

**Zeitschrift:** Studia philosophica : Schweizerische Zeitschrift für Philosophie =  
Revue suisse de philosophie = Rivista svizzera della filosofia = Swiss  
journal of philosophy

**Herausgeber:** Schweizerische Philosophische Gesellschaft

**Band:** 59 (2000)

**Artikel:** Deux maximes de rationalité émotive

**Autor:** De Sousa, Ronald

**DOI:** <https://doi.org/10.5169/seals-882925>

### **Nutzungsbedingungen**

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. 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. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. [Mehr erfahren](#)

### **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. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. [En savoir plus](#)

### **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. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. [Find out more](#)

**Download PDF:** 02.04.2026

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

## Emotion und Vernunft / Emotion et rationalité

---

RONALD DE SOUSA

### Deux maximes de rationalité émotive

*It is now widely acknowledged that emotions make an important contribution to rationality, and may themselves be assessed as more or less rational. But neither instrumental rationality nor epistemic rationality suffice to account for such assessments. Emotions are reducible neither to desires nor to beliefs. This is brought out by looking at two plausible maxims of emotional rationality. Both concern temporality: the first holds that certain objects of desire are more adequately envisaged under a given temporal aspect; the second enjoins a systematic proportionality of pleasures of anticipation to anticipated pleasures. Although neither can be reduced to classical principles of rationality, a broadly biological point of view sheds some light on why each one seems reasonable.*

#### 1. Emotions et valeurs

Les philosophes ont longtemps caressé l'espoir de montrer que la morale n'est autre que la rationalité bien comprise. Par ailleurs, on a voulu voir une opposition irréductible entre la raison et les émotions. Mises ensemble, ces deux doctrines relèguent les émotions à un rôle auxiliaire, sinon positivement néfaste, dans la quête d'une conduite morale : les émotions corrompent nos meilleures intentions et nos projets rationnels, et nous poussent à poursuivre des buts à courte échéance qui nuisent à nos propres intérêts, si étroitement égoïstes soient-ils.

Pourtant les émotions sont de toute évidence au centre de notre vie mentale, et le mépris des théories éthiques envers elles est paradoxal. En effet, la notion même de morale n'aurait aucun sens si les émotions n'existaient pas. C'est par ce que nous ressentons que nos actes (ou ceux des autres) et leurs conséquences nous intéressent. Cette observation banale nourrit l'espoir de réhabiliter les émotions et de leur reconnaître un rôle dans la rationalité, aussi bien que dans le fondement de la morale.

Dire que l'on accorde un rôle aux émotions dans la morale, ce n'est pas préciser en quoi ce rôle consiste. Plusieurs positions possibles s'étalent sur une échelle qui marque différents degrés d'objectivité. A gauche (si l'on peut dire) il y a l'*émotivisme*, qui incarne un subjectivisme extrême. Selon cette doctrine, toute opinion morale n'est que l'expression d'une émotion.<sup>1</sup> Si l'on pense corriger un jugement moral erroné, on ne peut qu'exprimer une autre attitude émotive. Tout à droite, on trouve un objectivisme platonicien qui verrait dans toute émotion une intuition qui nous permet d'accéder à un univers moral objectif. Ces positions extrêmes sont sans nul doute indéfendables. Elles laissent cependant assez d'espace entre elles pour que nous puissions distinguer au moins deux autres positions plus vraisemblables. A côté du subjectivisme, on peut caractériser le *naturalisme*; plus près de l'objectivisme, on peut camper l'*axiologisme*. C'est cette dernière position qui formera la toile de fond du présent exposé. Mais esquissons d'abord ce que j'entends par *naturalisme*.

Le *naturalisme* se démarque par le refus de la distinction entre le normatif et l'état de fait. Certains états d'esprit (à commencer par le plaisir et la douleur) sont *intrinsèquement* désirables ou indésirables. Par conséquent la valeur qu'on leur attribue se passe de justification. L'argument de Moore,<sup>2</sup> comme quoi on peut toujours demander si le plaisir est une bonne chose, se trouve congédié sans autre, qualifié de simple pétition de principe: on peut bien la poser, dira le naturaliste, mais la réponse affirmative, à l'image du *cogito* cartésien, sera toujours plus certaine que tout ce qu'on pourrait avancer dans le but de la soutenir ou de l'infirmier. Il faut donc renoncer à lui chercher un fondement.

Comme beaucoup de renoncements, cette perspective austère n'est pas sans avantages. Elle permet aux différences de valeur de se dessiner clairement en noir et blanc. Mais, par là même, elle manque de couleur. En vertu de l'attention qu'il prête à la motivation, et donc à la question du comportement, le naturalisme entraîne une sorte d'aplatissement des valeurs, et la notion d'émotion morale en demeure curieusement appauvrie. On peut s'en rendre compte lorsqu'on la compare avec les notions de *pensée morale* ou d'*acte moral*. En effet on ne saurait exiger que les pensées morales ou les actes moraux se rangent facilement dans un petit nombre de catégories fixes. Les émotions seraient-elles intrin-

1 La défense la plus sérieuse de ce point de vue demeure celle qu'entreprit Stevenson; cf. *Ethics and Language*.

2 Cf. MOORE, *Principia Ethica*.

sèquement simples ? Une telle conclusion ne ferait guère justice à la complexité émotionnelle d'une œuvre de Tolstoï ou de Proust. Une notion de valeur pluridimensionnelle s'impose.<sup>3</sup> Mais la multiplicité des échelles de valeurs, à son tour, rend celles-là impropres à la détermination directe du comportement. On peut dire, pour prolonger la métaphore de la couleur, que celle-ci élargit la gamme de notre expérience, mais sacrifie la gradation univoque qui va du noir au blanc.

C'est ce qui nous amène à l'*hypothèse axiologique*. Pour la schématiser de manière très sommaire, cette hypothèse repose sur les considérations suivantes :

- (i) Les émotions sont souvent motivantes, mais elles ne sont pas pour autant réductibles à des désirs. Le désir implique une tendance à l'action. Il exige une adéquation du monde à la pensée (« world-to-mind direction of fit »), ce qui n'est pas nécessairement le cas pour les émotions.
- (ii) Par ailleurs, les émotions sont aussi souvent cognitives, mais elles ne sont pas pour autant réductibles à des croyances, bien qu'elles partagent avec celles-ci l'impératif d'une adéquation de la pensée au monde.
- (iii) Ainsi les émotions doivent plutôt être considérées comme constituant une sorte de troisième instance, qui s'apparente au désir par son rapport à la motivation, et à la croyance dans la mesure où elle prétend porter un jugement sur le monde objectif.

L'hypothèse axiologique est naturaliste dans un sens large, mais elle n'est pas condamnée à accepter l'aplatissement des valeurs, car elle rejette l'idée qu'il soit possible de ranger tout ce qui s'inscrit dans le domaine de la valeur le long d'une seule échelle. Pour le déontologisme, toute alternative s'inscrit dans la dimension qui seule débouche sur le choix d'une action. Or l'axiologie garde ses distances envers toute conclusion strictement pratique. Elle peut donc envisager un nombre indéfini de nuances et de dimensions incommensurables. Le point de vue axiologique considère certaines émotions comme des *perceptions de valeurs*,<sup>4</sup> qui nous apprennent quelque chose sur le monde. Les émotions prétendent donc (du moins dans certains cas) à l'objectivité, mais à une objectivité *relative*. Le paradoxe n'est qu'apparent : en effet, si l'on entend par objectivité *indépendance par rapport à l'observateur*, les valeurs ne sont pas objectives, puisque rien n'aurait de valeur sans qu'il existe d'organismes capables de ressentir quelque chose.

3 Cf. notre article « The Good and the true ».

4 Cf. notre ouvrage *The Rationality of Emotion*, ch. 12, et TAPPOLET, « Les émotions et les concepts axiologiques ».

Voilà donc la différence cruciale apportée par le point de vue axiologique à la notion de rationalité. L'optique axiologique permet de discerner un niveau d'adéquation ou d'erreur indépendant de toute valeur instrumentale. Cette qualité d'adéquation est analogue à la vérité, conçue naïvement comme une sorte de correspondance. Cependant la différence de ce mode de correspondance est qu'il représente non pas une dimension unique comme celle de la Vérité, mais un nombre indéfini de dimensions dont chacune se rapporte à une propriété distincte qui caractérise l'objet formel de chaque état affectif. Les qualités qu'attribuent implicitement nos réactions esthétiques mettent bien ce point en relief : en effet, à part le jugement sommaire qu'on peut porter sur une œuvre d'art comme étant « bonne » ou « mauvaise », les réactions esthétiques n'ont aucun lien direct avec le comportement ; elles se distinguent les unes des autres par leur adéquation (plus ou moins grande) à différentes qualités. Ainsi la peur est jugée plus ou moins adéquate par rapport à la dimension du *danger* ; l'amusement, à la dimension de l'*amusant*, et ainsi de suite. Il s'agit là de ce qu'on pourrait appeler les émotions standard. Mais celles-ci ne forment qu'une petite proportion de toutes les émotions qui constituent notre répertoire émotionnel. Souvent ni les émotions ni leurs objets formels n'ont de noms dans nos langues. Elles n'en sont pas moins spécifiques, et susceptibles d'être plus ou moins adéquatement liées les unes aux autres. Je pense, pour donner un exemple, à la première représentation que j'ai eu l'occasion de voir de *En attendant Godot*. L'œuvre de Beckett venait de sortir ; j'avais treize ans. A côté de moi se trouvait une dame très grave qui n'y voyait que sens tragique. Pendant un échange de bouffonnerie particulièrement cocasse qui déclenchait le rire général, elle se pencha vers moi pour me chuchoter bien haut : « *Mais il ne faut pas rire ! Ce n'est pas censé être drôle !* » Cette dame avait sans doute objectivement tort, et pourtant elle percevait aussi quelque chose d'objectif, un élément sans nom, qui ressemblait au sens tragique sans l'être à proprement parler, et envers lequel elle ressentait une émotion, sans nom elle aussi, dont l'adéquation à cette scène dans sa totalité faisait défaut.

Cette brève esquisse devra suffire à établir la toile de fond sur laquelle se dessineront les deux maximes que je veux proposer. Il s'agit de principes de rationalité émotionnelle qui semblent irréductibles aux principes conventionnels qui régissent la croyance et l'action. Par *principes conventionnels*, j'entends le syllogisme pratique d'une part, et la théorie de la décision d'autre part. Le syllogisme pratique pose une fin en guise

de prémisses majeures, et déduit une action à exécuter à l'aide de prémisses mineures qui relient les moyens à la fin posée par la majeure :

J'ai faim  
Manger cette nourriture assouvirait ma faim  
Donc, il faut la manger

Quant au calcul délibératif de la théorie de la décision, on peut le concevoir comme fondé sur le modèle bayésien. C'est le type de raisonnement qui gouverne explicitement les jeux de hasard, et implicitement tous les autres actes, puisque tout acte peut être considéré comme un pari au sens large. Exemple : je prise la fumée, mais aussi ma santé. Si je puis établir un degré de désirabilité  $f$  au plaisir de fumer, et un degré  $c$  au cancer, je peux soupeser les branches de l'alternative qui se présente à moi en les pondérant de leur probabilité et de celle de leurs contraires. Ainsi ma décision de continuer de fumer ou de ne plus fumer tiendra compte de la probabilité conditionnelle du cancer et du plaisir de fumer, étant donné que je continue ou que je cesse de fumer. Il sera rationnel de continuer de fumer, pourvu que la somme des désirabilités pondérées de la fumée plus le cancer et de la fumée sans cancer, reste plus haute que la somme des désirabilités pondérées de la privation de fumer plus le cancer et de la privation sans cancer.

Cette formule présente un puissant instrument critique. Mais il est notoire qu'elle n'est pas en mesure de fixer les valeurs originales des paramètres concernés, et particulièrement des degrés de désirabilité. Par ailleurs, il semble évident que certaines modifications systématiques de ces paramètres sont plus rationnelles que d'autres. Ainsi on peut toujours se défendre de l'accusation d'être incohérent dans ses préférences d'après la norme bayésienne, en faisant valoir un changement d'avis. S'il est incohérent de préférer A à B, B à C, et en même temps C à A, on peut toujours soutenir qu'entre le moment de manifester les premières et les dernières préférences, on a changé de point de vue ou révisé ses priorités. La théorie bayésienne est sans ressources pour sanctionner de tels changements. A moins de se résigner à abandonner tout jugement de rationalité qui transcende l'instantané, il faut donc recourir à d'autres principes, qui posent certaines contraintes complémentaires par rapport à la théorie de la décision.

L'existence de tels principes ne fait guère de doute. L'estimation de la valeur de certaines options plus ou moins lointaines, par exemple, donne vraisemblablement lieu à un certain escompte. Appelons *hédô* l'unité hédonique. Il semble entièrement raisonnable de n'accorder à un

plaisir de 100 hédos recevables dans un an que l'équivalent de 90 hédos sur le moment. On peut n'être pas d'accord sur le taux d'escompte qui convient ; mais il semble raisonnable de faire reposer ces principes sur la supposition que la distance dans le temps ou l'espace a un effet direct sur la probabilité des divers aboutissements, et par là sur leur désirabilité totale. George Ainslie<sup>5</sup> a montré les conséquences importantes qui découlent du taux d'escompte, qui semblerait défini par une sorte de loi psychologique caractérisée par une formule hyperbolique. Une telle formule n'est pas totalement arbitraire, dans la mesure où on peut montrer qu'elle se calque sur une simple image géométrique qui représente les transformations que subit la grandeur apparente d'une forme lointaine à mesure qu'on s'en approche :

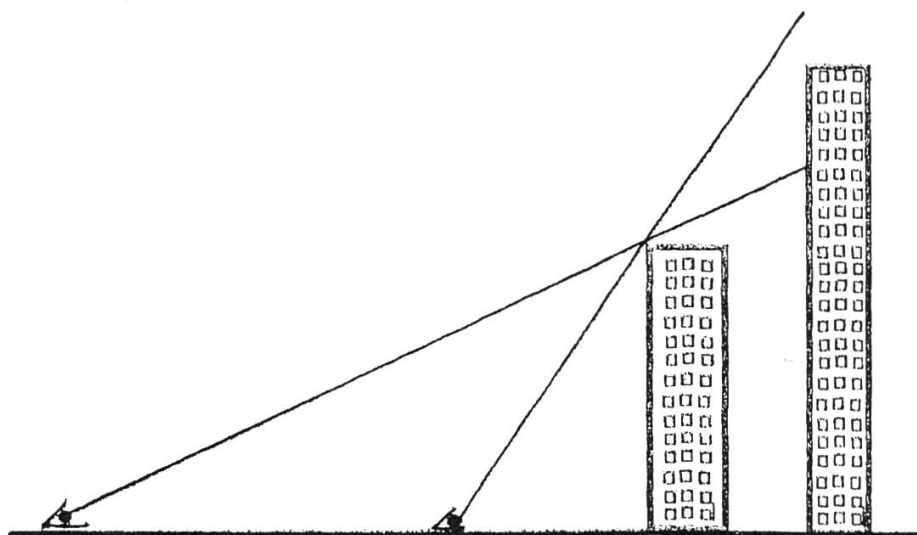


Fig. 1

J'ai déjà commenté cette formule ailleurs, et je n'y reviendrai pas.<sup>6</sup> Je compte plutôt me pencher sur deux autres principes qui se rapportent à la temporalité. Il s'agit de deux principes qui semblent gouverner la rationalité de certaines émotions. Ils sont complémentaires de la rationalité qui régit la simple adaptation des moyens à leurs fins, sans qu'on puisse les déduire ni de la logique, ni de la rationalité bayésienne. Néanmoins, nous verrons plus loin qu'une certaine perspective biologique serait peut-être en mesure d'expliquer l'intuition sur laquelle ils reposent.

5 Cf. AINSLIE, *Picoeconomics: the strategic interaction of successive motivational states within the person*.

6 Cf. notre article «What can't we do with economics? Reflections on Ainslie's Picoeconomics».

Le premier concerne le côté « aspectuel » de certaines expériences. Il s'agit du principe *d'adéquation aspectuelle du désir à son objet*. Il s'agira en second lieu d'un principe qui régit le rapport entre une émotion ressentie dans l'imagination d'un état futur, et l'évaluation que l'on fait du plaisir effectif que nous procurera cet état futur lui-même. Je défendrai la plausibilité de ce principe, avancé pour la première fois par Platon, et que, pour honorer son origine, je désignerai par le nom de *Principe de Philèbe*. Après avoir décrit ces deux principes, je tâcherai de répondre à trois questions : tout d'abord, ces deux principes sont-ils effectivement irréductibles à la rationalité bayésienne orthodoxe ? Deuxièmement, peut-on les expliquer comme étant des principes fonctionnels qui dérivent de la logique de la sélection naturelle ? A cette deuxième question, je tenterai de répondre par un « peut-être » qui penchera plutôt vers l'optimisme pour ce qui est du principe de Philèbe, mais plutôt dans l'autre sens pour ce qui est de la notion de rationalité aspectuelle. Troisièmement, peut-on tout de même espérer intégrer ces principes dans un cadre biologique dans le sens large du mot ? A cette troisième question, je tenterai dans un esprit spéculatif de donner une réponse affirmative.

## 2. Le côté aspectuel de l'expérience

Aristote notait déjà une distinction importante, mais souvent oubliée aujourd'hui, entre diverses façons d'envisager le déroulement des événements dans le temps. Il distinguait en effet les notions d'*état (hexis)* – par exemple : le fait d'avoir faim –, d'*accomplissement (kinesis)* – par exemple : gagner une course, écrire un article – et d'*activité (energeia)*.<sup>7</sup> Une activité, comme un état, n'a pas de durée fixe qui soit déterminée par sa nature même. En revanche, un accomplissement suppose une série de phases plus ou moins distinctes, et n'existe comme tel qu'après que soit terminée la dernière de ces phases. Un état, comme une activité, peut cesser d'exister pour une raison ou une autre, mais ni l'un ni l'autre ne contiennent de fin intrinsèque (au sens de *finalité*) qui constitue en même temps leur fin (au sens de *terminus*).

Pour Aristote, l'*eudaimonia* était surtout affaire d'activité et non d'accomplissement ou d'état. Ces catégories ne sont pas sans leur côté flou (la pensée est le cas par excellence d'une *energeia*, par exemple ;

7 ARISTOTE, *Ethique à Nicomaque*, II, 5 et X, 4.

cependant la solution d'un problème intellectuel ne serait-elle pas un accomplissement ?) Cependant l'idée que l'activité est une composante essentielle d'une vie heureuse exprime fort vraisemblablement une certaine sagesse. Elle suscite deux sortes de questions.

On peut d'abord se demander *pourquoi* l'activité jouerait un rôle particulièrement important dans le bonheur. En effet, certaines composantes de l'*eudaimonia* semblent plus naturellement se ranger parmi les états ou les accomplissements. Le plaisir et la santé, par exemple, sont facilement conçus comme étant des états ; et d'aucuns diraient certainement que le succès – athlétique, professionnel ou autre – joue un rôle essentiel dans une vie heureuse ; or le succès semble être un exemple typique d'accomplissement. Pourquoi donc Aristote privilège-t-il l'*energeia* ? Je mets cette question de côté pour y revenir plus bas.

La deuxième question surgit dès que l'on remarque que les types de propriétés que distingue ainsi Aristote rappellent les catégories purement syntactiques que les linguistes appellent *aspects*. On apprend aux étudiants en grec que l'aoriste n'est pas à proprement parler un temps, puisqu'il indique non pas une position relative d'un événement dans la dimension temporelle, mais une certaine façon d'envisager son déroulement dans le temps : ponctuel, duratif, fréquentatif, ou parfait, c'est-à-dire achevé. De fait, ces distinctions ne s'appliquent pas qu'aux langues étrangères ; ce qu'on apprend à l'école à appeler les temps des verbes en français ou en anglais sont souvent mieux compris comme étant des aspects. L'imparfait en français exprime d'habitude un aspect *continu*, le passé simple exprime plutôt un aspect *ponctuel*, et le passé composé typiquement un *parfait*. Le présent simple en anglais n'est pas en réalité un temps simple, comme on peut s'en rendre compte en comparant « he works at the university » et « he is working at the university ». Seule la seconde expression nous dit ce qu'il fait au moment présent : la première n'exclut pas qu'il ne soit, au moment en question, ni en train de travailler ni à l'université.

On peut donc poser la question : quel est le rapport entre les catégories aspectuelles grammaticales et les catégories distinguées par Aristote ? Il semble bien qu'elles ne soient pas sans rapport. Mais au premier abord, il semblerait que l'aspect grammatical ne saurait distinguer les propriétés intrinsèques des événements auxquels il se rapporte. *J'ai fait du ski ; je fis du ski ; je faisais du ski* : ces expressions peuvent servir à attirer l'attention sur un même événement, mais envisagé d'un point de vue approprié à des contextes différents. Or il est peu vraisemblable que l'apport de bonheur qui découle d'un événement donné ne tienne

qu'à la façon dont on envisage son déroulement dans le temps, plutôt qu'à ses propriétés intrinsèques.

Comment une distinction purement grammaticale pourrait-elle influencer le rôle que joue une certaine chose dans la détermination de l'eudémonie ? C'est à *travers le désir* suscité par la perspective d'un événement, que le caractère objectif de l'activité se voit lié à la subjectivité qui s'exprime à travers l'aspect grammatical de la formule qui le définit. Tout désir doit être (quelque confusément que ce soit) formulé ; on peut donc lui attribuer des propriétés syntactiques qui ne sauraient être directement liées à des événements concrets. L'établissement de ce rapport entraîne la possibilité de poser la question de l'adéquation de l'aspect du désir au caractère intrinsèque de son objet.

Pour mieux mettre en relief la signification de la notion d'aspect, voyons d'abord comment on pourrait soulever la question de la rationalité par rapport à la simple position temporelle de l'événement au moment où le sujet l'envisage. On a vu qu'une option à venir est escomptée en fonction de la période qui nous en sépare. Mais comment choisirait-on parmi des options dont l'une est à venir et l'autre est dans le passé ? Y a-t-il une formule rationnelle qui permette de répondre à la question : « Que préfères-tu : recevoir 100 \$ demain, ou *avoir reçu* 100 \$ hier ? »

On peut imaginer des contextes où la question pourrait ne pas être absolument absurde. Dans un cas d'amnésie, par exemple, où un sujet, qui ignore tout de sa situation et ne sait même pas qui il est, a l'occasion de se choisir une identité : derrière un tel *voile d'ignorance*,<sup>8</sup> l'étrangeté de la situation rendrait la question intelligible, dans la mesure où elle en traduirait une autre : préféreriez-vous être un personnage qui a reçu 100 \$ ou qui va les recevoir ? Mais dans ces circonstances, et en toute autre circonstance normale, une telle question n'aurait de sens que dans la mesure où elle ne serait qu'un détour qui se rapporte au moment présent.

Il n'en est pas de même lorsque au lieu de poser une alternative en fonction du passé et de l'avenir, on la pose en fonction d'un *aspect temporel*. Car certaines perceptions ou certaines émotions semblent naturellement se présenter sous un certain aspect.<sup>9</sup> En particulier, le désir

8 La notion est évidemment empruntée à Rawls (cf. *Théorie de la justice*). Il va sans dire que l'usage que j'en fais n'a aucun rapport avec le sien.

9 Le beau titre du livre de Janet Landman (*Regret: the persistence of the possible*) fournit l'exemple du regret, le qualifiant de *persévérance du possible*.

peut parfois avoir comme objet un événement considéré sous un aspect ou sous un autre. Le désir de courir dans une course, par exemple, peut se présenter comme le désir d'une activité – la course elle-même – ou au contraire elle peut se présenter comme un désir d'avoir accompli quelque chose, au parfait.

Les cas les plus intéressants sont ceux qui fournissent une alternative, dont une branche semble plus «raisonnable» que l'autre. Voyons par exemple *le cas du touriste pressé*. «Que préfères-tu : *visiter le Louvre*, ou *avoir fait le Louvre* ? » Plus d'un touriste pressé, s'il était tout à fait sincère, donnerait la deuxième réponse. Pourtant on en serait gêné, parce qu'on se rend compte que dans un certain sens c'est la mauvaise réponse. Je ne sais pas trop comment *justifier* un tel verdict, pourtant il me semble évident qu'il est normalement irrationnel de préférer certaines choses sous l'aspect du parfait : vouloir *avoir visité* un musée plutôt que *visiter* un musée, *avoir fait l'amour* plutôt que *faire l'amour*, *avoir écouté un beau morceau de musique* plutôt que *l'écouter*, toutes ces formules laissent entendre que l'activité en question n'est pas véritablement source de jouissance en elle-même. Or ce sont là justement (à certaines exceptions près) précisément le genre d'activités qui se justifient par leur qualité intrinsèque et non comme moyen pour accéder à une valeur distincte. Inversement (et sans vouloir nier qu'on puisse là aussi imaginer des contre-exemples), il me semblerait tout à fait rationnel de préférer *avoir rangé ma chambre* au processus de ranger ma chambre.

Voilà donc le principe de *l'adéquation aspectuelle de l'émotion à son objet*.<sup>10</sup> Quel pourrait être le fondement d'un tel principe ? D'où viendrait sa vraisemblance ?

Beaucoup, comme Pascal, nous incitent à jouir du moment, à vivre dans le présent. Pourtant on ne peut envisager un accomplissement sous l'aspect de la durée, puisque tout instant, y compris celui qui transforme un processus en un accomplissement, est sans durée. Par contre rien ne m'empêche de désirer arriver à l'instant qui conclut un certain processus ou une certaine activité. De plus, il n'y a rien là d'irrationnel, pour autant que cette activité vise explicitement à l'accomplissement d'une

10 On pourrait se demander pourquoi le principe est formulé en fonction de l'émotion, alors qu'il semble concerner principalement le *désir*. C'est que justement, la distinction entre les aspects temporels semble concerner la façon d'envisager l'objet, chose qui touche à l'optique axiologique, plutôt qu'au simple fait de vouloir réaliser l'objet – qui touche à la fonctionnalité du désir comme tel, c'est-à-dire comme facteur déterminant de l'acte.

tâche ou à l'achèvement d'un processus dont la valeur dépend en partie ou en tout de l'atteinte de son but. Cependant ce n'est pas toujours le cas, comme le montrent les exemples que je viens de citer. En particulier, si Aristote a raison de voir dans l'activité libre de toute contrainte une condition essentielle de l'*eudaimonia*, il est peut-être irrationnel de placer le gain monétaire, en tant qu'accomplissement, au-dessus de l'activité qu'est la contemplation philosophique.

Un tel verdict peut pourtant sembler arbitraire. Voilà qui nous rappelle à la première question posée plus haut : à quoi tient la supériorité de l'activité, ou de ce qui peut être appréhendé sur l'aspect du continu ? Cette question peut maintenant se diviser en deux :

- (a) Quelle est l'origine de la valeur supérieure de l'activité ?
- (b) Pourquoi dire qu'un désir qui ignore la supériorité de cette valeur est irrationnel plutôt que simplement mal avisé ?

A la première question, on peut se risquer à proposer une spéculation. Ce qu'Aristote appelle *energeia*, c'est aussi ce que l'on pourrait appeler l'activité *ludique* par opposition au *travail*. Le travail, c'est ce qui se justifie par un but dont il ne fait pas lui-même partie. L'activité ludique, par contre, porte en elle sa justification et sa valeur.<sup>11</sup> Dans la réalité, ces deux catégories sont rarement disjointes. Cependant la notion d'*energeia* pose comme idéal une activité ludique pure, et l'on peut se demander sous quel angle on pourrait justifier la prééminence d'une telle idéalisation. Après avoir dûment noté le fait que l'activité, par définition, n'a pas de rendement précis sous forme de résultat concret, on peut vraisemblablement poser l'hypothèse que les organismes dits « supérieurs » ont besoin d'une disponibilité cognitive qu'aucun apprentissage, ni aucune capacité ne peut suffire à façonner si elle se limite à maîtriser un projet utilitaire quel qu'il soit. En effet la nature même d'un projet utilitaire, c'est de trouver sa justification dans la solution d'un certain problème concret. Or il est vraisemblable que chez les enfants comme chez les scientifiques mûrs, les niveaux de réorganisation cognitive qui constituent le progrès intellectuel ne sont possibles qu'à travers une exploration désintéressée.<sup>12</sup> Sur le plan biologique, *seule l'utilité de l'activité ludique repose précisément sur le fait qu'elle n'en a pas.*

11 Aristote distingue le loisir (*scholé*) de l'amusement (*paidia*) (*Ethique à Nicomaque*, X, 7). Sans vouloir effacer la distinction, je range les deux dans la seule catégorie du ludique, par contraste au travail. Cf. Bernard SUITS, *The Grasshopper: Games, Life, and Utopia*.

12 GOPNIK & MELTZOFF, *Words, thoughts, and theories*.

Comment cette adéquation entre-t-elle dans l'esprit du programme axiologique ? C'est qu'elle semble objective dans le sens qu'elle dépasse l'individu ; par ailleurs, elle se base sur certaines réalités humaines plutôt que logiques ou métaphysiques.

La seconde question, qui concernait la différence entre une accusation d'irrationalité et une simple erreur, trouve par la même occasion sa réponse : la différence entre une simple erreur et un cas d'irrationalité est que celle-ci est de second ordre, dans le sens qu'elle gouverne les conséquences générales et statistiques d'une certaine politique. Être irrationnel, c'est manquer à minimiser les erreurs à la longue. Dans un cas spécifique, on peut discuter l'à-propos d'une certaine attitude en fonction de circonstances particulières. Mais en tant que tactique émotionnelle générale, l'oubli du ludique est irrationnel parce qu'il empêche le développement de capacités essentielles à l'être humain.

### 3. Le principe de Philèbe

Le deuxième principe que je voudrais préconiser touche au problème bien connu de l'*acrasie*, ou « faiblesse de volonté ». Ce problème se présente tout d'abord – dans le *Protagoras* de Platon où il fait pour la première fois son apparition – comme un dilemme pour l'hédonisme psychologique. En effet son existence semble impliquer que dans la lutte entre différents mobiles, le plaisir est à la fois gagnant et perdant dans une seule et même situation. Gagnant, puisqu'il semble évident que c'est sous la coupe du plaisir que nous agissons de façon à contredire notre « meilleur choix », et perdant, puisque dans l'optique de l'hédonisme, ce meilleur choix lui-même ne saurait être défini qu'en termes de *supériorité globale du plaisir impliqué dans ce choix*. Cependant une certaine conception radicale de l'hédonisme, qu'on peut baptiser l'*hédonisme immédiat*, rend compte assez facilement de ce phénomène. Il suffit de constater que la désirabilité  $d(A)$  du résultat futur d'une action  $A$  prise sur le moment n'a qu'un rapport contingent, et non logique, avec la désirabilité du choix de faire  $A$  dans le moment présent. La situation de l'acratie peut alors se caractériser à l'aide de la formule suivante :

$$(1) \quad d(A) > d(B) \ \& \ d(\text{Choisir}(A)) < d(\text{Choisir}(B))$$

En langage clair, je juge que  $A$  est préférable à  $B$ , tout bien considéré ; mais l'acte de choisir  $A$ , lui, est jugé moins agréable que l'acte de

choisir B. Or le choix est déterminé par le plaisir qu'apporte l'acte de choisir lui-même, et non par celui qui est envisagé dans la planification qui amène ce dernier.

Dans l'optique de l'hédonisme immédiat, on peut postuler que tout désir suppose un *plaisir anticipé*. Cette dernière expression est ambiguë. Elle désigne à la fois le plaisir que l'on ressent au moment où l'on envisage l'événement futur, et le plaisir que cet événement procurera lorsqu'il aura lieu. La situation de l'acratique implique une disproportion entre la mesure hédonique du plaisir que l'on anticipe, et celle du plaisir apporté par l'anticipation elle-même. La formule (1) s'applique donc bien à l'état d'esprit de l'acratique. Mais si cette formule n'a rien de contradictoire, l'état qu'elle décrit n'en est pas moins irrationnel. On a donc raison de convaincre l'acratique d'irrationalité. Son irrationalité ne se passe pas uniquement (comme on le suppose d'ordinaire) au niveau de son choix, mais au niveau de son affectivité même. Son irrationalité consiste justement en ce qu'il ne respecte pas *le principe de Philèbe*, qui se formule simplement ainsi :

(PP) Le plaisir de l'anticipation doit rester proportionnel au plaisir anticipé tel qu'il sera en réalité.

Dans le *Philèbe*, Platon défend l'idée qu'on peut parler de *plaisirs faux*, qu'il ne faut pas confondre avec les *faux plaisirs*. Les interlocuteurs de Socrate soutiennent qu'on peut bien ressentir un plaisir qui repose sur une fausse croyance – on peut se réjouir d'avoir gagné le gros lot d'une loterie, disons, alors qu'en réalité on n'a rien gagné du tout. Cependant, poursuivent-ils, le plaisir lui-même n'en est pas moins véritable. C'est un *vrai plaisir* tout de même.<sup>13</sup>

Platon ne le nie pas. Il insiste cependant sur le fait que le plaisir d'anticipation a pour objet immédiat une sorte de peinture dans l'âme, qui peut être plus ou moins ressemblante en tant que portrait du plaisir anticipé.<sup>14</sup> Le plaisir d'anticipation se réfère donc à un autre plaisir, le plaisir anticipé, et peut être considéré comme une *représentation erronée* de ce plaisir futur.

Il n'est pas nécessaire de suivre Platon jusqu'au bout de ce raisonnement : je ne m'obstinerai donc pas à parler d'un plaisir faux. Je constaterai seulement qu'il existe une relation contingente entre les deux plaisirs, et qu'il est *irrationnel* que cette relation ne maintienne pas une certaine proportionnalité constante.

13 PLATON, *Philèbe*, 38 a.

14 *Ibid.*, 39 a–40 d.

Il vaut la peine de souligner que l'envers de la possibilité de l'acrasie est la possibilité du progrès moral et émotionnel. Une telle possibilité suppose en effet qu'il peut y avoir, en matière de morale comme d'émotions, des erreurs susceptibles d'être corrigées. La disproportion entre le plaisir anticipé et le plaisir d'anticipation en fournit justement un exemple. L'acrasie est un état conflictuel : or, sans conflit, pas de progrès. Le principe de Philèbe s'inscrit donc clairement dans l'optique générale de l'axiologisme, pour laquelle les émotions constituent, du moins parfois, des perceptions de valeurs. Comme toutes les perceptions, celles-ci peuvent à l'occasion être erronées. Mais comme dans d'autres cas de perception erronée, l'erreur ne se corrige que moyennant d'autres perceptions.

#### 4. L'arrière-plan biologique

Il est temps d'aborder nos trois questions. La première concernait le rapport entre nos principes et la théorie de la décision classique. Il est clair, je crois, que ni le principe d'adéquation aspectuelle ni le principe de Philèbe ne sont des théorèmes de la théorie de la décision. L'adéquation aspectuelle ne saurait être assimilée à un principe de la théorie de la décision, puisqu'elle ne concerne pas la désirabilité elle-même, mais bien la façon dont cette désirabilité est envisagée dans son déroulement temporel.

Quant au principe de Philèbe, on peut mieux faire ressortir son indépendance du calcul bayésien classique si on lui oppose un cas typique d'irrationalité, assez proche de celui qui m'occupe, mais qui semble entièrement réductible au schéma classique. Je veux parler d'une certaine sorte de *projet irréalisable*. Envisageons le cas d'un agent qui est incapable de résister à telle tentation (disons : la tentation de faire T) dans les circonstances C.<sup>15</sup> S'il se connaît lui-même, il saura que le projet suivant est irréalisable : se mettre dans la situation C, et ne pas faire T. Il semble donc irrationnel de former ce projet, dans la mesure où il est irrationnel de former un projet irréalisable.

La différence du cas qui nous occupe tient au fait qu'il ne s'agit ici d'aucun projet, irréalisable ou non. Il s'agit simplement d'une anticipation émotive. Cette anticipation émotive peut agir sur moi comme un mobile, mais qu'elle le fasse ou non ne fait l'objet d'aucune nécessité

15 Cf. ELSTER, *Ulysses and the Sirens*.

logique. C'est donc bien la relation entre l'émotion du moment et son objet futur qui est taxée d'irrationnelle.

On pourrait pourtant objecter que le choix de A ne peut être considéré séparément de A, puisque par hypothèse la désirabilité  $d(A)$  inclut tous les facteurs quelconques qui pourraient être considérés comme pertinents. Le plaisir ou la peine que nous donne le choix lui-même doivent donc être compris dans la valeur qu'on attribue à  $d(A)$ . Et le choix de A, à son tour, n'a de sens que s'il se base sur  $d(A)$ .<sup>16</sup>

Malgré la justesse de ces observations, cette objection n'est pas probante pour les raisons suivantes. On peut, certes, envisager la force motivante d'un désir présent comme étant le résultat d'un amalgame de la valeur attribuée à l'événement en perspective, d'une part, et de celle qui s'attache au choix sur le moment, de l'autre. Mais on ne saurait insister, sans commettre de pétition de principe, que ces deux composantes ne peuvent être départagées. Tout d'abord, ce serait là exclure *a priori* la possibilité de l'acrasie, ce qui semble arbitraire en regard de l'ampleur des débats que cette question a suscité depuis Socrate. Deuxièmement, il est difficile de nier que l'on puisse évaluer une perspective donnée plus ou moins correctement, et difficile aussi de nier qu'une telle évaluation pourrait ne pas engager de jugement passionnel. De là il n'y qu'un petit pas à la possibilité qu'un certain engagement passionnel soit *erroné* par rapport à l'événement envisagé, dans le sens précis requis par l'hypothèse présentée ici. Mais ce qui tranche nettement la question, c'est que l'attitude émotive qui est amenée par mon jugement quant à  $d(A)$  et qui motive mon choix est liée causalement, mais non logiquement, à la valeur émotionnelle du choix lui-même. Par conséquent il est toujours possible que la relation entre eux soit perturbée. Le principe de Philèbe exprime la contrainte normale qui définit cette relation, et sa violation constitue donc le cas pathologique.

Il reste cependant à expliquer pourquoi cette contrainte existe. C'est ce qui nous ramène là la deuxième des questions posées plus haut : si elle ne se réduit pas à la rationalité bayésienne, quel est le fondement de cette contrainte ? Une réponse semble s'imposer : c'est que sans un principe de corrélation systématique tel que le principe de Philèbe, on semblerait condamné à ce qu'il n'y a qu'une relation aléatoire entre ce que je ressens maintenant et ce que je ressentirai plus tard. Or la corrélation entre les deux sert à quelque chose, du point de vue évolutif. L'émotion

16 Fabienne Pironnet a fait valoir cette objection à un colloque à l'Université Montréal en février 2000.

que je ressens aujourd'hui envers quelque perspective d'avenir ne m'est utile que si elle peut me fournir une indication fiable quant à ce que je ressentirai plus tard. Bafouer ce principe, c'est garantir que l'avenir que l'on choisit aura une valeur aléatoire par rapport à la valeur de ce choix lui-même. C'est un peu comme si, dans une décision portant sur des activités en plein air, on se laissait guider par des prévisions du temps qui ne se rapportent que par hasard à la période envisagée.

Ici comme dans le cas de l'adéquation aspectuelle, il est clair qu'il s'agit bien d'irrationalité, et non simplement d'erreur dans un cas particulier, car il n'est pas de relation particulière qui soit requise dans chaque cas de plaisir d'anticipation, mais bien une proportionnalité en général. S'il fallait juger de la rationalité d'un cas particulier où le principe ne se trouve pas respecté, il serait tout aussi raisonnable de dire qu'on a meilleur temps de maximiser tout plaisir, même si c'est pour être déçu plus tard. Encore une fois, c'est par son effet à la longue que le principe de Philèbe s'impose : il s'agit donc bien là d'un principe de rationalité.

Si les principes dont j'ai parlé ne sont pas équivalents, faut-il qu'ils s'organisent en une hiérarchie fixe en cas de conflit ? Le principe d'escompte implique une inversion des valeurs relatives de deux plaisirs dont l'un est proche et l'autre lointain. Or une telle inversion semble violer le principe de Philèbe (sans pourtant toucher au principe d'adéquation aspectuelle). Par contre, les principes aspectuel et de Philèbe vont parfois dans le même sens. En effet, on peut acratiquement désirer avoir vécu quelque chose que l'on sait, malgré le plaisir d'anticipation que l'on prend à envisager son accomplissement, devrait plutôt être vécu et donc désiré dans la durée. Dans ces deux types de cas, et dans les autres où il peut y avoir interférence entre les divers principes en question, doivent-ils être appliqués dans un ordre fixe ? Faut-il que l'un d'entre eux prime systématiquement ? C'est là une question d'un grand intérêt, mais qui restera sans réponse ici. Je me contenterai de risquer l'hypothèse suivante : la question de la primauté d'un principe sur l'autre est une question qui n'est normative qu'en apparence. En réalité, c'est une question empirique, dont la réponse tient aux conditions biologiques particulières dans lesquelles a évolué la tendance naturelle à se soumettre à ces différents principes. On bute aussi sans doute parfois sur des différences individuelles de tempérament, comme celle qui sépare le calcul hédonique « normal » de celui qui préfère remettre la souffrance à plus tard, de celui que l'on pourrait nommer le *faux-masochiste*, qui préfère *avoir vécu* les expériences douloureuses auxquelles il

pense ne pouvoir échapper, et qui donc choisit de remettre le plaisir à plus tard.<sup>17</sup>

## 5. Conclusion

Tout compte fait, le point de vue biologique se doit de s'appuyer sur le comportement, puisque seul le comportement peut donner directement prise à la sélection naturelle. Mais les différents principes qui régissent la rationalité peuvent se situer plus ou moins en recul de la sélection naturelle. Dans cette optique, les deux principes dont il a été question ici occupent chacun une position différente. Le principe de Philèbe, malgré son aspect d'abord paradoxal du point de vue de la maximisation de la totalité hédonique, n'est apparemment qu'à une opération près de la rationalité classique. En effet, il se justifie – du moins en partie – dès que l'on prend compte de l'influence sur la motivation – et par là sur le comportement – que peut avoir le plaisir qu'on éprouve dans la contemplation d'une perspective à venir.

Du fait même qu'il donne lieu à une explication – fût-elle incomplète – à partir de postulats biologiques, ce principe a de plus un caractère qui permet de le ranger dans la catégorie que j'ai qualifiée de *noir-et-blanc* : il se comprend finalement par son utilité pratique.

Le principe d'adéquation aspectuelle, par contre, semble plus éloigné de l'exigence biologique. Il aurait plutôt sa source en commun avec les principes esthétiques, dans leur acception traditionnelle qui veut qu'ils soient coupés de tout intérêt pratique.<sup>18</sup> C'est ce qui lui permet d'épouser des valeurs qui ne portent pas directement sur les chances de succès du comportement, et qui s'étalent sur de multiples dimensions incommensurables.

Même là, pourtant, il est permis de penser que la biologie n'y est pas pour rien. En effet, ce qui caractérise surtout les êtres supérieurs que nous nous targuons d'être, c'est la *disponibilité cognitive*. Or, cette dis-

17 Ces questions, ainsi que le cas du faux-masochiste, s'inspirent de très pertinents commentaires de Bernard Baertschi et de Paul Dumouchel, auxquels je suis bien conscient de n'avoir aucune réponse satisfaisante à donner.

18 Ceci n'implique nullement que le clivage entre les considérations désintéressées de l'esthétique et les considérations pratiques puisse être tenu pour acquis. Au contraire, j'ai expliqué dans «Fetishism and objectivity in aesthetic emotion» pourquoi je me suis vu contraint d'abandonner l'espoir que l'on puisse préserver un domaine d'esthétique désintéressée pure.

ponibilité repose justement sur l'activité ludique : chez l'enfant comme chez le scientifique, il saute aux yeux que l'apprentissage le plus important est celui qui ne porte sur aucun but particulier au moment même. Il porte plutôt sur l'élargissement de nos facultés cognitives. Et celles-ci comprennent, dans la perspective axiologiste, nos facultés émotives dans toute leur complexité.

### Références

- AINSLIE, G., *Picoeconomics: the strategic interaction of successive motivational states within the person*, Cambridge, CUP, 1992
- DE SOUSA, R., « The Good and the true », in *Mind*, vol. 83, 1974
- *The Rationality of Emotion*, Cambridge (Mass.), MIT Press, 1987
- « What can't we do with economics ? Reflections on Ainslie's Picoeconomics », in *Journal of Philosophical Research*, vol. 22, 1997
- « Fetishism and objectivity in aesthetic emotion », in M. HJORT & S. LAVER (dir.), *Emotion and the arts*, Oxford, OUP, 1997
- ELSTER, J., *Ulysses and the Sirens. Studies in rationality and irrationality*, Cambridge, CUP, 1979
- GOPNIK, A. & MELTZOFF, A.N., *Words, thoughts, and theories*, Cambridge (Mass.), MIT Press, 1997
- LANDMAN, J., *Regret: the persistence of the possible*, Oxford, OUP, 1993
- MOORE, G.E., *Principia Ethica* [1903], éd. T. Baldwin, Cambridge, CUP, 1993
- PLATON, *Philèbe et Protagoras*, in *Œuvres complètes*, trad. L. Robin, Paris, Gallimard (Bibliothèque de la Pléiade), 1959
- RAWLS, J., *Théorie de la justice*, trad. C. Audard, Paris, Seuil, 1987
- STEVENSON, CH. L., *Ethics and Language*, New Haven, Yale UP, 1944
- SUITS, B.H., *The Grasshopper: Games, Life, and Utopia*, Toronto, University of Toronto Press, 1972
- TAPPOLET, CH., « Les émotions et les concepts axiologiques », in P. PAPERMAN & R. OGIEN (dir.), *Raisons pratiques*, Paris, éd. de l'École des hautes études en sciences sociales, vol. 4, 1995