- *Proposition aristotélienne* : nos actions ne sont rationnelles que dans la mesure où elles réalisent les fins qui sont objectivement bonnes.

Je ne veux pas évaluer ces deux positions à la lumière de la conception de Hume, mais souligner simplement une conséquence commune aux deux. Les deux excluent en effet la possibilité de distinguer entre *actions rationnelles* et *actions moralement bonnes*. En effet, selon les kantien, une action est moralement bonne si et seulement si elle est entreprise par respect de la loi morale. Or, la loi morale formule l'exigence de justification impersonnelle qui est le critère central de la rationalité pratique. Pour les aristotéliens, une action est rationnelle seulement si elle manifeste les vertus qui expriment ou qui promeuvent le bon fonctionnement des gens en tant qu'êtres humains.

⁵ Par exemple WALLACE, 1978; NUSSBAUM, 1990, ch. 2; MCDOWELL, 1998, ch. 13.

Pourquoi cette impossibilité de distinguer entre actions rationnelles et actions moralement bonnes est-elle problématique ? Rappelons-nous que, pour les anti-humiens, l'expression «être justifiée» semble se référer, dans le contexte de la rationalité pratique, à une *valeur morale*, à savoir la valeur de posséder des raisons qui démontrent ou qui suggèrent de manière satisfaisante la *bonté* de nos actions. Cela présuppose que nous connaissions les conditions qui doivent être satisfaites pour qu'une action soit bonne. Il y a ici une analogie avec la justification théorique. Si nos *croyances* sont justifiées parce que nous possédons des raisons qui démontrent ou qui suggèrent de manière satisfaisante la *vérité* de ce que nous croyons, nous devons connaître les conditions qui doivent être satisfaites pour qu'une proposition soit vraie. Nous avons donc besoin d'une *théorie de la vérité* pour évaluer la rationalité de nos croyances. De manière semblable, les anti-humiens ont besoin d'une *théorie de la valeur* pour évaluer la rationalité de nos actions.

Ces théories ne sont plausibles que si elles respectent certaines contraintes. Une de ces contraintes est, sans doute, que ces théories tiennent compte de nos conceptions naïves de la nature de la vérité et de la valeur telles que le sens commun nous les suggère. Selon la *conception naïve de la vérité*, une proposition est vraie si et seulement si elle *représente correctement la réalité*. Cette conception de la vérité repose sur une intuition réaliste selon laquelle la vérité d'une proposition ne dépend pas de nos attitudes vis-à-vis de cette proposition, – mais de la relation entre cette proposition et la réalité. Selon cette conception naïve, il est possible qu'une proposition soit vraie bien que personne ne la croie, ou que tout le monde partage une croyance fautive. Puisque l'intuition réaliste de la vérité est presque universellement partagée, on peut considérer la conception naïve comme *théorie par défaut*. Cela veut dire qu'en cas de divergences au sujet des conditions qui doivent être satisfaites pour qu'une proposition soit vraie – quand aucune des différentes théories concurrentielles de la vérité ne convainc suffisamment –, la conception naïve est le standard minimal qui s'applique.

Y a-t-il aussi une *conception naïve de la valeur* que le sens commun nous suggère ? Je crois qu'il y en a au moins deux. Certains considèrent comme évident que nos actions sont bonnes parce qu'elles sont *généralement approuvées*, tandis que d'autres considèrent comme évident que nos actions sont bonnes parce qu'elles *expriment ou promeuvent de manière adéquate ce qui nous rend humains*. Les premiers se laissent guider par une *intuition subjective* et pensent qu'il serait absurde de dire qu'une action est bonne, si elle était généralement désapprouvée, car c'est finalement grâce à notre appréciation que les actions sont appelées bonnes ou mauvaises. Les seconds se laissent guider par une *intuition objectiviste* et pensent qu'il est absurde de dire qu'une action est bonne parce qu'elle est généralement approuvée, car la seule raison d'approuver une chose est précisément qu'elle est bonne. Puisque ces deux intuitions sont mutuellement incompatibles, nous pouvons conclure que le sens commun ne suggère pas de théorie par défaut dans le cas de la valeur.

Rappelons-nous maintenant que selon les conceptions kantienne et aristotélicienne de la rationalité pratique, nous ne pouvons pas évaluer la rationalité de nos actions indépendamment de leur bonté. S'il est vrai que le sens commun ne nous fournit aucune théorie par défaut pour évaluer la bonté de nos actions, nous ne pouvons dire qu'une action est rationnelle sans référence à une des théories concurrentielles de la valeur (par exemple une théorie d'inspiration kantienne ou aristotélicienne). Et cela a une conséquence fâcheuse : en cas de divergence au sujet de ce qui rend une action bonne, nous divergeons *a fortiori* sur ce qui la rendrait rationnelle. Tandis que nous pouvons nous référer au sens commun pour évaluer la rationalité des croyances en cas de divergence entre théories concurrentielles de la vérité, le recours à une théorie par défaut n'est pas possible en cas de divergence concernant les conditions qui doivent être satisfaites pour qu'une action soit bonne. Et l'expérience la plus commune nous suggère que ces divergences sont fréquentes.

S'il est vrai que les conceptions kantienne et aristotélicienne requièrent un rattachement de la rationalité pratique à la moralité, elles ne sont apparemment pas des alternatives attrayantes à la conception de Hume. Premièrement, elles sont vulnérables à *l'objection de futilité* ; car les théories morales (au moins sous la forme d'une théorie normative de la valeur) sont essentiellement contestables⁶. Deuxièmement, elles sont vulnérables à *l'objection d'inutilité* ; car la théorie de la rationalité pratique qu'elles proposent ne peut être utilisée dans l'étude empirique du comportement humain. Je voudrais seulement ici expliquer brièvement la deuxième objection.

Dans les sciences humaines, on a besoin d'une conception de la rationalité pratique qui soit axiologiquement neutre, afin de pouvoir expliquer le comportement humain sans l'évaluer – et c'est précisément ce que les conceptions kantienne et aristotélicienne excluent. Or, si les critères de rationalité pratique proposés par les philosophes ne sont pas pertinents pour la recherche en sciences humaines, ils sont simplement inadéquats. Un cas analogue se présente d'ailleurs si l'on considère la rationalité théorique. Les philosophes qui prétendent qu'il ne peut y avoir de croyance rationnelle qui contredise un dogme religieux ou culturel cherchent à établir des critères de rationalité théorique qui ne sont pas pertinents pour la recherche scientifique et pour cette seule raison simplement inadéquats. Supposons que les sciences humaines en viennent à embrasser une conception de la rationalité pratique qui est rattachée à une théorie morale. Dans ce cas, elles devraient abandonner le principe de neutralité axiologique (*Wertfreiheitsprinzip*). Puisque l'abandon de ce principe rendrait l'étude scientifique du comportement humain impossible, nous avons besoin d'une conception de la rationalité pratique qui ne s'expose pas à cette objection. Je voudrais maintenant défendre la thèse selon laquelle c'est précisément le cas pour les conceptions néo-humiennes de la rationalité pratique.

⁶ Cf. BAIER 1985, part 2.

2. *La consistance des préférences*

Les philosophes néo-humiens cherchent à compléter plutôt qu'à réfuter Hume⁷. Comme Hume, ils maintiennent que la raison ne nous indique pas quelles sont les bonnes fins ; ce sont nos émotions qui nous présentent les fins éligibles de nos actions comme plus ou moins désirables. Cependant, les néo-humiens insistent sur le fait que la raison nous rend capables *d'éviter le choix de fins indésirables* bien plus souvent que Hume ne voulait l'admettre. Selon eux, ce dernier aurait négligé le fait que nos fins peuvent être mutuellement *consistantes* ou *inconsistantes*, dès que nous devons choisir parmi une pluralité d'options plus ou moins attractives. En effet, nous ne pouvons pas choisir parmi ces options sans classer, selon un ordre de priorité, les états de choses qu'elles sont censées produire.

Je défendrai ici une variante de la position néo-humienne qui emprunte un certain nombre d'éléments à la théorie de la décision⁸. Au lieu de parler de fins, nous pouvons donc parler de préférences pour des états de choses, en supposant que leur réalisation peut être causée, ou rendue plus probable, par le choix de certaines options plutôt que d'autres. Pour qu'un ordre de préférences puisse garantir que nous choisissons les options que nous valorisons le plus, cet ordre doit assurer la consistance des préférences. D'après ce point de départ, la première tâche des philosophes néo-humiens est d'indiquer les critères de consistance entre préférences. Quels sont donc ces critères ?

La consistance concerne la compatibilité logique de relations de préférences (ou des propositions qui les expriment). Ainsi, si par exemple quelqu'un préfère

- (A) «la conservation de sa santé»
à
(B) «la gloire pour ses exploits sportifs»,
et s'il préfère également
(B) «la gloire pour ses exploits sportifs»
à
(C) «les plaisirs sensuels qu'une drogue lui procure»,
alors cette personne a des préférences inconsistantes si elle préfère
(C) «ses plaisirs sensuels ...»
à
(A) «la conservation de sa santé»,

⁷ Cf., par exemple, BAIER, 1985, part 2; 1997; BLACKBURN, 1998; WILLIAMS, 1990; mais aussi SUGDEN, 1998 et HARDIN, 1988.

⁸ Une bonne introduction à la théorie de la décision pour philosophes est RESNICK, 1987.

car cette dernière préférence viole la *transitivité* des préférences selon laquelle

$$A > B \ \& \ B > C \Rightarrow A > C^9$$

La satisfaction du critère de transitivité fait partie de ce que l'on entend par consistance des préférences.

Pourquoi est-il mal de violer ce critère ? Quand toutes les options (A , B , C) sont disponibles en même temps, la violation du principe de transitivité entraîne une indécision qui peut motiver une procrastination semblable à celle dont souffre l'âne de Buridan ; quand les options sont disponibles seulement en paires (AB ; BC ; AC), elle peut conduire, au fil du temps, à l'indigence¹⁰. Ainsi, quelles que soient les préférences d'une personne, si elle cherche à les satisfaire au mieux, elle devrait veiller à la transitivité de ses préférences.

En plus de la transitivité, deux autres critères au moins doivent être satisfaits pour que les préférences d'une personne soient mutuellement consistantes, dont l'un concerne la neutralité du temps et l'autre l'indifférence au poids de la probabilité pour la valeur d'une fin.

Le critère de *neutralité temporelle* est satisfait si l'ordre de nos préférences ne subit pas de changement qui soit dû uniquement au passage du temps. Supposons que je doive choisir entre

(A) «la représentation d'une pièce de théâtre demain soir»

et

(B) «la représentation d'un opéra après-demain soir».

Ne pouvant pas aller aux deux représentations, je préfère l'opéra ($B > A$). Si, le lendemain, le seul rapprochement temporel de la représentation de la pièce de théâtre me conduit à renverser mes préférences de telle sorte que je préfère maintenant la représentation de la pièce de théâtre ce soir à la représentation de l'opéra demain soir ($A > B$), je viole le principe de neutralité du temps, qui dit que

$$\text{Temps 1 } (B > A) \Rightarrow \text{Temps 2...n } (B > A).$$

Pourquoi est-il mauvais de violer ce critère ? Sans observer ce principe, une personne ne pourrait guère poursuivre des plans, même si elle reconnaissait que la poursuite des plans avancerait son bien-être. Et si nous admettons que nous devons souvent incorporer nos préférences dans la structure complexe de plans afin de pouvoir les réaliser, nous avons intérêt à veiller à la neutralité temporelle de nos préférences.

Le critère d'*indifférence au poids de la probabilité* est satisfait si l'ordre de nos préférences ne subit pas de changements uniquement parce que nous aimons être plus (ou moins) confiants d'atteindre une fin. Si, par exemple, je préfère

⁹ J'utiliserai «>» pour «est préféré à».

¹⁰ Cf. HAUSMAN & McPHERSON, 1996, p. 28.

(A) «obtenir 10 francs avec une probabilité de 50%»
à
(B) «obtenir 9 francs avec une probabilité de 50%»,
je ne peux pas, de manière consistante, préférer
(B*) «obtenir 4.50 francs avec certitude»
à
(A) «obtenir 10 francs avec une probabilité de 50%»,
car cela violerait un critère de consistance disant que

$$[A > B \ \& \ \neg (B > B^*) \ \& \ \neg (B < B^*)] \Rightarrow A > B^*.$$

Pourquoi est-il mauvais de violer ce critère ? Simplement parce que sa violation a pour conséquence que l'on ne réalise pas sa fin préférée : ne pas être indifférent au poids de la probabilité revient à programmer des regrets.

Certains théoriciens de la décision pensent que la rationalité pratique requiert, en plus, que les préférences d'une personne soient *complètes*, c'est-à-dire que la personne soit capable de comparer toutes les options possibles en termes de préférences, et que les préférences d'une personne soient *continues*, ce qui implique, par exemple, que cette personne n'ait pas de relations de préférence lexicographiques. Il n'est pas nécessaire de discuter la pertinence de ces critères supplémentaires dans le présent contexte. Je crois qu'il suffit de compléter la position de Hume en ajoutant les trois critères de consistance que je viens d'exposer. Bien entendu, cela n'est que le premier pas dans l'élaboration de la position néo-humienne que je souhaite défendre. Il s'agit maintenant de répondre aux deux questions suivantes : (1) Comment pouvons-nous satisfaire ces critères de consistance ? (2) Comment la satisfaction de ces critères nous permet-elle d'éviter des fins indésirables ?

Je voudrais esquisser une réponse à ces questions à partir de quelques considérations empiriques. Ainsi, bien que mon point de départ diverge de celui qu'occupent la plupart des philosophes néo-humiens contemporains, j'arriverai à des résultats semblables aux leurs. Je voudrais montrer que les compétences qui permettent aux gens d'agir sur la base de préférences consistantes leur permettent également de résister au choix de fins indésirables.

Examinons d'abord la première de ces questions. Puisqu'il semble facile de comprendre l'avantage d'avoir des préférences consistantes, on devrait pouvoir admettre que les gens observent généralement ce critère de rationalité pratique. Or, certaines expériences nous apprennent que les gens violent fréquemment ce critère. Ces expériences sont problématiques parce qu'elles risquent de *saper le rôle normatif de la rationalité* même si son rôle critique est bien défini. En effet, si nous devons agir rationnellement, il faut que nous puissions le faire.

Considérons pour commencer le critère de neutralité temporelle. Je suppose qu'il n'est pas inconsistant pour les mortels de préférer avoir un bien plus tôt que plus tard. Et si l'on a le choix entre deux biens de valeurs différentes, et

que le plus petit est disponible quelque temps avant le plus grand, il n'est pas non plus inconsistant d'attendre pour avoir le plus grand bien. Dans ce deuxième cas, il serait cependant inconsistant de changer ses préférences en faveur du plus petit bien simplement parce que le moment approche où l'on pourrait l'avoir. Pour illustrer ce cas, imaginez que l'on vous présente deux situations de choix. Dans la première situation, vous devez dire si vous préférez obtenir

(A) 100 francs sur-le-champ

ou

(B) 200 dans deux ans.

Dans la deuxième situation vous devez dire si vous préférez obtenir

(C) 100 francs dans 6 ans

ou

(D) 200 dans 8 ans.

Si dans la première situation vous choisissez (A) et dans la deuxième (D), vous choisissez comme la majorité des gens – en violant le critère de neutralité temporelle des préférences. En effet, dans les deux situations, les sommes ainsi que l'écart temporel entre leur disponibilité sont identiques. C'est uniquement le rapprochement ou l'éloignement temporel qui semble conduire au renversement des préférences.

On sait que la majorité des gens optent pour (A) et (D), parce que c'est le résultat d'une des nombreuses études empiriques qui ont été conduites pour savoir si les gens ont des préférences temporelles consistantes. D'autres résultats confirment l'hypothèse que le critère de neutralité temporelle des préférences est souvent violé, par exemple dans le choix entre une période courte ou une période longue d'absence d'un bruit dérangeant, dans le choix des femmes d'accoucher sous péridurale ou non, et dans le choix des enfants de 4 ans entre deux collations d'attractivité différente. Même les animaux semblent préférer une proie immédiatement disponible aux alternatives supérieures mais différées ou une petite quantité de nourriture tout de suite à une plus grande quantité plus tard ¹¹.

Ces études ne cherchent pas à nier l'avantage d'avoir des préférences consistantes à travers le temps, mais elles suggèrent que cet avantage n'est pas gratuit. Pour en bénéficier, il faut acquérir des compétences qui nous permettent de rester neutres face à la proximité ou à la distance temporelles, afin de maintenir nos préférences à travers le temps. En langage de tous les jours, cela veut dire qu'il faut apprendre à se contrôler soi-même. Ce contrôle requiert la familiarité avec un ensemble de stratégies d'engagement à l'avance (*precommitment*), dont on peut distinguer au moins quatre types ¹². Dans

¹¹ Cf. A INSLIE, 1992, ch. 3; LOEWENSTEIN & ELSTER, 1992.

¹² Cf. A INSLIE & HASLAM, 1992.

certaines situations, il faut savoir *se priver d'une partie de sa liberté* en donnant aux autres un certain pouvoir sur nous, comme Ulysse ordonnant à ses compagnons, à l'approche des sirènes, de l'attacher au mât et d'ignorer ses protestations. Dans d'autres circonstances, il faut savoir *se priver de certaines informations*, et, par exemple, ne pas regarder le programme de télévision pour la soirée que l'on préfère passer avec des amis. Parfois, il faut savoir *préparer ses réactions émotionnelles*, comme le végétarien par principe qui se fait une habitude d'évoquer des images répulsives d'usines d'animaux et d'abattoirs pour ne pas être tenté de manger de la viande. Et de temps en temps, il faut savoir *se tenir strictement à des règles personnelles*, par exemple à la règle de ne jamais fumer une seule cigarette quand on a réussi à se libérer de la dépendance. La recherche au sujet de ces stratégies devrait nous informer de plus en plus exactement sur la manière dont elles nous permettent de satisfaire le critère de neutralité temporelle.

Considérons maintenant le critère d'indifférence au poids de la probabilité. Je suppose qu'il n'est pas inconsistant pour les mortels d'être légèrement hostiles au risque. Dans une situation de choix, ils pourront ainsi préférer obtenir un petit bien avec une grande probabilité à obtenir un grand bien avec une petite probabilité, même si la valeur escomptée de la seconde option plus risquée est supérieure à celle de la première. Mais tant que l'aversion au risque est limitée, il n'est pas inconsistant de préférer l'option plus risquée quand sa valeur escomptée dépasse de loin celle de l'option moins risquée. Cependant, il est inconsistant de préférer d'abord l'option moins risquée avec une valeur escomptée plus petite, et ensuite l'option avec une valeur escomptée plus grande mais plus risquée. Pour illustrer le cas, imaginez que l'on vous présente deux situations de choix. D'abord vous devez choisir entre deux bons, le bon A et le bon B, qui vous permettent de tirer aveuglément une parmi dix boules numérotées de 1 à 10. Voici les opportunités liées à ces bons :

Bon (A) : Vous gagnez 10 francs si vous tirez n'importe quelle boule à l'exception de la boule numéro 1

Bon (B) : Vous gagnez 100 francs seulement si vous tirez la boule numéro 1

Ensuite, vous devriez dire quel serait le prix minimum auquel vous seriez prêt à vendre chacun des deux bons, si vous les possédiez. Si, dans la première situation vous choisissez (A) et dans la deuxième situation vous demandez un prix plus élevé pour (B), vous agissez comme la majorité des gens – en violant le critère d'indifférence au poids de la probabilité. Si vous préférez d'abord (A) à (B), vous ne pouvez pas ensuite valoriser (B) plus que (A) sans modifier votre attitude face au poids de la probabilité d'un cas à l'autre, car la valeur escomptée des options reste inchangée.

De nombreuses études en psychologie et en économie expérimentale suggèrent que les gens renversent fréquemment leurs préférences en violation du

critère d'indifférence au poids de la probabilité¹³. La plupart d'entre elles suggèrent que ces violations s'expliquent par la présence de certains mécanismes psychologiques tels que la surévaluation de la certitude comparée au risque, ou le changement de l'attitude à l'égard du risque selon le type de conséquences, ou la représentation des conséquences en termes de gain ou de perte (*framing*), etc¹⁴.

Comment peut-on contrôler ces tendances ? On en sait peu. La familiarité avec des stratégies d'engagement à l'avance (*precommitment*) semble être utile dans les cas où il s'agit simplement de résister à la tentation d'insouciance. On peut, par exemple, conclure des contrats d'assurance dans lesquels on s'engage à mettre sa ceinture de sécurité, ou à ne pas fumer, sous peine de conséquences financières graves. Cette privation de liberté peut aider les gens à avoir des préférences consistantes concernant la valeur de leur intégrité physique. Dans d'autres cas, le problème semble dû au fait que différentes procédures d'évaluation sont utilisées d'une situation de choix à l'autre et que ces procédures ne fournissent pas le même type d'information. Si l'on vous demande de dire si vous préférez le

Bon (A) { 10 francs avec une chance de 9/10 }

ou le

Bon (B) { 100 francs avec une chance de 1/10 },

vous devez *établir un ordre de priorité*. Si l'on vous demande d'indiquer le prix minimal acceptable pour le bon (A) et pour le bon (B), vous devez *corrélér chaque bon avec une somme d'argent*. Il n'est pas sûr que vous soyez attentif aux mêmes aspects dans les deux situations de choix. Je suppose que l'on peut acquérir la perspicacité qui nous permet de veiller à la compatibilité entre différentes procédures d'évaluation. Et il y a certainement d'autres compétences qui nous rendent plus aptes à respecter le critère d'indifférence au poids de la probabilité. Il suffit ici de constater que le respect de ce critère est coûteux.

Puisqu'on peut imaginer facilement comment les renversements de préférences décrits peuvent conduire à la violation du critère de transitivité, je ne veux pas considérer ce troisième critère à part. Les exemples discutés suggèrent (a) que nous ne pouvons satisfaire les critères de consistance des préférences sans effort, et (b) que cet effort consiste dans le contrôle de soi-même et dans l'exercice d'une certaine perspicacité. Maintenant, il s'agit de voir comment ces *compétences décisionnelles*, comme je vais les appeler, nous permettent d'éviter des fins indésirables. C'était la deuxième question.

Pour y répondre, je voudrais commencer par trois observations peu controversées à propos de l'acquisition de ces compétences décisionnelles. La première est que nous les acquérons dans la pratique, d'abord en observant et en imitant autrui et en nous soumettant à sa critique, et ensuite en expérimentant

¹³ TVERSKY & KAHNEMAN, 1986; THALER 1992, ch. 7.

¹⁴ Cf. SUTHERLAND, 1992.

nous-mêmes. Ce type d'apprentissage peut être appelé *learning by doing*¹⁵. Deuxièmement, on peut observer que l'acquisition de ces compétences nous rend capables d'éviter un certain nombre de pièges dans nos prises de décisions, mais sans nous donner en même temps les moyens d'expliquer exactement comment et pourquoi. Le savoir acquis est avant tout un *savoir-faire tacite*¹⁶. La troisième observation est que l'acquisition de ces compétences débouche sur *un certain degré de maîtrise*. Le savoir acquis reste partiel.

Si ces observations sont correctes, elles suggèrent que l'acquisition des compétences décisionnelles ne dépend pas seulement des activités d'un individu, mais également du fait que les gens qui l'entourent partagent certains *standards* que l'emploi de ces compétences devrait normalement satisfaire. Comme les compétences linguistiques supposent des règles de grammaire partagées par les membres d'un groupe linguistique, les compétences décisionnelles supposeraient un *standard partagé de prudence*. Une moindre satisfaction du standard partagé de grammaire indique – *ceteris paribus* – une compétence linguistique moindre (par exemple dans le cas des petits enfants). Une moindre satisfaction du standard partagé de prudence indiquerait – *ceteris paribus* – une compétence décisionnelle moindre (par exemple dans le cas d'un apprenti). L'existence de ces standards permet de répondre à la question de savoir si quelqu'un possède les compétences requises par un type de choix, et le cas échéant, si elles sont suffisamment développées pour choisir rationnellement dans un cas particulier. Les deux exemples suivants devraient illustrer cette idée.

Imaginons d'abord un gourmand qui veut garder la ligne. Il peut se demander s'il dispose de stratégies d'engagement à l'avance (*precommitment*) pour accompagner régulièrement ses collègues de travail à la cafétéria sans succomber à la tentation de manger. Si l'observation de ses réactions dans d'autres situations comparables ne lui permet pas de répondre, il peut demander l'avis de quelqu'un qui maîtrise davantage ces stratégies. Il est important de voir qu'il s'agit ici uniquement de savoir si le gourmand possède les compétences décisionnelles requises – ce qui veut dire, dans le cas présent, s'il est capable de se contrôler – et non pas si ses préférences satisfont le critère de neutralité temporelle. Nous jugeons de ses compétences décisionnelles en vertu du standard partagé de prudence. Et si nous constatons qu'il satisfait le standard partagé de prudence, nous avons alors de bonnes raisons pour *présumer* qu'il satisfait également le critère de neutralité temporelle.

Imaginons maintenant une jeune mère qui, après quelques années sans activité en dehors du foyer, décide soit de reprendre un travail, soit d'entreprendre des études. Si elle opte pour le travail, il est très probable qu'elle ne trouvera qu'une occupation peu intéressante avec un salaire modeste. Si elle réussit ses études, elle aura des chances de trouver un emploi intéressant et

¹⁵ Cf. OAKESHOTT, 1991.

¹⁶ Cf. POLANYI, 1989.

mieux payé. Elle peut se demander si elle dispose de la perspicacité requise pour envisager plusieurs procédures d'évaluation afin de choisir ce qu'elle désire le plus. Si l'observation de ses réactions dans d'autres situations comparables ne lui permet pas de répondre, elle peut demander l'avis de quelqu'un qui possède davantage de perspicacité. Comme dans l'exemple précédent, il ne s'agit pas d'examiner si les préférences de la jeune mère satisfont le critère d'indifférence au poids de la probabilité, mais seulement de savoir si la jeune mère possède (ou peut s'approprier) les compétences décisionnelles requises, c'est-à-dire, dans le cas présent, la perspicacité suffisante pour choisir ce qu'elle veut. Comme dans l'exemple précédent, nous jugeons de ses compétences décisionnelles en vertu du standard partagé de prudence. Et nous pouvons *présumer* que si c'est le cas, son choix satisfera également le critère d'indifférence au poids de la probabilité. En effet, si elle est perspicace, nous avons de bonnes raisons pour pronostiquer qu'elle satisfera ce critère.

Pourquoi ces compétences décisionnelles comme les stratégies d'engagement à l'avance (*precommitment*) et la perspicacité nous permettent-elles d'éviter des fins indésirables ? La réponse est que ces compétences fonctionnent comme des filtres. Si le choix d'une fin ne peut pas se faire en satisfaisant les standards de contrôle de soi et de perspicacité propres à la situation donnée, cette fin ne passe pas le filtre. Et cela signifie que cette fin est *prima facie* indésirable. La raison en est que l'observation de ces standards nous donne la meilleure assurance que nous agissons à partir de préférences consistantes.

Il y a certainement d'autres compétences décisionnelles dont l'exercice contribue à assurer que nos préférences soient mutuellement consistantes, mais je ne saurais pas en fournir un inventaire. Notons, cependant, que nous semblons nous servir d'outils mentaux similaires pour satisfaire les critères de rationalité de nos croyances. C'est du moins ce qu'un nombre croissant de psychologues suggèrent¹⁷. Selon eux, nous déployons un certain nombre de techniques heuristiques lorsque nous recherchons et traitons des informations. Ces techniques heuristiques nous permettent d'éviter certaines distorsions dans le processus de formation et d'organisation de nos croyances et contribuent ainsi à assurer leur rationalité.

Voici un exemple. Une défaillance cognitive répandue résulte du penchant à préférer systématiquement des raisons basées sur une plus grande quantité d'informations. Or, ces raisons ne sont pas toujours les meilleures¹⁸. À la bourse, les entrepreneurs ignorant délibérément une grande quantité d'informations disponibles semblent pouvoir faire de meilleures prédictions concernant l'attractivité des titres¹⁹. Ils bénéficieront donc d'une heuristique qui leur indique quand il faut arrêter la recherche d'information, afin d'établir une telle prédiction. Notons qu'une telle heuristique n'aura pas la forme d'un algorithme

¹⁷ Cf. GIGERENZER & AL., 1999; KLEIN, 1999.

¹⁸ Cf. HEINER, 1983.

¹⁹ Cf. GIGERENZER & AL., ch. 3.

pour calculer la valeur ajoutée de chaque unité supplémentaire d'information, car cette méthode ne diminue guère la quantité d'informations requise pour faire des prédictions. Elle consisterait plutôt dans l'application de certaines stratégies de simplification que l'on apprend par voie d'imitation et d'expérimentation, comme par exemple la recherche d'une seule raison saillante²⁰. En s'appropriant une telle heuristique, on saisit également le standard qui doit être satisfait par celui qui maîtrise son application. Ainsi, la maîtrise de l'heuristique devrait assurer au mieux la rationalité des croyances concernant nos opportunités et nos moyens.

Revenons maintenant à nos exemples intuitifs d'irrationalité pratique. Pourquoi est-il irrationnel de sacrifier ma santé à la poursuite de plaisirs sensuels ? Parce que quelqu'un qui dispose de stratégies d'engagement à l'avance (*precommitment*) préfère réduire sa poursuite de plaisirs sensuels pour conserver sa santé – à moins qu'il ne dévalorise le futur de manière excessive, ce qui est une position presque impossible à tenir de manière consistante, c'est-à-dire, sans provoquer systématiquement des regrets. La difficulté des adolescents à vivre au jour le jour sans souffrir de leur propre inconstance illustre cette observation. Pourquoi est-il irrationnel de renoncer au développement de ses talents et de sa personnalité pour la poursuite du prestige et du statut social ? Parce que quelqu'un de perspicace préfère les biens intrinsèques aux biens positionnels et considère cet échange comme une perte²¹ – à moins qu'il ne dévalorise son autonomie personnelle, par exemple, par un conformisme excessif. Encore qu'il me semble extrêmement difficile d'adopter de manière consistante la position d'un conformiste excessif : elle semble requérir un effort continu de reniement de ses besoins naturels, et la recherche empirique suggère que le conformisme devient ainsi une source importante de troubles psychologiques²².

Si les néo-humiens peuvent reconnaître que ce sont en effet des exemples d'irrationalité pratique, qu'est-ce qui est alors problématique avec notre manière de raisonner dans ces deux exemples ? La réponse est que ces exemples ne passent pas le test qui nous indique si nous préférons effectivement ce que nous croyons préférer. Ce test n'a aucun autre but que de constater si ce que nous croyons préférer dans un cas est consistant avec ce que nous croyons préférer dans un autre cas. Si le résultat est négatif, nous avons une bonne raison de croire que nous ne préférons pas ce que nous croyons préférer. Mais le test ne dit rien sur la bonté ou la désidérabilité de nos fins.

²⁰ Cf. GIGERENZER & AL., part 3.

²¹ Cf. SCITOVSKI 1992.

²² Cf. RYAN & DECI 2000.

Bibliographie

- AINSLIE, George (1992), *Picoeconomics : The Strategic Interaction of Successive Motivational States within the Person*, Cambridge, Cambridge University Press.
- AINSLIE, George & Nick HASLAM (1992), «Hyperbolic Discounting», in LOEWENSTEIN & ELSTER (1992), p. 57-92.
- BAIER, Annette (1985), *Postures of the Mind : Essays on Mind and Morals*, London, Methuen & Co.
- BLACKBURN, Simon (1998), *Ruling Passions : A Theory of Practical Reasoning*, Oxford, Clarendon Press.
- GIGERENZER, Gerd, Peter M. TODD AND THE ABC RESEARCH GROUP (1999), *Simple Heuristics that Make Us Smart*, Oxford, Oxford University Press.
- HAMPTON, Jean E. (1998), *The Authority of Reason*, Cambridge, Cambridge University Press.
- HARDIN, Russell (1988), *Morality within the Limits of Reason*, Chicago, University of Chicago Press.
- HEINER, Ronald A. (1983), «The Origin of Predictable Behaviour», *American Economic Review* 73/4, p. 560-595.
- HUME, David (1739), *A Treatise of Human Nature*, éd. L.A. Selby-Bigge, Oxford 1983, Clarendon Press.
- KLEIN, Gary (1999), *Sources of Power : How People Make Decisions*, Cambridge, MIT Press.
- KORSGAARD, Christine (1996), *Creating the Kingdom of Ends*, Cambridge, Cambridge University Press.
- LOEWENSTEIN, George & Jon ELSTER (eds.) (1992), *Choice over Time*, New York, Russell Sage Foundation.
- MCDOWELL, John (1998), *Mind, Value and Reality*, Cambridge, Harvard University Press.
- NAGEL, Thomas (1970), *The Possibility of Altruism*, Princeton, Princeton University Press.
- NUSSBAUM, Martha (1990), «The Discernment of Perception : An Aristotelian Conception of Private and Public Rationality», in *Love's Knowledge : Essays on Philosophy and Literature*, New York, Oxford University Press, p. 54-105.
- OAKESHOTT, Michael J. (1991), *Rationalism in Politics and Other Essays*, new and expanded edition, Indianapolis, Liberty Press.
- POLANYI, Michael (1989), *La logique de la liberté*, Paris, P.U.F.
- RESNICK, Michael D. (1987), *Choices : An Introduction to Decision Theory*, Minneapolis, University of Minnesota Press.
- RYAN, Richard M. & Edward L. DECI (2000), «Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being», *American Psychologist* 55/1, p. 68-78.
- SCITOVSKY, Tibor (1992), *The Joyless Economy : The Psychology of Human Satisfaction*, revised edition, New York, Oxford University Press.

ASSOCIATION DES SOCIÉTÉS DE PHILOSOPHIE
DE LANGUE FRANÇAISE (ASPLF)

SOCIÉTÉ AZURÉENNE DE PHILOSOPHIE
(CÔTE D'AZUR – FRANCE)

CENTRE DE RECHERCHES
D'HISTOIRE DES IDÉES (CRHI)
UMR 6045 CNRS

DÉPARTEMENT DE PHILOSOPHIE
FACULTÉ DES LETTRES, ARTS
ET SCIENCES HUMAINES
UNIVERSITÉ DE NICE – SOPHIA ANTIPOLIS

Le XXIX^e Congrès de l'Association des Sociétés de philosophie
de langue française (ASPLF) aura lieu à Nice (France)

du mardi 27 août au dimanche 1^{er} septembre 2002
(le dimanche 1^{er} septembre sera consacré à une excursion)

Le thème du congrès est :

Avenir de la raison, devenir des rationalités

Les communications se disposeront en 3 sections :

1^{ère} section : *Les figures de la raison aujourd'hui*

2^{ème} section : *Les rationalités scientifiques*

3^{ème} section : *Unité et pluralité de la raison pratique*

Vous êtes invité(e) à participer à ce congrès.

Votre proposition de communication doit nous parvenir avant le 10
septembre 2001, accompagnée d'un résumé dactylographié n'excé-
dant pas 25 lignes de 60 signes. Veuillez adresser votre envoi à :
Congrès ASPLF, Département de philosophie, Université de Nice,
98, bd E. Herriot, B.P. 3209, F - 06204 Nice – Cedex 3, E-mail :
leboulan@unice.fr