

AKRASIA PRATIQUE ET AKRASIA EPISTEMIQUE

Pascal Engel
 Université de Genève
 Programme FNS Connaissance, raisons et normes

Introduction

L'*akrasia* désigne en grec l'absence de contrôle. Mais il y a plusieurs sortes d'absence de contrôle, et plusieurs sortes de contrôle. On peut parler de contrôle sur nos actions, sur nos intentions, sur nos émotions ou nos passions, et peut être – c'est ce qui est en question ici – de contrôle sur nos croyances. Ordinairement l'*akrasia* est définie comme un manque de contrôle sur nos actions, et particulièrement sur nos intentions et notre volonté. C'est pourquoi on l'appelle souvent *faiblesse de la volonté*. Mais il y a aussi plusieurs sortes de manque de contrôle. On peut perdre totalement le contrôle, ou bien partiellement, sur l'un ou l'autre de ses états. Les deux ne reviennent pas au même. L'*akrasia* n'est pas, en général, une perte de contrôle *total*. C'est seulement une perte de contrôle partiel. C'est ce qu'implique la comparaison d'Aristote avec l'homme ivre qui récite les vers d'Empédocle, qui perd en partie le contrôle de lui-même, mais pas totalement (*Ethique à Nicomaque*, VII, 1147b 12). L'akratique n'est pas celui qui perd les pédales. Hume parlait de passions calmes, et Austin évoque le *don* d'Oxford qui, à la *high table*, mange calmement, froidement, toute la bombe glacée¹. L'akratique n'est pas un aliéné ni un sujet compulsif. Davidson évoque Médée : « Je sais bien quelle vilénie j'ai l'intention de faire. Mais plus forte que ce que je peux penser après coup est ma fureur »² et appelle « principe de Médée » le principe selon lequel l'agent perd totalement le contrôle de lui-même et est vaincu par la passion. Mais l'agent akratique ne perd pas le contrôle à la manière de Médée. Il agit délibérément. Il est également conscient de ce qu'il fait.³

La définition la plus générale de l'*akrasia* est donc : la capacité à faire *librement, délibérément ce que l'on juge devoir ne pas faire ou contre son meilleur jugement*. Dans l'*akrasia*, décrite de la manière la plus générale il y a donc deux conditions :

- (1) l'action d'une personne diverge de ce qu'elle juge raisonnable de faire
- (2) cette action est accomplie librement et délibérément

¹ J.L. Austin "A Plea for Excuses", in *Philosophical papers*, 1962, tr. Fr Seuil Paris, cite par Davidson, "How is Weakness of the Will possible?", in *Essays on Actions and Events*, Oxford, Oxford University Press, 1984, tr. Fr. Actions et événements, Paris, PUF 1993, p.48

² Davidson, op.cit, tr.fr. p.45

³ Ce trait n'exclut pas les explications de type freudien ou par l'inconscient (si on les prend, ce qui n'est pas mon cas), mais cela exclut l'idée que l'*akrasia* serait un type de comportement compulsif. Cela ne veut pas dire que des analyses en termes d'émotions, comme celle d'Ariela Lazar ou celle de Mark Johnston soient exclues non plus. Mais ici je ne cherche pas à analyser ce type d'explications. (cf Johnston, Mark, « Self-Deception and the Nature of Mind », in B. McLaughlin, A. Oksenberg Rorty, dirs., *Perspectives on Self-Deception*, Berkeley : University of California Press, 1988., 63-91; A. Lazar, « Deceiving Oneself Or Self-Deceived ? On the Formation of Beliefs « Under the Influence » », *Mind*, 108, 430, 1999, 265-290)

Ainsi décrite l'*akrasia* est certainement une forme d'irrationalité au sens où l'agent akratique est dans un état contradictoire. Le conflit akratique est constitué par la coexistence des trois propositions Tel que Davidson le caractérise, elle réside dans la non compatibilité des trois propositions suivantes :

P1 si un agent veut faire x plus qu'il ne veut faire y et s'il se croit libre de faire x ou de faire y , il fera intentionnellement x s'il fait x ou s'il fait y intentionnellement

P2 si un agent juge qu'il serait meilleur de faire x que de faire y , alors il veut faire x plus qu'il ne veut faire y

P3 Il y a des actions incontinentes (ou akratiques) ⁴

On peut penser que cette définition fait beaucoup de pétitions de principe, et notamment : les actes akratiques sont-ils vraiment libres ? portent-ils sur une ou des actions particulières ou sur un caractère ? jusqu'à quel point sont-ils intentionnels ?). La plupart des auteurs qui ont écrit sur ces sujets ont admis qu'il y avait une famille d'action akratiques, plutôt qu'un seul type d'action de ce genre. Le phénomène, ou les phénomènes de l'*akrasia* soulèvent beaucoup de problèmes. Il n'est pas question ici de les examiner tous. Celui qui m'intéresse le plus particulièrement est le suivant : y-t-il des cas de perte de contrôle, donc en ce sens d'*akrasia* dans le domaine non pas simplement des actions et de la volonté et de l'intention, mais dans le domaine des *croyances* ?

La réponse semble être positive. Il y a, d'abord, une similitude importante entre le problème de l'*akrasia* et celui de la duperie de soi (*self deception*).

Dans la duperie de soi telle qu'elle est couramment caractérisée :

- (a) le sujet croit que P
- (b) mais il désire croire que non P
- (c) ce qui le conduit à croire que non P

Ici il ne s'agit pas de l'action, mais de la croyance, sauf à considérer que la croyance peut être formée ou acquise d'une manière intentionnelle, comme une action (je reviens sur ce point plus bas). La similitude vient du fait qu'il y a une forme d'incohérence, cette fois dans les croyances, et dans le fait que, tout comme dans l'*akrasia*, cette incohérence semble induite par un désir. Mais cela s'arrête là

En revanche, certains auteurs ⁵ont soutenu qu'il peut exister de l'*akrasia* épistémique. Un cas d'*akrasia* épistémique serait, selon Scanlon, le suivant :

« Je peux savoir, par exemple, qu'en dépit des protestations de Jones qu'il est un ami loyal, il est en fait un habile dissimulateur. Et pourtant quand je suis en sa présence, je peux trouver son apparence de chaleur et d'amitié si forte que, bien que je sache à quoi m'en tenir avec lui, qu'on peut quand même lui faire confiance. » (Scanlon, *op.cit*: 35)

Dans un tel cas, j'ai des raisons de croire que Jones est un hypocrite, mais je crois néanmoins qu'il est fiable quand je suis en sa présence. Autrement dit, l'*akrasia* épistémique serait un cas dans lequel on croit quelque chose (que Jones est un ami sincère) mais où en même temps on croit que l'on ne devrait pas le croire.

⁴ Davidson , op cit, tr.fr ; p. 39-40

⁵ J. Heil "Incontinent Believing", Journal of Philosophy, 1983, Rorty, A. 1983 "Akratic Believers", American Philosophical Quarterly, 20, 175-183 1983, T. Scanlon *What we Owe to each Other*, Harvard, Harvard University Press 1996

Mais quelle est exactement la différence entre l'*akrasia* épistémique et la *self deception* ?

Dans le cas décrit par Scanlon, on a, semble-t-il quelqu'un qui croit à la fois que Jones est un hypocrite et qui se retrouve croyant le contraire.

Il y a cependant une différence. C'est que dans l'*akrasia* épistémique à la différence de ce qui se passe dans la *self deception*, du moins selon la caractérisation ordinaire, il y a chez l'agent une *conscience* du fait qu'il *ne devrait pas* croire que P.

S'il doit y avoir un parallélisme entre *akrasia* pratique et *akrasia* épistémique, il faut supposer que le sujet a, en un sens quelconque un *contrôle* sur sa croyance. Ce n'est pas absolument clair dans l'exemple de Scanlon, que l'agent ait un contrôle sur sa croyance que Jones est sympathique. Pour qu'il y ait un parallèle complet avec le cas pratique, il faut qu'on ait les conditions suivantes modelées sur (1)-(2):

(1*) Un sujet a une croyance du premier ordre qui diverge de sa croyance du second ordre au sujet de ce qu'il est raisonnable pour lui de croire

(2*) La croyance du premier ordre est formée de manière délibérée et libre

David Owens appelle la première condition « condition de jugement » et la seconde condition de contrôle⁶. Il donne un exemple un peu différent de celui de Scanlon : je crois que Martin est coupable (toutes les données pointent vers sa culpabilité), mais je reste convaincu de son innocence. Le parallèle entre action akratique et croyance akratique est alors clair : dans les deux cas il y a une divergence entre d'une part une action et une croyance du premier ordre, et d'autre part un jugement du deuxième ordre au sujet de ce qu'il est raisonnable de croire ou de faire, et dans les deux cas, il y a une forme d'activité libre de l'agent. Est-ce que ce genre de cas sont possibles ? Avant de voir si c'est possible il est nécessaire de revenir sur l'*akrasia* selon Davidson.

2. La première analyse de l'*akrasia* de Davidson

Selon Davidson⁷ une action est un événement qu'on peut décrire comme le produit causal d'une intention ou des raisons d'un agent. Une raison d'agir consiste au minimum en une pro-attitude, une valeur ou un but de l'agent, et dans sa croyance qu'en agissant ainsi il atteindra ce but. Des raisons plus complexes peuvent impliquer une multiplicité d'attitudes et de croyances. Le sens que donne Davidson au terme « désir », recouvre une multiplicité d'attitudes motivationnelles. Deux relations entre raisons et actions sont constitutives du fait que les actions peuvent être expliquées par des raisons :

- une relation logique qui fait que l'explication de l'action la rationalise, ou la rend raisonnable aux yeux de l'agent (la montre désirable à la lumière de ses désirs et de ses croyances
- une relation causale, qui est telle que les raisons d'agir causent l'action

Dans la conception classique, socratique et aristotélicienne, du raisonnement pratique, le raisonnement pratique est quasi déductif et de forme

Il est bon d'avoir X

⁶ D. Owens "Epistemic Akrasia", *The Monist* (2002) 85 pp. 381-97.

⁷ Davidson, « Actions, reasons, and causes » (1963), in Davidson 1980, op.cit, tr.fr p.16-36. Davidson a ensuite revise sa conception initiale

Accomplir telle action de type A conduira à obtenir X [ceci est une action de type A]

donc je ne fais pas A (ou non A)

Il y a contradiction, et donc en ce sens conflit akratique, parce que la conclusion inverse devrait être détachée des deux prémisses.

Davidson soutient aussi, apparemment, un principe internaliste – au sens où si un agent juge qu'il est correct (obligatoire, requis) de faire A, alors il est nécessairement motivé à faire A si l'occasion s'en présente⁸) quant aux raisons, qui est très proche de P2 ci-dessus.

La conception que se fait Davidson du raisonnement pratique dans « How is Weakness of the Will possible ? »⁹ est en gros la suivante : l'agent a diverses croyances et désirs qu'il tient comme pertinents pour l'action sur laquelle il délibère. L'agent tient certaines de ces paires de croyances et de désirs comme pesant à l'encontre du fait d'accomplir l'action. Mais le fait que l'agent ait à la fois des raisons *pro et contra* n'implique pas que l'agent soit dans un état contradictoire : l'agent juge certains traits comme rendant l'action désirable et d'autres comme ne la rendant pas désirable. Ces croyances et désirs sont, selon Davidson des raisons *prima facie* pour accomplir l'action. Davidson compare ces jugements à des jugements de probabilité.

Soit $\text{Pr}(m1, p)$ où « m1 » est une considération pertinente relative aux données disponibles et soit p un état de choses

Il soutient que l'on ne peut pas détacher directement et sans restrictions p d'un argument où $(m1, p)$ figure et où $m1$ est une prémisses parce qu'il peut y avoir une autre considération qui rend non p probable. Mais une conjonction de considérations qui respectivement augmentent et diminuent les perspectives qu'un événement ait lieu ne constituent pas une contradiction. Par exemple :

Si baromètre tome, il est très probable qu'il va pleuvoir

Baromètre tombe

Il est très probable qu'il va pleuvoir

Si cieux rouge le soir, il est très probable qu'il ne pleuvra pas

Cieux rouge le soir

Il est très probable qu'il ne pleuvra pas

$\text{Pr}(Rx, Fx)$

Fa

$\text{Pr}(Ra, \text{pr}(Rx, Sx) \text{ et } Fa)$

$\text{Pr}(\text{non } Rx, Sx)$

Sa

$\text{Pr}(\text{non } Ra, \text{Pr}(\text{non } Rx, Sx \text{ et } Sa))$

Dans le raisonnement pratique l'agent pèse les raisons pour et les raisons contre une certaine action. Il forme alors un jugement pratique basé sur les raisons qu'il passe en

⁸ Sur l'internalisme en théorie de la motivation, cf. notamment Michael Smith, *The Moral problem*, Oxford, Blackwell, 1994. Ces distinctions sont bien exposées dans R. Ogien, ed. *Le réalisme moral*, Paris PUF 2000.

⁹ « How is Weakness of the Will possible ? » (1969), in Davidson 1980, op.cit, tr.fr. *Actions et événements*, ch. 2., pp. 36-65

revue. Idéalement il considère toutes les raisons pertinentes pour sa décision. L'agent fait, dans la terminologie de Davidson, un jugement « tout bien pesé » (*all things considered*) qui est conditionnel :

Tout bien pesé, je devrais faire x

Toutes les considérations étant prises en compte, je devrais faire x

Ce jugement tout bien pesé, conditionnel, est suivi d'un jugement additionnel, cette fois non conditionnel et non hypothétique, de la forme « Je devrais faire x » ou « x est désirable ». ce jugement est « catégorique », ou « all out », inconditionnel. C'est un engagement de la part de l'agent de faire l'action en question.

La structure du raisonnement pratique est alors :

Prima facie il est meilleur de faire X que Y ¹⁰

Tout bien pesé je devrais faire X

Je devrais faire X

Je fais X

La conclusion est l'intention (je devrais faire X) ou l'action (je fais X). Elle est non conditionnelle, car on n'agit pas conditionnellement. Mais elle n'est pas en contradiction, quand on échoue à la tirer, avec la prémisse

Tout bien pesé, je devrais aller enseigner en Suisse plutôt qu'en France.

Voici une occasion de le faire

Je ne le fais pas

Une personne peut juger que, tout bien pesé, il est meilleur pour elle de faire quelque chose, mais ne pas avoir l'intention d'accomplir l'action, et ne pas l'accomplir. C'est cette indépendance de l'intention (ou du jugement inconditionnel, car le jugement inconditionnel *est*, pour Davidson l'intention) qui fait que l'agent peut être akratique.

On voit ainsi que Davidson rejette son modèle antérieur causal strict de « Actions, raisons et causes (1963), qui assimilait l'intention à une paire désir-croyance (ce que l'on peut appeler la théorie causale simple).

Selon la nouvelle conception du raisonnement pratique dans « Comment la faiblesse de la volonté est-elle possible ? » l'agent qui forme un jugement inconditionnel (intention) agit en accord avec son meilleur jugement. Dans le cas akratique, l'agent agit contre son meilleur jugement. Mais il n'y a pas de contradiction, bien qu'il y ait une divergence entre le jugement conditionnel tout bien pesé et le jugement inconditionnel qui constitue l'intention.

Ce qui rend l'*akrasia* paradoxale, c'est le fait que les raisons de l'agent akratique ne « rationalisent » pas son action. Comme Davidson le dit dans « Paradoxes of irrationality », les raisons qui sont tenues comme les meilleures par l'agent ne sont pas les *causes* de son action. Dans les cas normaux d'action non akratiques, l'action de l'agent réfléchit le contenu de ses jugements. Dans les cas akratiques, elle entre en

¹⁰ *prima facie* pourrait être traduit par l'expression « de primesaut » que Montaigne utilise souvent.

conflit avec eux. On a affaire à des causes qui ne sont pas les causes « appropriées », ou à des causes qui ne sont pas des raisons *pour* l'action.

La définition de l'*akrasia* selon Davidson est :

« Un agent ϕ ie pour une raison R tout en croyant qu'il y a une autre option d'action ψ qui lui est ouverte, qui est incompatible avec ϕ , et pour laquelle il a une raison R' qui inclut R et sur la base de laquelle il juge que l'option est meilleure que ϕ [« Comment la faiblesse de la volonté est-elle possible ? »tr.fr. p.62]

Le cas le plus fort d'*akrasia* est celui dans lequel l'agent agit contre une raison basée sur toutes les considérations pertinentes (i.e un jugement tout bien pesé) L'agent dans ce cas juge que, selon tous ses croyances et désirs pertinents. Dans le cas de la croyance, Davidson invoque le principe *d'évidence totale* de Carnap : ne croyez qu'en fonction de l'ensemble des données disponibles. Dans le cas du raisonnement pratique, Davidson propose d'invoquer un principe parallèle :

Principe de continence : accomplissez l'action jugée la meilleure sur la base de toutes les raisons pertinentes disponibles (tr.fr. p.64)

Par là il veut dire, les raisons théoriques (croyances justifiées par les données) et les raisons pratiques (désirs, valeur, etc.)

Mais il s'agit plus d'une norme ou d'un idéal que de ce qui se passe en réalité. L'agent la plupart du temps agit sur la base d'un sous ensemble de ses données disponibles et de ses considérations disponibles. Mais l'analyse de Davidson ne requiert qu'une chose ; que l'agent agisse contre ce qu'il juge le meilleur, tout bien considéré, et moyennant éventuellement une révision possible de son évaluation. L'analyse de l'*akrasia* est par conséquent la suivante :

L'agent fait un jugement tout bien pesé en faveur de ϕ ier tout en formant l'intention de ψ ier, et il fait un jugement catégorique conduisant à son action.

L'agent ne se contredit pas au sens où il tiendrait à la fois que P et que non P, car il n'y a pas de contradiction entre juger que, tout bien pesé, il vaut mieux faire X que Y, et faire Y. Comme le premier jugement est tout bien pesé, il n'est pas catégorique, et n'entre donc pas en conflit avec le second. Cela ne veut pas dire, bien entendu qu'il n'y ait pas une tension entre prémisse et conclusion, mais elle n'est pas *formelle* ni déductive.

Une première difficulté ressort de cette analyse : l'agent, s'il ne se contredit pas, semble être rationnel. Car après tout, qu'y a-t-il d'irrationnel, quand on juge que quelque chose est meilleur tout bien pesé à finalement ne pas le faire ? D'autres considérations peuvent avoir pesé en dernier lieu. Pour qu'il y ait conflit, il faut qu'il y ait une stabilité du jugement tout bien pesé.

3. La seconde analyse de l'akrasia de Davidson

Dans « Paradoxes of irrationality » (1980)¹¹, Davidson donne une caractérisation différente de celle qu'il donnait en 1969 dans « Comment la faiblesse de la volonté est-elle possible ? ». Il soutient que l'agent soutient le principe de continence comme un principe du second ordre qui conduit l'agent à se contredire :

« Une pure incohérence interne intervient si je soutiens aussi, comme en fait c'est le cas, que je devrais agir d'après mon meilleur jugement, selon ce que je juge meilleur ou obligatoire, tout bien considéré »

L'agent se contredit car il va contre son principe en agissant. On retrouve la caractérisation ci-dessus, en termes de premier et second ordre

Davidson reprend l'exemple célèbre de Freud, dans *L'homme aux loups*, de l'homme qui voit une branche en travers de l'allée dans un parc, et qui, dans le but d'éviter qu'un passant ne se blesse, la jette dans un buisson, continue sa marche, puis, revient sur ses pas pour poser à nouveau la branche à son lieu initial, de peur qu'elle ne blesse quelqu'un en restant dans le buisson. En retournant dans le parc l'agent va contre son principe du second ordre. Le désir de retourner au parc était une raison pour y retourner, mais il fonctionnait aussi comme une raison pour ignorer le principe du second ordre.

- (i) toutes choses égales par ailleurs je ne devrais pas retourner au parc
- (ii) je dois agir en fonction de mon meilleur jugement
- (iii) je désire retourner au parc
- (iv) je retourne au parc

Davidson caractérise le cas ainsi :

« Il a un motif pour ignorer le principe, à savoir qu'il désire, peut être très fortement, remettre la branche dans sa position initiale. Cependant bien que son motif pour ignorer le principe était une raison pour l'ignorer, ce n'était pas une raison contre le principe lui-même, et quand il est entré de cette seconde façon, il était non pertinent en tant que raison. »(*Paradoxes de l'irrationalité*)

Davidson distingue des raisons pratiques pour les actions et des raisons théoriques ou épistémiques pour les croyances. L'action de retourner au parc est causée par une raison pratique (le désir de remettre la branche en position initiale), mais elle n'est pas très forte. (elle est contrebalancée par de meilleures raisons pratiques. Mais le désir de retourner au parc est si fort, qu'il agit contre son principe du deuxième ordre.

Cette analyse pose deux problèmes. Le premier est que Davidson semble introduire deux actions : l'action accomplie d'une part, et celle d'ignorer le principe de l'autre. La seconde action facilite la première. Mais il ne semble pas y avoir deux actions distinctes. Et de plus si l'on considère que ignorer le principe est quelque chose que l'agent a choisit de faire, sa seconde action n'est pas en contradiction avec son jugement tout bien pesé. Il n'y a que le fait d'ignorer le principe qui est contradictoire avec le jugement tout bien pesé.

Le second problème est plus sérieux. C'est le fait que soutenir le principe de continence qui rend l'action irrationnelle, parce que l'irrationalité ne peut venir que de la contradiction de croyances internes à l'agent :

¹¹ In R. Hopkins et R. Wolheim eds, *Philosophical Essays on Freud*, Cambridge, Cambridge University Press 1980, tr.fr. in *Paradoxes de l'irrationalité*, L'Eclat, Nîmes, 1993

« Ce n'est que quand des croyances sont incompatibles avec d'autres croyances selon des principes défendus par l'agent lui-même – quand il y a une incohérence – qu'il y a un cas clair d'irrationalité » (ibid)

Autrement dit il n'y a irrationalité que quand il y a contradiction *reconnue*. Si l'agent n'endosse pas le principe de continence, l'action akratique ne sera pas irrationnelle de son propre point de vue.

Comme on l'a déjà noté, dans la première analyse de Davidson, il n'est pas clair que l'action de l'agent soit irrationnelle. Si Davidson ajoute la contradiction avec le jugement du deuxième ordre, c'est sans doute qu'il a conscience de ce fait.

Mais est-ce que le fait de croire au principe explique l'irrationalité et la préserve ? Non, car un agent peut avoir le principe du troisième ordre de ne pas suivre son principe de continence du deuxième ordre (il me semble que c'est ce qui se passe avec les trois grosses dames d'Antibes de Somerset Maugham, dont j'évoque le cas dans ma préface à *Paradoxes de l'irrationalité*). Mais si l'on postule un principe du troisième ordre cela conduit à une régression à l'infini du type de celle de la tortue de Lewis Carroll.

Le principe de continence ne peut pas figurer explicitement dans la pensée de l'agent.

Il ne peut pas non plus y avoir de *choix* de l'agent de ne pas tenir compte du principe.

La seconde caractérisation de Davidson conduit à rapprocher (sans les identifier) l'*akrasia* de la *self deception* : l'agent raisonne mal, du fait d'un désir qui obstrue son raisonnement. Dans la première caractérisation au contraire l'agent raisonne correctement, il délibère correctement. Mais il n'adhère pas à son raisonnement. Il n'a pas les bonnes intentions.

4. *L'akrasia épistémique est-elle possible ?*

Les difficultés de Davidson sur l'*akrasia* pratique illustrent la difficulté qu'il y a à considérer qu'il peut y avoir de l'*akrasia* épistémique.

Il est légitime de soutenir qu'il y a des normes de la croyance comme le principe d'évidence totale et le principe selon lequel une croyance correcte et une croyance vraie. Je ne défendrai pas ici ces normes, et ne j'en dirai que peu ici.¹²

Norme de vérité (NV) Ne croyez que P que si P est vrai

Norme d'évidence (NE) Ne croyez que P que si P est basé sur des raisons probantes

(NV) et (NE) sont souvent dites être des normes de la croyance au sens où c'est un trait conceptuel ou nécessaire de notre concept de croyance que l'on ne doit croire que P que si P est vrai, et si l'on a suffisamment de données pour croire que P. En d'autres termes, on ne peut être dit comprendre le concept de croyance que si l'on est en mesure d'apprécier cette norme. C'est ainsi que, selon moi on doit comprendre la célèbre expression "les croyances visent la vérité". Cette conception normative de la norme de la croyance n'implique pas que la vérité soit, en un sens téléologique, le but ou l'objectif que le sujet entend viser quand il acquiert une croyance. Mais ce fait conceptuel quant à la croyance ne nous dit rien quant à la

¹² cf. notamment P. Engel et R. Rorty, *A quoi bon la vérité*, Paris, Grasset, 2005, et « Truth and the Aim of Belief », in D. Gillies, ed. *Laws and Models in science*, King's College, London, 2005, p. 77-97, "Belief and Normativity", *Disputatio*, à paraître 2008

manière dont la norme peut être appliquée et fonctionne. Il nous dit que (NV) est une caractéristique conceptuellement nécessaire de la croyance, mais il ne nous dit pas comment la norme doit être appliquée, ou comment elle peut être violée. Les normes ne sont pas des nécessités logiques ou des lois de la nature. Elles reposent sur de telles nécessités (par exemple en logique) ou sur de telles vérités conceptuelles, mais elles ne sont pas *simplement* des nécessités (en l'occurrence des nécessités sur le concept de croyance).

Elles peuvent ne pas être appliqués dans certains cas, ou même ne s'appliquer que *ceteris paribus*. En particulier, nous voulons pouvoir essayer d'expliquer, dans certains cas comment il est possible pour quelqu'un de croire qu'il a suffisamment de données pour croire une proposition, mais néanmoins de ne pas parvenir à la croire, et vice versa.

Il est donc particulièrement important, pour notre problème, de déterminer si un phénomène tel que l'*akrasia* épistémique est possible. Car d'un côté, sa possibilité, si elle est réelle, semblerait menacer la thèse que (NE), et peut-être (NV) soit une norme constitutive et authentique de la croyance. Car s'il y a des agents doxastiques qui peuvent ne pas croire, et qui ne croient pas en fait, dans un certain nombre de circonstances, ce qu'ils ont d'excellentes raisons « évidentielles » (relatives à l'*evidence*) de croire, et qui de plus sont conscients d'être dans cette situation, comment pourrait-on dire qu'une norme comme celle de la croyance (NE) est réellement une norme de la croyance ? Il y aurait toutes sortes de formations de croyance fondées sur des données suffisantes, et d'autres qui ne le sont pas, et même d'autres dans lesquelles les sujets reconnaissent que leurs croyances sont infondées mais sont néanmoins portés à avoir ces mêmes croyances. En d'autres termes, il n'y aurait pas de norme rationnelle intrinsèque pour la croyance. D'un autre côté, si l'*akrasia* épistémique est impossible, cela renforcerait la thèse selon laquelle c'est une impossibilité conceptuelle que l'on puisse croire une proposition dont on pense consciemment que l'on ne devrait pas la croire, et par conséquent cela renforcerait l'analyse de la rationalité des croyances en termes d'une norme constitutive. Mais alors, comme je l'ai suggéré à l'instant, on ne comprendrait pas non plus comment NE et NV peuvent être des normes de la croyance, parce que nous ne comprendrions pas comment elles peuvent *réguler* la croyance.¹³ Heureusement nous n'avons pas besoin de suivre l'une ou l'autre des branches de ce dilemme.

Quelles sont les conditions qui doivent être présentes pour que l'*akrasia* épistémique soit possible ? Pour que le concept d'*akrasia* ait un sens dans le domaine épistémique, il doit, au moins sous certains de ses aspects cruciaux, être parallèle au concept d'*akrasia* dans le domaine pratique. Comme on l'a vu, une action est akratique au sens pratique si elle accomplit par l'agent librement et consciemment à l'encontre de son meilleur jugement. Le cas parallèle dans le domaine épistémique serait le suivant : une croyance est formée et maintenue de manière akratique seulement si l'agent a une croyance de premier ordre insuffisante quant à ses données et une croyance du second ordre qu'il ne devrait pas entretenir la croyance du premier ordre, et il a la croyance du premier ordre de manière délibérée et volontaire : en d'autres termes les conditions (1*) et (2*) ci-dessus sont remplies.

Notons que ces deux conditions ne s'appliquent pas nécessairement à la duperie de soi ou *self deception* : dans la *self deception* (et encore moins dans le *wishful*

¹³ J'emprunte cette terminologie à Nishi Shah, « How Truth Regulates Belief », *Philosophical Review*, 2003 et le dilemme que je decris ici ressemble un peu à celui qu'il appelle «le dilemme du téléologiste ».

thinking), le sujet n'a pas en parallèle à son jugement de premier ordre un jugement de second ordre selon lequel il ne devrait pas croire ce qu'il croit au premier ordre, et il n'a pas la première croyance de manière volontaire. La dupe de soi-même peut être plus ou moins consciente que l'un de ses jugements entre en conflit avec l'autre, mais elle n'a pas à croire qu'elle ne devrait pas croire ce qu'elle croit.

Il y a donc deux manières par lesquelles on peut nier que l'*akrasia* épistémique soit possible: en premier lieu en niant (1*) soit possible, i.e en niant que l'on puisse manquer de croire ce que l'on considère que l'on ne doit pas croire, et en second lieu que l'on puisse avoir un contrôle sur de telles croyances. Comme la possibilité de (2*) dépend de celle de (1*), on doit examiner d'abord celle-ci.

Beaucoup d'auteurs nient que l'*akrasia* épistémique soit possible. Susan Hurley est l'un d'eux:

“ Au cas où cela serait possible, il pourrait y avoir un conflit à l'intérieur de l'agent , au sens où il y aurait des raisons qui s'affrontent pour exercer l'autorité sur le sujet. Mais s'agissant de ce que l'on devrait croire, il n'y a que la vérité qui gouverne et elle ne peut pas être divisée contre elle-même ni manifester des conflits. Cela fait sens de supposer que quelque chose puisse être bon sous certains aspects et pas sous d'autres; mais cela n'a pas de sens de supposer que quelque chose puisse être vrai sous certains aspects et pas sous d'autres. ¹⁴

En d'autres termes il peut y avoir de nombreuses raisons d'agir, mais il ne peut y avoir qu'une sorte de raison de croire, à savoir la vérité (NV) ou les données (NE). La norme de vérité ne peut pas être divisée. Comme le dit Scanlon (*op cit* 1996) il y a une unité des raisons de croire qui n'existe pas dans les raisons pratiques pour l'action. Dans le cas des actions il y a nécessairement plusieurs sortes de raisons d'agir, qu'on peut peser, tandis que dans le cas de la croyance, il n'y a qu'une seule sorte de raison : des raisons épistémiques. Dans la même veine,, Philip Pettit et Michael Smith écrivent :

“ Imaginez que vos croyances aillent à l'encontre de ce que les données et les faits indiquent. Dans un tel cas, vos croyances ne permettront pas à ces réquisits de demeurer visibles, parce que les croyances qui résistent vous donneront elles mêmes votre sens de ce qui est et de ce qui semble être. Vous vous voyez donc refuser une expérience dont le contenu est que vous êtes en train de croire quelque chose à l'encontre des réquisits de faits et des données. »¹⁵

Comme le notent Pettit et Smith, on peut avancer le même point plus simplement en disant qu'un sujet qui aurait de telles croyances exemplifierait le paradoxe de Moore. Jonathan Adler a précisément soutenu que l'on peut passer d'un « principe subjectif de raison suffisante qui nous dit que :

A. Quand on prête attention à ses propres croyances, on doit les juger être crues pour des raisons suffisantes

A la « norme objective de l'évidentialisme » qui nous dit

¹⁴ S. Hurley , *Natural Reason*, Oxford, Oxford University Press, 1993 : 133.

¹⁵ P. Pettit and M. Smith « Freedom in Belief and Desire », *Journal of Philosophy* 1996: 448

B. On ne doit croire que P que si l'on a des raisons adéquates de croire que P

(et qui n'est autre que (NE)

via le fait que l'agent ne peut pas se reconnaître soi-même comme croyant pleinement que P quand il juge ses propres raisons comme inadéquates.

Il n'y a ici pas de passage d'un *est* à un *doit* parce que (A) est déjà normatif et exprime la reconnaissance d'une norme.

Mais on notera qu'il y a dans ces arguments une assimilation implicite de (NV) et de (NE) , car on tend à assimiler des énoncés de la forme :

(1) Je crois que P, mais cette croyance est basée sur des données insuffisantes

à des énoncés de la forme :

(2) Je crois que P mais non P

Mais (1) et (2) ne sont pas des formes équivalentes du paradoxe de Moore. Je peux par exemple croire

(1') Moriarty a fait le coup mais cette croyance est fondée sur des données insuffisantes.

Mais si je dis

(2') Je crois que Moriarty a fait le coup mais il n'a pas fait le coup.

La seconde croyance n'est pas équivalente à la première. Je peux avoir une information qui menace, même très fortement, ma croyance que Moriarty a fait le coup, et pourtant toujours croire qu'il a fait le coup. En d'autres termes, contrairement à ce que disent Smith, Pettit et Adler, bien qu'il y ait une tension dans (1') ce n'est pas incohérent au sens où (2') l'est. Rappelons nous, dans la théorie de l'*akrasia* de Davidson, sa thèse selon laquelle il n'y a pas de tension entre un jugement « tout bien pesé » en prémisse majeure du syllogisme pratique et un jugement catégorique ou non conditionnel exprimant une intention ou conduisant à une action. Nous avons ici le même phénomène. Il semble parfaitement possible de croire à l'encontre des données dont on dispose, et au sens de la condition (1) ci-dessus, d'être akratique épistémiquement. Et de fait cela arrive souvent, et pas seulement dans les cas de croyances religieuses de type « tertulianesques » ou dans des cas pathologiques¹⁶. En ce sens la norme NE est souvent violée. Cela vient du fait qu'à la différence de la norme (NV), elle autorise des degrés dans les raisons (je peux avoir des raisons *plus ou moins bonnes* de croire que P, en fonction de mes données), tandis que ma croyance que P est vrai ou faux, est *catégorique*. Je propose d'appliquer l'analyse de l'*akrasia* de Davidson dans le domaine pratique à

¹⁶ *Credo qui a absurdum*. Sur les cas pathologiques, voir M. Davies et M. Coltheart, *Pathologies of Belief*, Blackwell, Oxford, 2000 et mon article « Peut-on parler de croyances délirantes ? » in J. Chemouni, dir. *Clinique de l'intentionnalité*, In-Press, Paris, 2001, 157-173

l'*akrasia* théorique, c'est-à-dire de voire le conflit dans les croyances de l'agent victime d'*akrasia* épistémique comme un conflit entre une croyance de premier qui est plus ou moins justifiée pour des raisons plus ou moins bonnes (donc une croyance qui n'est pas tenue comme pleinement vraie par le sujet) et une croyance catégorique du second ordre selon laquelle l'agent ne devrait pas croire la première. Cela rend compte du fait qu'il y a, comme le dit Davidson dans le cas de l'*akrasia* pratique, quelque chose de « sourd » dans le sujet ou de la conscience d'une tension, sans qu'il soit pour autant victime d'une contradiction formelle dans ses croyances.

Devons-nous alors conclure que (NE) n'est pas une norme , puisque l'*akrasia* épistémique est possible ? Non, car la condition (1*) est seulement nécessaire, mais pas suffisante pour l'*akrasia* épistémique. Pour que l'*akrasia* épistémique ait lieu de manière complète, la condition (2*) selon laquelle nous devons avoir un contrôle sur notre croyance déficiente du premier ordre et sur notre croyance de second ordre que cette croyance de premier ordre est déficiente, doit être remplie. Car autrement il n'y aurait aucune différence entre l'*akrasia* épistémique et le *wishful thinking* épistémique. Il n'y a que la première qui est volontaire et inclut un élément de contrôle. Celui qui prend ses désirs pour des réalités et croit ce qui lui plaît n'est pas un akratique (ni une dupe de soi, d'ailleurs). Or il est douteux que la condition (2*) de contrôle puisse être remplie.

Mais est-ce que cette analyse ne conduit pas à considérer le contrôle comme *trop* présent dans l'*akrasia* épistémique ? Ne risque-t-on pas de traiter le sujet victime d'*akrasia* épistémique comme ayant une forme de contrôle volontaire sur ses croyances, et par conséquent d'être exposé aux objections usuelles contre cette idée? Selon l'argument classique de Bernard Williams, nous ne pouvons pas avoir de contrôle de ce genre, tout simplement parce qu'il y a une connexion intime entre l'impossibilité de croire à volonté et le fait que l'on soit conscient du fait que l'on croie ainsi^{17 18}.

Je n'ai pas l'intention ici de reformuler l'argument d'impossibilité du volontarisme doxastique de Williams, que j'accepte pour l'essentiel. Il y a de nombreuses formes de contrôle volontaire et de volonté à distinguer ici, cependant. Mais la plupart des auteurs s'accordent au moins sur les points suivant : croire pleinement à volonté n'est pas possible s'il doit s'agir d'une acquisition directe et immédiate de croyance volontaire, et croire à volonté par des moyens indirects, en se provoquant soit même à croire à plus ou moins long terme (via des drogues, des bourrages de crâne, etc.) est possible, mais n'est pas un véritable cas de croyance volontaire. Pour qu'il y ait une croyance volontaire véritable, il ne suffit pas de vouloir croire, pour des raisons pratiques ou prudentielles, mais de vouloir croire *pour des raisons épistémiques*, au sens où la raison de croire est jugée par l'agent être telle quelle est une raison de la croyance même qu'il a, et pas une raison indirecte ou détournée. En d'autres termes, s'il est possible de vouloir croire , pour des raisons qui sont des raisons sur le *vouloir*, il ne semble pas possible de vouloir croire, pour des raisons qu'on tient appropriées à la croyance. On peut certainement croire que P pour une raison qui est distincte de celle pour laquelle on est justifié à croire que P (en particulier une raison pratique de croire que P), mais on ne peut

¹⁷ cf. Bernard Williams, « Deciding to Believe » (1970) dans *Problems of the Self* (Cambridge : Cambridge University Press, 1973). J'ai commenté cet argument notamment dans "Voluntarism and Volitionism about Belief" in A. Meijers, ed. *Belief, Cognition and the Will*, Tilburg, Tilburg University Press 1999.

¹⁸ Je m'accorde en ce sens avec Owens 2002.

pas délibérément et volontairement croire que P quand on juge que la raison ou les données qui justifient P sont insuffisantes. En ce sens Jonathan Adler a raison. Son principe de raison suffisante (A) pour les croyances empêche la formation d'une croyance de manière akratique au sens où il s'agirait d'une croyance *volontaire*.

Si ce qui précède est correct, par conséquent, la norme évidentialiste (NE) gouverne l'attitude de croyance constitutivement et conceptuellement. Mais cela ne veut pas dire que la norme ne puisse pas être partiellement violée quand la condition (1*) n'est pas remplie : on peut avoir conscience de croire que P pour certaines raisons insuffisantes relativement aux données, et cependant en même temps avoir conscience que l'on ne devrait pas avoir cette croyance. L'agent peut faire l'expérience d'une divergence *partielle* par rapport à la norme. En revanche la condition de volonté (2*) ne peut pas être remplie. C'est donc le fait que la première condition peut ne pas être remplie qui assure la possibilité (partielle, puisque la condition (2*) n'est pas remplie) de l'*akrasia* épistémique. Si la norme ne pouvait pas être violée, elle ne serait pas une norme¹⁹. *

¹⁹ Je reviens sur ce point dans Engel 2007, « Belief and Normativity »

* J'ai déjà analysé ailleurs la question de l'*akrasia* notamment dans « Aristote, Davidson et l'*akrasia* », Philosophie, 3, 1984, et celle de la *self deception* dans « Croyance, jugement et *self deception* », L'inactuel, 3, 1995, 105-122. D'excellents traitements récents sont ceux de R. Holton, *How is strength of the will possible?*, in C. Tappolet C and Stroud, S. 2004 *Weakness of will*, Oxford, Gérard Reach., *Pourquoi se soigne-t-on ?* Le bord de l'eau, Bordeaux, 2003, et de Jon Elster, *Faiblesse de volonté*, Paris, O.Jacob, 2007.