

OH ! CARROLL !
RAISONS, NORMES ET INFÉRENCE¹

Pascal Engel

1. Introduction

Dans un article fameux Bernard Williams (1982) a distingué deux sortes de raisons d'agir : d'une part les raisons internes, celles *pour lesquelles* un agent agit, et qui sont telles qu'il a accès à celles-ci dans sa délibération, et d'autre part les raisons *externes*, celles que l'on a, en général et objectivement parlant, d'agir, qui ne sont pas nécessairement accessibles à l'agent ni n'interviennent dans sa délibération. Williams donne l'exemple suivant. Vous désirez boire le verre qui est sur la table, parce que vous avez soif et croyez que c'est un verre d'eau. Ce sont vos raisons internes de boire ce verre. Mais à votre insu le verre en question est un verre d'essence, qu'il ne serait pas bon pour votre santé d'ingurgiter. Il y a donc des raisons externes de *ne pas* boire le verre. Il y a de multiples manières de comprendre cette distinction (Parfit 1997), mais la différence essentielle est que les raisons internes permettent d'expliquer l'action de l'agent et ses motifs, alors que les raisons externes ne le peuvent pas nécessairement, et sont – c'est le sens du terme « externe » - extérieures aux motivations de l'agent. Les premières sont appelées quelquefois motivantes, alors que les secondes sont appelées normatives : elles relèvent de ce que l'agent *devrait* faire, plutôt que de ce qu'il fait effectivement. Il est tentant de soutenir que les raisons externes correspondent aux lois de la rationalité, idéalement parlant. Mais les lois de la rationalité sont-elles extérieures à l'agent ? Williams pour sa part pensait qu'il n'y a que des raisons internes, qui dérivent toutes de ce qu'il appelait « l'ensemble motivationnel » de l'agent. C'est semble-t-il, conforme à la notion même de raison : car qu'est-ce qu'une raison à laquelle on n'aurait pas accès, une raison qui pourrait en principe n'être une raison pour personne ? Souvent on assimile à une position humienne sur la motivation l'idée qu'il n'y a que des raisons internes, et à une position utilitariste, ou kantienne, ou en tous cas à une forme de cognitivisme moral, l'idée qu'il pourrait y avoir des raisons externes.

¹ Cet article est l'une des versions de nombreux exposés que j'ai faits sur ces sujets et que j'ai publiés ailleurs sous des formes variées (Engel 1998b, Engel 2002, Engel 2005, 2005 a, 2005 b). Sa cause prochaine fut l'atelier où m'invita Gian-Andri Toendury à Fribourg en juin 2008. Merci à Julien Dutant, Adam Leite, Jim Pryor, Joshua Schechter, et Gianfranco Soldati pour leurs remarques. Et à Patrick Ducray pour l'invitation à tourner mes tâtonnements en quelque chose d'écrit.

On peut appliquer cette distinction aux raisons de croire. Il y a les raisons pour lesquelles un agent croit que P (internes) et les raisons qu'il y a de croire que P (externes). Par exemple Watson croit que Moriarty a fait le coup parce qu'il a vu ses traces de pas dans la neige. Mais il n'y a en fait aucune raison de croire cela, car ces traces ne sont pas celles de Moriarty, comme Holmes s'en est aperçu. Watson devrait croire que ce n'est pas Moriarty qui a fait le coup, mais ses capacités intellectuelles limitées font qu'il n'atteint pas cet objectif. Holmes a plus de chances de l'atteindre (même s'il n'est pas parfait).

Enfin on peut appliquer cette distinction aux raisons d'inférer : il y a les raisons pour lesquelles un agent infère Q de P et les raisons qu'il y a d'inférer Q de P. Les raisons d'inférer ne sont qu'une sous espèce des raisons de croire. Les règles de la logique déductive sont par excellence des raisons externes de croire, des raisons que l'agent rationnel devrait avoir d'inférer, même si dans les faits les agents sont loin de suivre ces raisons idéales. Une conception de la logique comme celle de Frege, selon laquelle les lois logiques sont des lois de « l'être vrai », semble tout à fait conforme à l'idée que la logique nous donne des raisons externes d'inférer, qui sont là *qu'il y ait ou non quelqu'un pour inférer*.

C'est l'application de la distinction de Williams aux raisons d'inférer qui m'intéresse ici. Mais pas seulement, car je voudrais essayer de rattacher cette distinction à l'idée selon laquelle il y a, en général, des normes de la croyance. Pour ce faire, je commencerai par examiner le fameux paradoxe de l'inférence de Lewis Carroll, puis je le confronterai aux difficultés que l'on rencontre quand on veut formuler la ou les normes de la croyance. Les deux me paraissent relever d'un même syndrome, qu'une meilleure conception de ce que c'est que d'être autorisé, épistémiquement parlant, à croire et à inférer permet de dissiper. Il y a une forme d'autorisation épistémique *prima facie* qu'on a à utiliser une règle d'inférence et à croire une proposition, qui explique comment des normes de la pensée peuvent réguler nos croyances.

2. Les nombreux problèmes de Carroll

L'histoire fameuse d'Achille et de la Tortue racontée par Lewis Carroll (1895) met en scène Achille présentant à la Tortue deux propositions :

- (A) Deux grandeurs égales à une même troisième sont égales entre elles
- (B) Les deux côtés de ce triangle sont des grandeurs qui sont égales à une même troisième

et la Tortue est invitée à conclure

(Z) Les deux côtés de ce triangle sont égaux entre eux

Elle refuse de tirer la conclusion. Achille lui demande alors si elle est prête à accepter la proposition

(C) Si A et B sont vrais alors Z doit être vrai

La Tortue accepte (C) mais continue de refuser de tirer la conclusion.
Alors Achille lui demande si elle est prête à accepter une autre proposition

(C') Si A et B, et (C) sont vrais, alors Z doit être vrai

Et la Tortue accepte, sans accepter (Z) pour autant. L'histoire veut que l'illustre guerrier et le modeste reptile en soient toujours à ajouter des propositions à cette liste sans que jamais ce dernier n'accepte la conclusion.

L'histoire de Lewis Carroll peut être comprise de multiples manières (Engel 1989, 1998, 2005, 2005 a), mais on peut formuler au moins trois problèmes principaux :

1) *problème de la justification de la déduction*. Etant donné que les prémisses (A) et (B) impliquent la conclusion (Z) (ce que personne ne conteste, puisque l'inférence est une inférence classique en forme de *modus ponens*), qu'est-ce qui justifie ce fait ? Qu'est-ce qui justifie objectivement les principes déductifs de la logique ? Il est clair que le problème de Carroll peut être lu comme une version de ce problème, tel que Quine (1936) l'a posé à propos du conventionnalisme de Carnap en logique : comment peut-on dériver les règles de la logique de conventions si cette dérivation elle-même fait appel à des règles de la logique ?

2) *problème de la garantie inférentielle*. Etant donné qu'on a de bonnes raisons de croire que A et B sont vrais, est-il raisonnable de croire que Z est vrai ? Quelle sorte de garantie inférentielle est-ce que l'on a quand on passe des prémisses à la conclusion ? Le point ici ne concerne pas les raisons « externes » qu'il y a d'user des règles de logique, mais celles, « internes », que le sujet lui-même a quand il raisonne des prémisses à la conclusion

3) *le problème psychologique de l'inférence ou du raisonnement*. A supposer qu'il soit pour moi raisonnable de croire Z sur la base de A et de B, qu'est-ce qui peut forcer notre

esprit à croire Z ? Quelle est la force normative de l'inférence logique sur le comportement inférentiel des agents ? Comment la logique peut-elle mouvoir l'esprit ?

Ces questions ne sont pas identiques. (1) porte sur les raisons ou les justifications objectives des lois de la logique. (2) porte sur les raisons internes que l'agent peut avoir d'inférer. (3) porte sur les motifs psychologiques qu'il a d'inférer.

Certes, pour Achille, c'est tout un que (1) l'inférence de A et B à Z soit objectivement valide, (2) que ce fait donne les meilleures raisons possibles à la Tortue de tirer la conclusion, et (3) que son esprit passe immédiatement des prémisses à la conclusion. Achille semble incarner une sorte de sens commun platonicien : les lois de la logique sont objectivement correctes, et on doit, de ce fait, les suivre. Mais le refus de la Tortue peut être interprété comme introduisant une disjonction entre ces diverses options. En premier lieu la Tortue peut refuser le *modus ponens* à titre de loi logique ou douter que cette loi soit justifiée absolument. Certains logiciens ont émis des doutes sur cette loi (Mc Gee 1985, Priest 1987), mais leur position est minoritaire. En second lieu la Tortue peut douter que le *modus ponens* nous donne toujours de bonnes raisons d'inférer une conclusion. Par exemple, elle pourrait vouloir dire qu'il est parfaitement possible de voir que de P, et si P alors Q, on doit inférer Q, mais que si Q est une proposition absurde on n'est pas tenu de l'accepter (par exemple si vous croyez que la lune a des trous et que si la lune a des trous alors elle est en gruyère, vous devriez croire que la lune est en gruyère : mais de toute évidence vous n'avez aucune raison de croire une pareille stupidité). Enfin on peut interpréter la Tortue comme mettant l'accent sur le problème (3) : il se peut bien qu'on ait les meilleures raisons du monde de croire que Z sur la base de A et B, mais cela ne nous donne pas pour autant un motif psychologique de croire que Z. Peut-être la logique, par elle-même, est-elle insuffisante pour « mouvoir l'esprit » (Blackburn 1995).

La force du problème de Carroll vient en partie de l'incompatibilité apparente des trois propositions suivantes :

- (i) quand une forme d'inférence est valide, c'est un fait objectif qu'elle le soit, et ce fait constitue une raison pour une personne d'inférer certaines choses d'autres choses
- (ii) on doit s'attendre à ce que quelqu'un qui comprend qu'une règle d'inférence est valide ou préserve la vérité tire la conclusion appropriée et raisonne d'après cette règle
- (iii) tirer une conclusion ou faire une inférence est, au moins en partie, un état mental ou psychologique.

Comme on vient de le voir, on peut interpréter la Tortue dans son refus de tirer la conclusion comme mettant en doute (i) ou (ii). Mais il est plus vraisemblable, compte tenu

du fait que la Tortue accepte (C) qui énonce la règle du modus ponens, qu'elle accepte tout à fait (i) et (ii). Ce qui se produit plus vraisemblablement est que la Tortue refuse d'accepter la force psychologique des raisons logiques. Elle émet des doutes quant à la capacité pour la logique de mouvoir l'esprit. Son point, selon cette lecture de l'histoire, porte sur la force normative des normes logiques.

Il est intéressant de noter, comme l'a fait Blackburn (1995), que le problème de Carroll ne concerne pas seulement le raisonnement théorique, qui passe de croyances à des croyances, mais aussi le raisonnement pratique, qui passe de croyances et de désirs à des intentions ou à des actions. Considérez la Tortue confrontée aux deux propositions suivantes :

(A') Je préfère la laitue aux souvlakis

(B') C'est le moment de prendre ma décision

et supposons qu'elle refuse de tirer la conclusion (action) :

(Z') Je choisis de prendre la laitue

Supposons qu'Achille lui propose le principe normatif

(C') Il faut préférer la laitue aux souvlakis

La Tortue ne serait-elle pas victime d'une forme d'akrasia ? (Engel 2005 a)

Nous pouvons généraliser le problème à la notion de norme : les normes, épistémiques, logiques ou morales, ne suffisent pas, par elles-mêmes, à mouvoir l'agent. En outre la *reconnaissance* d'une norme n'est pas suffisante pour garantir son efficacité.

(C) par exemple pourrait être:

(C*) c'est une *norme du raisonnement déductif correct* que si l'on accepte des inférences en forme de modus ponens comme valides, alors on doit, en présence d'instances de cette forme valide, accepter la conclusion

ou même, si une tortue venait à ne toujours pas tirer la conclusion,

(C**) Il y a des justifications rationnelles absolues pour la norme que si l'on accepte des inférences en forme de modus ponens comme valides, alors on doit, en présence d'instances de cette forme valide, accepter la conclusion

Si la Tortue venait à ne toujours pas tirer la conclusion, peut-être qu'Achille pourrait faire appel à un Habermas qui dirait

(C**) Il y a des normes de la rationalité communicationnelle selon lesquelles si l'on accepte des inférences en forme de modus ponens comme valides, alors on doit, en présence d'instances de cette forme valide, accepter la conclusion

On aurait beau taper du point sur la table, la Tortue resterait de marbre. Car (c'est ce que suggère l'histoire) les normes n'ont en elles-mêmes aucun pouvoir *motivant*. Il faut plus (le désir, l'habitude, une force indépendante).

Je vais laisser ici de côté le problème (1) de la justification de la déduction, et supposerai que la Tortue n'est pas une sceptique quant à cette dernière (Dummett 1973, Engel 1989, Boghossian 2003)². Je ne discuterai ici que le problème (2), le problème de la garantie inférentielle, et le problème (3), celui de la psychologie de l'inférence.

3. Externalisme et internalisme quant à l'inférence

Ceux qui sont familiers du paradoxe de Carroll trouveront sans doute que ce qui précède est *much ado about nothing*. Le problème, comme Carroll lui-même le notait dans une lettre à l'éditeur de *Mind* au moment de la publication de son texte, n'est-il pas simplement que la tortue fait une confusion élémentaire entre une proposition figurant dans une inférence à titre de *prémisse* (comme A et B dans l'inférence ci-dessus) et une *règle* d'inférence ? (C) est la règle d'inférence qui permet de passer de A et B à Z, ce n'est pas une prémisse que l'on pourrait ajouter simplement aux autres. Une fois qu'on reconnaît cela, la tortue n'a aucun lieu d'ajouter (C) à titre de prémisse, comme si elle était sur le même plan que A et B (Ryle (1954) formule le point excellemment).

Mais est-ce que cette solution traditionnelle peut résoudre les problèmes (ii) et (iii) de la garantie et de la force psychologique des inférences ? Non. Ces problèmes portent sur la question de savoir si la règle peut *par elle-même* être une raison d'inférer la conclusion, et de plus une raison motivante. La solution traditionnelle assume que le fait même d'être une

² J'ai aussi examiné le problème de la justification de l'inférence dans Engel 2009.

règle implique qu'on la suive. Une certaine sagesse wittgensteinienne l'assume aussi : « On suit la règle aveuglément ». Oui, mais pourquoi faudrait-il suivre la règle aveuglément ? La réponse wittgensteinienne semble se contenter d'un « C'est comme ça », qui a l'apparence du fait profond et immarcescible. Mais est-ce le cas ?

Passons en revue les solutions qui semblent clairement inadéquates. En premier lieu, il y a ce que l'on peut appeler le *cognitivism pur*. Selon le cognitivism, toute norme, qu'elle soit d'action ou de croyance, repose sur un fait et la seule reconnaissance de ce fait est suffisante pour à la fois justifier l'agent dans son action ou sa croyance. C'est un *fait* (logique) que le *modus ponens* est une forme d'inférence valide. La reconnaissance de ce fait doit conduire la tortue à effectuer l'inférence correspondante. Mais c'est clairement insuffisant. Comme le fait remarquer Wittgenstein dans sa discussion sur les règles, il n'y a pas de « rails » qui prédéterminent la manière dont on va inférer psychologiquement parlant. La solution platonicienne semble nous dire qu'on a simplement une intuition de faits transcendants quant à la validité des inférences. Mais cette intuition, et la nature de la faculté qui est supposée la sous-tendre, sont un mystère.

En second lieu, on peut souscrire à une forme d' *internalisme simple*. Selon l'internalisme quant à l'inférence, la reconnaissance de la validité d'une règle entraîne que l'on formule un jugement quant à la règle, lequel en retour justifie et cause la conclusion. En d'autres termes, l'agent est supposé passer par les transitions suivantes dans le cas de notre exemple carrollien :

- (i) toute inférence de la forme *modus ponens* (MP) est valide
- (ii) Cette inférence (de A et B à Z) est de forme MP
- (iii) donc cette inférence (de A et B à Z) est valide

Mais c'est précisément cette transition qui conduit à la régression à l'infini de Carroll. Car si le fait d'effectuer l'inférence dépend d'un jugement préalable quant à la reconnaissance de la validité de MP, alors un jugement supplémentaire est requis pour effectuer l'inférence, et ainsi de suite.

En troisième lieu, on peut défendre une forme d' *externalisme simple*, selon lequel les règles ne sont jamais explicites. Cela a le mérite d'éviter la régression à laquelle conduit l'internalisme simple. Stephen Pinker, examinant le paradoxe de Carroll, donne une « solution » de ce genre :

« Aucun système d'inférence ne suit jamais de règles explicites. A un certain point, il y a simplement à obéir, et c'est tout. Cela veut dire qu'on ne pose plus de questions, et que la règle n'a qu'à être exécutée par la simple force mécanique du système (Pinker 1997, tr. fr :109). C'est aussi la réponse du humien : « Il y a toujours quelque chose d'autre, quelque chose qui n'est ni sous le contrôle du fait ni sous celui de la raison, qu'on doit supposer comme une force brute extrinsèque, si la délibération est supposée conduire à déterminer la

volonté » (Blackburn 1995 : 695). Cette « solution » externaliste a pour elle le mérite de la simplicité et elle est certainement phénoménologiquement correcte. Mais résout-elle le problème? Non. Elle se contente de l'éliminer. Car toute solution décente au paradoxe doit être capable de nous expliquer pourquoi notre croyance en la vérité des prémisses et en la validité de l'inférence *nous justifie* en même temps qu'elle nous *force* à inférer la conclusion. Evoquer une simple force causale brute ne suffit pas. Toute autre solution, y compris celle qui, avec Wittgenstein nous dit qu'on « suit la règle aveuglément », par la force d'une habitude ou d'une disposition, est vouée à ignorer le problème.

Enfin, on peut soutenir que le problème vient de ce que la Tortue ne *comprend* pas les termes employés dans l'inférence concernée (Brown 1954, Black 1970, Stroud 1977). On est tenté de dire que quelqu'un qui refuse de tirer une conclusion en *modus ponens* est tout simplement quelqu'un qui ne comprend pas le sens du conditionnel « si...alors ». Comprendre cette signification, c'est être capable d'effectuer l'inférence. C'est correct, mais cela ne résout pas non plus le problème. Car le fait que la Tortue accepte la règle (C) semble bien indiquer qu'elle comprend la règle du *modus ponens* associée à « si ». Elle la comprend, mais ne l'applique pas. Une solution satisfaisante de ce type devrait être capable de nous dire en quoi le fait de comprendre la signification d'un terme est en même temps une justification de son usage. Elle doit nous dire en quoi le fait de posséder certains concepts nous permet de comprendre une inférence, en même temps que les propriétés de ces concepts justifient les inférences de ce genre. Peacocke (1992) et Boghossian (2003) ont proposé des solutions de ce type au paradoxe, en termes de rôles conceptuels « constitutifs » des termes logiques, qui justifient nos inférences. Posséder le concept de « si » (ou de tout autre terme, logique ou non), c'est *ipso facto* être disposé à effectuer des inférences impliquant ce terme, parce que les conditions de correction de l'usage du terme sont des conditions qui, par définition, nous autorisent à l'utiliser de telle ou telle manière. Mais ce type d'analyse court le risque d'être circulaire, car elle nous dit que ce qui nous autorise à inférer (par exemple une conclusion en MP), c'est l'existence d'une règle constitutive ayant la propriété de nous autoriser à faire ce type d'inférence. Mais en vertu de quoi, et comment, la règle en question nous autorise à faire l'inférence en question est précisément ce que nous souhaiterions expliquer (Enoch et Schechter 2006).

Le cognitivisme simple, l'internalisme simple, l'externalisme simple et la théorie des rôles conceptuels ne sont pas – du moins sans amendements - des réponses adéquates au problème de Carroll tel que nous l'avons posé. Mais ce rapide passage en revue nous permet de diagnostiquer à quelles conditions on pourrait répondre au problème. Tout d'abord, comme je l'ai dit au sujet de l'externalisme simple, toute solution adéquate doit être capable de nous dire en quoi le sujet qui effectue une inférence à la fois a des *raisons* de la faire, et de *bonnes* raisons (garanties). Cela semble exclure toute forme

d'externalisme qui expliquerait l'inférence des prémisses à la conclusion sous la forme d'un processus seulement causal, qui serait par principe inaccessible à l'agent. Plus exactement, cela exclut toute forme simple de cette solution, qui ferait de la transition inférentielle un processus psychologique aveugle ou l'effet d'une simple habitude. Cela montre aussi que d'une manière ou d'une autre il faut qu'il y ait, chez celui qui effectue l'inférence, une conscience au moins potentielle du fait qu'il y a une *norme* ou une *règle* qui nous dit que si on a telle prémisses, alors on *doit* tirer la conclusion. La conception causale simple n'a pas de place pour ce fait. Elle n'a pas de place non plus pour l'idée que la conclusion de l'inférence est *justifiée* par les prémisses et par la transition qui va de celles-ci à la conclusion. Si le processus inférentiel n'est qu'un processus causal/ psychologique, il ne justifie rien et le problème de Carroll n'a rien à voir avec le caractère justifié des inférences logiques. Or il a bien quelque chose à voir avec cela. La Tortue n'est pas simplement stupide, ignorante ou frappée d'amnésie. Son refus de tirer la conclusion a quelque chose à voir avec le caractère justifié ou non de celle-ci. Ensuite ce que la régression de Carroll montre est que si on considère que la force normative de l'inférence implique l'existence d'une croyance ou d'un jugement quant à la norme elle-même, qui opère la transition entre les prémisses et la conclusion – un jugement intériorisé qui nous dit qu'il est correct d'inférer des prémisses de tel type une conclusion de tel type – alors on tombe dans la régression en question. C'est non seulement inacceptable logiquement – aucune régression ne permet de justifier quoi que ce soit – mais aussi c'est parfaitement dénué de toute plausibilité psychologique. Quand il s'agit d'inférences logiques élémentaires comme celles en *modus ponens*, mais aussi d'inférences plus complexes, il semble clair que l'on ne raisonne pas en jugeant : « Cette inférence est de la forme X, les inférences de la forme X sont valides, donc cette inférence est valide » et en inférant, sur la base de ce jugement, une conclusion. En d'autres termes, quand nous inférons, en passant de croyances à d'autres croyances, nous raisonnons *avec* nos croyances, et directement sur leur contenu. Pour parler comme John Broome (2002) il n'y a pas de « croyances-ponts » portant sur nos croyances, qui seraient des jugements de second-ordre sur nos croyances servant à effectuer la transition. En d'autres termes, comme le dit Paul Boghossian, à un moment quelconque, il doit être possible d'utiliser une règle de raisonnement pour parvenir à une conclusion *justifiée* sans que cet usage ait besoin d'être sous-tendu par une connaissance de la règle sur laquelle on s'appuie. Il doit être possible de passer d'une pensée à l'autre d'une manière qui engendre une croyance justifiée, sans que ce mouvement soit fondé dans la croyance justifiée du sujet quant à la règle de raisonnement (Boghossian 2003 : 27).

Une inférence n'est pas un processus aveugle, contrairement à ce que soutient l'externalisme simple. Elle n'est pas non plus basée sur un jugement réflexif de second-

ordre, contrairement à ce que soutient l’internalisme simple. Quand on passe des prémisses à la conclusion, on ne raisonne pas *sur* la règle, mais *avec* la règle et directement sur les contenus des croyances. Mais en même temps ce passage n’est pas une simple compulsion, il est rationnel et il est conscient. Certains auteurs ont parlé ici de « compulsion par la raison » (Brewer 1995). Mais il est bien difficile de voir ce que cela peut vouloir dire, sauf à réintroduire le type de réponse que l’on avait rejeté d’entrée de jeu, consistant à invoquer une faculté d’intuition rationnelle mystérieuse³. Comment la raison peut-elle par elle-même exercer une sorte d’action à distance et avoir une force normative ? C’est tout aussi mystérieux que la solution externaliste. Pourtant cette remarque peut nous guider vers une solution plus satisfaisante.

Une transition inférentielle logique ressemble bien plus à un jugement de perception qu’à une forme d’inférence basée sur un jugement réflexif. Dans un jugement de perception comme dans une inférence logique, on ne passe pas par une inférence pour atteindre le jugement. Quand on juge qu’un objet est rouge, on ne se dit pas « un objet rouge doit avoir telles caractéristiques, donc cet objet est rouge ». La croyance sur laquelle s’appuie notre jugement est non inférentielle. Elle est également *prima facie*. Quand je juge qu’un objet est rouge, je possède une justification *prima facie* que l’objet est rouge. Cela veut dire à la fois que quelque chose *m’autorise* à juger que l’objet est rouge et que ce jugement est *défaisable* : je suis dans des circonstances normales de vision, mes organes fonctionnent correctement, mais je pourrais être le jouet des apparences, et l’objet pourrait seulement m’avoir *semblé* être rouge. On a appelé ce type de justification *prima facie* et on a parlé d’autorisation épistémique (Burge 1997, Peacocke 1999, Wright 2004, Engel 2007). Quand nous les employons dans le raisonnement naturel, la justification que nous avons des règles logiques de base, telles que le *modus ponens*, ou le syllogisme disjonctif, ou la plupart des règles de déduction naturelle, semble être de ce type. De même que les jugements de perception semble réglés par des normes épistémiques selon lesquels il est *a priori* correct, *prima facie*, de juger qu’un objet perçu X a la propriété F, les jugements inférentiels (les transitions inférentielles) semblent réglés par des normes épistémiques selon lesquelles telle ou telle règle logique est *prima facie* correcte. Certes les jugements de perception et les jugements logiques ne sont pas du même type : les jugements logiques sont généraux, et vrais en vertu de leur forme, et ce n’est pas le cas des jugements de perception, qui sont toujours particuliers. Mais il ne s’agit pas d’assimiler le type de garantie *prima facie* que nous avons dans la perception à la garantie *prima facie* des règles logiques. Il s’agit seulement de dire que dans les deux cas elles sont garanties, en quelque sorte, *par défaut*.

³ Cf. Bonjour 1998. J’ai appelé ailleurs, à la suite d’Owens 2000, « internalisme non réflexif » ce type de solution.

De même les règles logiques ne sont pas les mêmes que celles de la perception : elles sont *plus sûres* que les règles qui gouvernent la perception. Mais elles ne sont pas toujours correctes dans tous les cas, notamment parce que, dans la vie courante, nous n'inférons pas seulement sur la base de la logique : nous usons de notre savoir d'arrière-plan. Pour reprendre un exemple de Harman (1986), si Marie croit que si elle va chercher dans le placard de la cuisine elle y trouvera des *Cheerios*, et si Marie va vers le placard (et donc croit qu'elle y va) et n'y trouve que des *Rice Creespies*, elle ne conclura pas, comme la logique devrait l'y conduire, qu'elle est face à des *Cheerios* ; au contraire elle cessera de croire cela, et croira plutôt qu'il y a des *Rice Creespies* dans le placard.

On peut laisser de côté la question de la justification des inférences logiques, qui n'entre pas en jeu directement dans le problème de Carroll. La garantie, ou l'autorisation, à laquelle on a affaire dans des transitions inférentielles simples, est de type *prima facie*. Elle est non inférentielle, ce qui veut dire qu'il n'est pas nécessaire, pour être justifié à effectuer une transition inférentielle, de recourir à une autre inférence qui justifierait la première, et ainsi de suite. Elle suppose donc une forme de connaissance de base, de type non inférentiel, du même ordre que celle qui nous fait, dans la perception, reconnaître immédiatement des formes et des couleurs. Elle est telle, pour reprendre la terminologie de Wright (2004), que la garantie que l'on a de la vérité des prémisses se *transmet* à la conclusion.

Comparer la justification des règles logiques de base de la déduction naturelle à celle des jugements perceptifs ne revient pas non plus à traiter notre reconnaissance des règles comme une forme d'intuition immédiate. Ce serait revenir à une variante de la position platonicienne faisant appel à une faculté mystérieuse. La compétence que nous avons quant aux règles logiques n'est pas basée sur une faculté d'intuition. La comparaison avec les jugements perceptifs tient seulement au fait que nous avons une garantie *prima facie*.

Il y a plusieurs sortes de notions d'autorisation épistémique (Engel 2007). Mais on peut suggérer que c'est ce type d'autorisation qui forme la base de notre garantie inférentielle et qui explique la transition rationnelle dans une inférence comme celle qu'on offre à la Tortue. La Tortue, dans l'histoire de Carroll, nie simplement qu'on ait une autorisation épistémique de ce genre. Mais c'est faux : si l'on comprend les prémisses, on *a* une telle autorisation. Avec les théoriciens du rôle conceptuel, Peacocke et Boghossian, cette réponse admet que l'autorisation épistémique est associée aux concepts logiques (ici le conditionnel). Mais elle n'en découle pas strictement. De même que nous formulons des jugements de perception dans des contextes particuliers, qui sont vrais dans les conditions normales de perception et *mutatis mutandis*, nous formulons des jugements inférentiels dans des contextes particuliers. Mais cette réponse ne minimise-t-elle pas la force du *doit* logique ? La tortue n'est pas simplement autorisée à inférer la conclusion ; elle *doit* inférer la conclusion à partir des prémisses. C'est une obligation, et non pas une simple permission

ou autorisation. Il est vrai que les règles logiques déductives sont par définition contraignantes, alors que celles qui concernent la perception, ou l'inférence inductive ne le sont pas au même sens. Mais il ne s'ensuit pas que l'application d'une règle particulière dans un contexte de raisonnement naturel soit toujours contraignante. Il n'y a aucun doute que si je crois que la Terre a été créée en six jours et que si la Terre a été créée en six jours, elle a été créée en moins d'une semaine, je *dois* croire que la Terre a été créée en moins d'une semaine. Mais évidemment, ma raison de croire que cela ne tient pas qu'à la logique. Il est parfaitement absurde de croire que la Terre a été créée en moins d'une semaine. Cela n'a rien à voir avec la logique. Dans la plupart des contextes quotidiens, le raisonnement logique n'entre qu'en partie dans nos conclusions. Mais le problème de Carroll n'est-il pas précisément un contexte où l'on offre *seulement* un raisonnement élémentaire ? Oui, en effet, mais cela n'empêche pas les inférences particulières que l'on fait, même dans ces contextes, sont relatives à un contexte d'hypothèses (par exemple sur la géométrie dans laquelle on travaille). Je n'entends pas par là dire que les inférences logiques ont quelque chose de contextuel. Au contraire leur justification est absolue. Mais dans le raisonnement courant, la plupart de nos justifications sont de type *prima facie*.

La réponse que je viens d'esquisser au paradoxe de Carroll, que l'on peut appeler *internalisme sophistiqué*, n'exclut pas nécessairement une réponse de type externaliste, selon laquelle notre connaissance *prima facie* des règles de logique est basée sur un savoir tacite, lui-même basé sur des compétences psychologiques réalisées dans le cerveau. Par définition ces règles sont inaccessibles au sujet de manière consciente, et elles font partie de son équipement inné. Cette version de l'histoire n'est nullement exclue par la réponse internaliste sophistiquée (cf. Engel 2005). Mais elle est en elle-même insuffisante, car il y a bien une garantie rationnelle dans la transition inférentielle. Le sujet doit pouvoir avoir accès à sa justification.

Quelle réponse l'internalisme sophistiqué apporte-t-il au problème des raisons de croire ? Elle est double. Dans une inférence logique élémentaire, on a des raisons de croire la conclusion sur la base des prémisses, en vertu de normes ou de lois, qui sont celles de la logique. Ces raisons sont absolues, non relatives, et externes : ce sont des raisons que tout sujet *devrait* reconnaître. Mais ces raisons ne suffisent pas, par elles-mêmes, à justifier la transition inférentielle. Il faut pour cela aussi des raisons internes, accessibles à l'agent, qui justifient *pour lui* son inférence. Ce sont celles que je propose d'appeler des autorisations épistémiques ou *prima facie*. Elles opèrent sur la base d'un savoir non inférentiel.

4. Les normes de la croyance

Je voudrais à présent essayer de tirer quelques leçons de cet examen du paradoxe de Carroll pour une question plus générale encore, celle de la nature des normes épistémiques, et de la croyance en particulier.

Il semble plausible de soutenir qu'il existe des normes épistémiques qui gouvernent la croyance correcte. La plus évidente semble être la norme de vérité : une croyance est correcte quand elle est vraie (dire d'une croyance qu'elle est fausse semble être une objection fatale contre elle, ou en tous cas une raison suffisante pour l'abandonner). La croyance est également gouvernée par la norme de raison suffisante : une croyance est correcte seulement si elle est fondée sur des raisons suffisantes. Enfin une croyance est simplement fondée sur une norme de rationalité : une croyance est correcte seulement si elle est rationnelle (logiquement cohérente). Ces conditions ne définissent pas la croyance (il y a des croyances fausses, non fondées, irrationnelles), mais elles sont des normes de la croyance *correcte*.

Tout le monde n'accepte pas l'existence de telles normes. Les philosophes sceptiques quant à l'existence de normes épistémiques (Stich 1990), déflationnistes (Horwich 1998) ou les naturalistes (Papineau 1999) rejettent carrément l'idée qu'il y ait des normes de ce genre, et qu'elles gouvernent en quoi que ce soit nos croyances. D'autres philosophes soutiennent qu'il y a des valeurs épistémiques, et peut-être des buts épistémiques (le vrai est le but de la croyance), mais n'acceptent pas l'idée de normes (Steglich-Petersen 2006). Je laisserai de côté ici ces discussions (cf. Engel 2007, à paraître) et supposerai simplement que la conception normative est correcte. La question est de savoir en quel sens elle l'est. Car si elle l'est, il se pose pour la régulation de la croyance un problème semblable à celui qui se pose pour la régulation de l'inférence par des règles logiques.

Admettons qu'il y a une norme de la croyance, que l'on peut formuler

(NV) Une croyance est correcte si et seulement si elle est vraie (Wedgwood 2002, Engel 2005)

Mais que veut dire le fait que la vérité soit une norme de la croyance ? Cela veut-il dire que celui qui croit que P a le *but* de croire quelque chose de vrai (Velleman 2000), ou bien qu'il est prêt à obéir à des impératifs du type de

Si P est vrai, croyez que P !

Ne croyez P que si P est vrai !

Ou à des prescriptions comme

On ne doit croire que ce qui est vrai

Une norme n'est pas une raison. Certes si c'est la norme de manger avec ces doigts et non avec une fourchette ou des baguettes dans tel pays, cela me donne une raison de manger avec mes doigts, mais je peux respecter la norme sans avoir aucune raison de manger avec mes doigts (cela me dégoûte). De même avec les croyances : si je crois que P, je crois *eo ipso* ce que P implique, parce que la croyance est soumise à une norme de rationalité, mais cela n'implique pas que j'aie une raison de croire que P. Je peux, pour reprendre le même exemple, admettre que si je crois que la Terre a été créée en six jours elle a été créée en moins d'une semaine, sans pour autant avoir la moindre raison de croire cela (Millar 2005 : 70 sq). Une norme est plutôt du type de ce que Broome (2000) appelle un « réquisit de rationalité ».

Mais comment les normes opèrent-elles ? Comment exercent-elles leur normativité ? C'est ici que nous retrouvons le problème de Carroll. Si l'on admet qu'une norme gouverne le comportement (ici le comportement doxastique, dans le cas du paradoxe de Carroll, le comportement inférentiel) par le biais d'un impératif (« crois que ce qui est vrai ! ») ou d'une prescription (*on ne doit croire que ce qui est vrai*), et si ce jugement est explicite, alors ne faut-il pas, pour appliquer la norme ou la règle, un autre jugement qui nous dise quand il est correct d'appliquer la règle, et ainsi de suite ? On retrouve le problème classique que posait Wittgenstein au sujet des règles. De même pour appliquer la norme NV il faudrait faire un jugement selon lequel la norme s'applique. Cela conduit nombre d'auteurs à traiter les normes comme *implicites* dans une pratique (voir par exemple Brandom 1994).

Mais (comme me l'a fait remarquer Crispin Wright) traiter les normes comme implicites, tout comme, dans la régression de Carroll, traiter les jugements « ponts » comme implicites ne résout pas le problème. Car même si l'on suppose qu'il existe un mécanisme inconscient ou tacite qui passe par des transitions du type « le modus ponens est une règle ; donc il faut appliquer le modus ponens », la régression aura lieu dans le système lui-même. Où le système trouvera-t-il la ressource pour échapper à la régression ? Et si l'on se réfugie dans le fait brut – le système passe *de facto* d'un état à l'autre, on ignore le fait que l'on infère pour une raison. Le paradoxe de la norme guette tout autant la conception normative de la norme de la croyance que le paradoxe de Carroll guette la conception normative internaliste de l'inférence.

Il semble donc tout aussi peu plausible de soutenir que la norme de vérité pour la croyance implique que le sujet, au moment où il forme une croyance, obéit à un impératif – si P est vrai, crois que P et si P est faux, crois que non P ou ne crois pas que P. En fait

quand on se demande si l'on doit croire que P, on ne se demande jamais si l'on croit que P, ni même s'il est correct de croire que P si P est vrai. On se demande seulement si P est vrai, et si la réponse est positive, on croit que P. En d'autres termes, une fois qu'on a découvert que P est vrai, il n'y a plus lieu de se poser de question. Comme le disent les auteurs qui ont, à la suite de Gareth Evans (1982 : 225), mis en avant ce trait, la vérité de P est *transparente* par rapport à la question de savoir si l'on croit que P : une fois que P est reconnu vrai, la question ne se pose plus (Moran 2000, Shah 2003). Quand on a établi pour soi-même que P est vrai, il n'y a plus lieu de se demander si l'on a une raison de croire que P. On a ainsi reconnu de manière directe et explicite la norme de correction de la croyance, qui est la vérité. Au contraire quelqu'un qui se demande s'il a une raison de croire que P entretient nécessairement une relation indirecte avec la vérité de P. Mais la vérité n'est pas dans une relation indirecte avec la croyance : elle est directe et immédiate, et est intrinsèque à l'affirmation du croire. La transparence de la vérité à la croyance veut dire que quand je délibère quant à savoir si je dois croire que P, je ne fais pas un raisonnement dont la prémisse serait

Je ne croirai que P que si P est vrai
P est vrai
Donc je dois croire que P

A partir du moment où je reconnais la vérité de P, j'ai réglé la question de savoir si je dois le croire. La norme de correction fait partie du concept même de croyance.

La même difficulté se présenterait si nous analysions, comme on le fait souvent, le fait de suivre une règle au moyen de la notion d'intention, sous la forme suivante :

J'ai l'intention de croire ce qui est en accord avec NV
Croire que P est en accord avec NV
Donc j'ai l'intention de croire que P

Mais une telle inférence, qui est supposée nous conduire de l'intention de suivre la règle (ou la norme) pour croire que P implique l'existence d'une autre croyance, celle selon laquelle ma croyance est en accord avec la norme. Et la régression à nouveau menace. La relation entre la norme de vérité et la croyance n'est pas une relation entre une croyance au sujet de la norme et la croyance qui est acquise. Elle est une relation entre la norme et le concept de croyance, que comprend toute personne qui a effectivement ce concept et cette relation est immédiate, comme le montre le phénomène de la transparence. La croyance est le seul état mental qui manifeste cette transparence. Conjecturer que P, ou imaginer que P

ne sont pas, en ce sens, transparents. L'épistémologie des raisons de croire, pas plus que l'épistémologie des raisons d'inférer, n'est une épistémologie inférentielle. En ce sens, Wittgenstein avait bien vu : la relation que nous avons aux raisons est directe et d'accès immédiat. Si, face à la considération d'une proposition P, nous nous demandions : « Vais-je croire que P ? » et devons faire appel à une norme – ne croire que P que si P – notre explication serait circulaire. Mais il n'y a aucune circularité de ce genre.

À ce point, la Tortue pourrait revenir vers nous et nous dire : « Soit, admettons que vous ayez proposé une réponse satisfaisante à la question de savoir comment des normes logiques ou épistémiques viennent réguler notre comportement inférentiel ou doxastique et dissipé le spectre de la régression à l'infini qui nous menace dès que nous supposons qu'un jugement au sujet de la correction de ces normes intervient dans notre motivation à les appliquer. Admettons que vous ayez raison sur le fait qu'il y a une relation immédiate et directe entre la norme et son application, que vous appelez « autorisation *prima facie* » ou « transparence ». Cela ne m'explique toujours pas pourquoi je devrais être *forcée* à la conclusion, ou dans le cas de la croyance être forcée à croire que P »

Je ne vois évidemment pas d'autre réponse possible à cette question que celle qui consiste à soutenir que l'individu qui possède le concept de *conditionnel* ou celui de *croyance* est *disposé* à faire des inférences en *modus ponens* et à considérer que la vérité de la proposition qu'il va croire est la réponse correcte à la question « dois-je croire que P ? » Et l'on retrouve alors nécessairement la difficulté au sujet de la notion « de suivre une règle » que Kripke, dans sa lecture de Wittgenstein, avait soulignée : les dispositions à elles seules ne peuvent pas expliquer pourquoi l'agent qui suit une règle ou une norme est *forcé* à la conclusion. De fait avoir une disposition à inférer, ou à utiliser le concept de croyance, n'implique pas que l'on exerce toujours ce concept quand les circonstances l'exigent. Et après tout, ne croyons nous pas bien souvent des choses fausses, et n'inférons-nous pas souvent très mal ? Comme l'ont noté les critiques de Kripke sur ce point, invoquer des circonstances idéales ne suffit pas. De plus invoquer une disposition semble impliquer la postulation d'une force causale brute et revenir à recourir à un type d'explication externaliste dont nous avons vu qu'elle ne permet pas de rendre compte du paradoxe de Carroll.

Mais toutes les dispositions ne sont pas des dispositions causales, comme la solubilité ou la fragilité. Certaines dispositions sont des dispositions rationnelles, des dispositions propres à la seule raison. Quand on ne les a pas, il n'y a rien à faire. Et le cas de la tortue l'illustre. Si elle n'a pas ces dispositions, il n'y a rien ni personne qui puisse jamais la forcer à la conclusion.

RÉFÉRENCES

- Boghossian, P. 2003, “Blind Reasoning”, *Proceedings of the Aristotelian Society*, Supp. vol. LXXVII, 225-249.
- Brandom, R. 1994. *Making it Explicit*, Harvard, Harvard University Press.
- Brewer, B. 1995, “Compulsion by Reason”, *Proceedings of the Aristotelian Society*, Supp vol. LXIX, 237-254.
- Black, M. 1970, “The Justification of Logical Axioms”, *Margins of Precision*, Cornell: Cornell University Press.
- Blackburn, S. 1995, “Practical Tortoise Raising”, *Mind*, 104, pp.696-711.
- Boghossian, P. and Peacocke, C.eds. *New Essays on the A priori*, Oxford: Oxford University Press
- Boghossian, P. 1999, “Knowledge of Logic”, in Boghossian & Peacocke 1999.
2002, “How are objective Reasons Possible ? in Bermudez & Millar 2002.
2003, “Blind Reasoning”, *Proceedings of the Aristotelian Society*, 77, 225-248.
- Bonjour, L. 1998. *In Defense of Pure Reason*, Cambridge, Cambridge University Press
- Brewer, B. 1995, “Compulsion by Reason”, *Proceedings of the Aristotelian society*, sup.vol. LXIX, 237-254.
- Broome, J. 2000, “Normative Requirements”, in Dancy, J., *Normativity*, Blackwell, Oxford, 78-99.
2002, “Practical Reasoning”, in Bermudez and Millar 2002.
- Brown, D. G. 1954, “What the Tortoise taught us”, *Mind*, 63, 250, 170-179.
- Carroll, L. 1895, “What the Tortoise said to Achilles”, *Mind*, 1, 278-280.
- Engel, P. 1989, *La norme du vrai*, Paris, Gallimard, tr. Engl. 1991, *The Norm of Truth, An Introduction to the Philosophy of Logic*, Hemel Hempstead, Harvester Wheatsheaf and University of Toronto Press.
1998, “La logique peut-elle mouvoir l’esprit?”, *Dialogue*, XXXVII , I, 1-27.
2005, “Logical Reasons”, *Philosophical Explorations*, 8, 1, 21-38.
2005 a, “Dummett, Achilles and the Tortoise”, in R. Auxier & L. Hahn, eds. *The Philosophy of Michael Dummett*, Library of Living Philosophers, Open court, LaSalle Illinois.
2005 b, “Truth and the Aim of Belief”, in D. Gillies, *Laws and Models in Science*, London, King’s college publications.
2006, “How to Resist a Tortoise”, in J. Goodenough, ed. *Paradoxes*, Norwich Lectures in Philosophy, University of East Anglia, Norwich.
2007. *Va savoir! De la connaissance en général*, Paris, Hermann.
2007 a, « Belief and Normativity », in *Disputatio* II, 23, 153-177.

- 2007 c, « Logique », in *Philosophie*, Ed. Eyrolles, 49-76.
- 2009, « La démonstration », in J.-L. Poirier (éd.), *Actes du colloque "Enseigner la philosophie"*, Ministère de l'Éducation Nationale, disponible à l'adresse suivante : <http://www.touteduc.fr/file/colloque%20philo.pdf>
- À paraître. « Epistemic Norms », in Berneker S. & Pritchard, D. (eds.). *The Routledge Companion to the philosophy of Knowledge*, Routledge, London.
- Enoch D & Schechter, J. "Meaning and Justification: The Case of Modus Ponens" *Noûs* 40 (4), 687-715.
- Harman, G. 1986. *Change in View*, Cambridge Mass: MIT Press.
- Horwich, P. 2000. *Meaning*, Oxford, Oxford University Press.
- Mc Gee V.1985. "A Counterexample to Modus Ponens", in *Journal of Philosophy* , 82(9) : 462-471.
- Millar, A. 2005 *Understanding People, Normativity and Rationalising Explanation*, Oxford, Oxford University Press.
- Moran, R. 2000. *Authority and Estrangement*, Princeton, Princeton University Press
- Owens, D. 2000. *Reason without Freedom: the Problem of Epistemic Normativity*, London, Routledge.
- Papineau, D. *The Roots of Reason*, Oxford, Oxford University Press.
- Parfit, D. 1997, "Reasons and Motivation", in *Proceedings of the Aristotelian Society*, supp. vol. 71, pp. 99-130.
- Pinker, S. 1997, *How the Mind Works*, Cambridge Mass, MIT Press ; tr. fr. *Comment fonctionne l'esprit*, Paris, O. Jacob 2000.
- Priest, G. 1987. *In Contradiction: A Study of the Transconsistent*. Dordrecht, Martinus Nijhoff Publishers, 2nd ed. Oxford University Press 2006.
- Quine W. V. O. 1936, "Truth by Convention", in *the Ways of Paradox*, Harvard, Harvard University Press.
- Shah, N. 2003, "How truth governs Belief", in *Philosophical Review*, Vol. 112, No. 4, 447-482.
- Steglich Petersen, A. 2006, "No Norm needed: On the aim of Belief", in *Philosophical Quarterly*, (225):499-516.
- Stich, S. 1990. *The Fragmentation of Reason*, Cambridge, Mass, MIT Press.
- Stroud, B. 1979, "Inference, Belief and Understanding", in *Mind*, 179-196.
- Velleman, D. 2000, "On the aim of Belief", in *The Possibility of Practical Reason*, Oxford, Oxford University Press.
- Wedgwood, R. 2002, "The aim of Belief", in *Philosophical Perspectives*, 16, 267-297.
- Williams, B. 1981, "Internal and External Reasons", in *Moral Luck*, Oxford, Oxford University Press.

Wright, C 2002. “On Basic Logic Knowledge”, in Bermudez, J. and Millar, A. (eds.), *Reason and Nature*, Oxford, Oxford University Press.

2004, “Warrant for nothing, and foundations for free?” *Proceedings of the Aristotelian society*, volume 78 (1), 167–212.