

Réflexivité épistémique et défense forte du sens commun. Remarques sur l'épistémologie de Pascal Engel

Jean-Baptiste Guillon

DANS **PHILOSOPHIA SCIENTIÆ** 2017/3 (21-3), PAGES 5 À 37

ÉDITIONS **ÉDITIONS KIMÉ**

ISSN 1281-2463

DOI 10.4000/philosophiascientiae.1299

Article disponible en ligne à l'adresse

<https://www.cairn.info/revue-philosophia-scientiae-2017-3-page-5.htm>



Découvrir le sommaire de ce numéro, suivre la revue par email, s'abonner...

Flashez ce QR Code pour accéder à la page de ce numéro sur Cairn.info.



Distribution électronique Cairn.info pour Éditions Kimé.

La reproduction ou représentation de cet article, notamment par photocopie, n'est autorisée que dans les limites des conditions générales d'utilisation du site ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Toute autre reproduction ou représentation, en tout ou partie, sous quelque forme et de quelque manière que ce soit, est interdite sauf accord préalable et écrit de l'éditeur, en dehors des cas prévus par la législation en vigueur en France. Il est précisé que son stockage dans une base de données est également interdit.



Philosophia Scientiæ

Travaux d'histoire et de philosophie des sciences

21-3 | 2017

N'allez pas croire !

Réflexivité épistémique et défense forte du sens commun. Remarques sur l'épistémologie de Pascal Engel

Epistemic reflexivity and a robust defence of common sense. Thoughts on Pascal Engel's epistemology

Jean-Baptiste Guillon



Édition électronique

URL : [http://](http://philosophiascientiae.revues.org/1299)

philosophiascientiae.revues.org/1299

DOI : [10.4000/philosophiascientiae.1299](https://doi.org/10.4000/philosophiascientiae.1299)

ISSN : 1775-4283

Éditeur

Editions Kimé

Édition imprimée

Date de publication : 30 octobre 2017

Pagination : 5-37

ISBN : 978-2-84174-816-7

ISSN : 1281-2463

Distribution électronique Cairn



CHERCHER, REPÉRER, AVANCER.

Référence électronique

Jean-Baptiste Guillon, « Réflexivité épistémique et défense forte du sens commun. Remarques sur l'épistémologie de Pascal Engel », *Philosophia Scientiæ* [En ligne], 21-3 | 2017, mis en ligne le 30 octobre 2019, consulté le 04 décembre 2017. URL : <http://philosophiascientiae.revues.org/1299> ; DOI : [10.4000/philosophiascientiae.1299](https://doi.org/10.4000/philosophiascientiae.1299)

Tous droits réservés

Réflexivité épistémique et défense forte du sens commun. Remarques sur l'épistémologie de Pascal Engel

Jean-Baptiste Guillon
Collège de France (France)

Résumé : Dans cet article, je discute l'épistémologie de Pascal Engel, en particulier sa stratégie de réponse aux arguments sceptiques dans *Va Savoir!*. Après avoir présenté de manière synthétique (section 1) les grands axes de cette stratégie (c'est-à-dire la stratégie du « mooréanisme internaliste »), je reviens avec plus d'attention sur deux éléments de cette stratégie avec lesquels je suis en désaccord : (i) le rejet par Engel de tout principe de réflexivité épistémique, et (ii) le rejet par Engel d'une défense « forte » du sens commun. Je défends qu'un certain principe de réflexivité (à savoir le principe de Méta-Cohérence, défendu par Michael Huemer) ne peut pas être raisonnablement rejeté (section 2) ; puis je montre que l'épistémologie d'Engel ne permet pas d'apporter une réponse satisfaisante au défi sceptique que pose le principe de Méta-Cohérence (section 3). Enfin, je soutiens que la seule manière de répondre de manière satisfaisante à ce défi est de recourir (comme Reid, mais pas comme Moore, Lemos ou Engel) à une défense « forte » du sens commun, c'est-à-dire à une stratégie dans laquelle les propositions de sens commun sont justifiées parce que elles sont de sens commun (section 4).

Abstract: In this paper, I discuss Pascal Engel's epistemology, in particular his anti-sceptical strategy as developed in *Va Savoir!*. After a first section, where I give a summary of the main ideas of his anti-sceptical strategy (i.e., "internalist neo-Mooreanism"), I concentrate my study on two elements of this strategy with which I disagree, namely (i) Engel's rejection of any principle of epistemic reflexivity, and (ii) Engel's rejection of a "robust" defence of common sense. In section 2, I argue that there is at least one reflexivity principle—namely Michael Huemer's Meta-Coherence Principle—which cannot be reasonably rejected. In section 3, I show that Engel's epistemology does not

have the resources to tackle the sceptical challenge that a sceptic may raise by using the Meta-Coherence Principle. Finally, in section 4, I argue that the most satisfactory way to tackle sceptical arguments based on Meta-Coherence is to adopt a “robust” defence of common sense, i.e., a defence (adopted by Reid but not by Moore, Lemos or Engel) according to which the propositions of common sense have some kind of justification in virtue of the very fact that they are common sense propositions.

C'est un honneur et une joie pour moi d'être invité à discuter les écrits d'un philosophe qui a eu sur l'évolution de mes réflexions une telle influence, tant au niveau de ses thèses qu'au niveau de la conception de la tâche philosophique. Chacun sait l'importance de l'œuvre de Pascal Engel au niveau académique international. Ce qui me frappe, personnellement, c'est une qualité philosophique qui n'est pas si fréquente parmi les auteurs de première importance – à savoir, un sens du rôle de l'activité philosophique dans nos sociétés. Engel permet d'aller au-delà de l'alternative désespérante entre une philosophie comme pur jeu intellectuel entre spécialistes et à l'extrême inverse une « philo » populaire qui croit avoir une signification sociale parce qu'elle s'adresse à monsieur tout le monde, mais pour ne lui livrer que du « prêt à philosopher », ou des maximes de vie façon « fortune cookies ». Engel rappelle, et démontre par ses écrits, que le philosophe a un rôle véritable, qui est de promouvoir et développer, notamment chez les intellectuels (ou les « clercs »), le sens de la rationalité et la valeur de la recherche de la vérité. D'un point de vue théorique, Engel a notamment défendu une théorie de la vérité comme valeur, et une théorie des normes épistémiques en général. Mais il a également mis en pratique cet attachement aux valeurs épistémiques par une activité infatigable de discussion et de critique de ceux, parmi ses contemporains, qui s'éloignent de ces valeurs ou les négligent. Il est pourtant si facile – tant de philosophes le font, et de tous bords – de pratiquer le type de philosophie que l'on préfère sans se soucier de ce qu'écrivent les autres écoles ; mais Engel a toujours montré une disposition remarquable à entrer en dialogue constructif avec ses adversaires pour inviter, voire provoquer, à la rationalité ceux qui seraient trop heureux de pouvoir tranquillement « déraisonner en rond ». À ce titre, Engel offre un exemple dont j'espère que je saurai m'inspirer.

Quant au contenu philosophique, aux thèses particulières d'Engel, elles ont eu une grande influence sur ma réflexion en philosophie de la connaissance. Mais à la différence d'autres philosophes, qui m'ont inspiré essentiellement par les *désaccords* que j'avais avec eux, je dois dire que je me suis trouvé presque toujours d'accord avec les thèses défendues par Engel – ce qui rend plus difficile l'exercice présent de dispute philosophique... En cherchant bien, je suis néanmoins parvenu à trouver deux questions (nous verrons qu'elles sont liées) sur lesquelles je pense être en désaccord, à savoir la théorie du sens commun et la question de la réflexivité épistémique.

Engel a souligné dans diverses publications qu'il faisait grand cas de la tradition philosophique du sens commun [Engel 2007, 2008, 2012]. Mais il y a des points de désaccord au sein même de cette tradition. Un point de dissension qui est relevé par Noah Lemos est le suivant : pour Thomas Reid, le fait qu'une croyance soit de sens commun – qu'elle soit une croyance profondément ancrée et largement partagée – « constitue une certaine raison pour la croire ». Mais un certain nombre de ses successeurs, notamment G. E. Moore, ne l'ont pas suivi sur ce point : bien sûr, Moore défend que « certaines propositions de sens commun sont épistémiquement justifiées pour nous », mais il ne s'engage pas à affirmer « qu'elles sont épistémiquement justifiées pour nous *en vertu* du fait qu'elles sont des propositions de sens commun, ou en vertu du fait qu'elles sont profondément ancrées ou largement répandues » [Lemos 2004, 6]. On pourrait dire que Moore propose une défense *du sens commun*, tandis que Reid propose plutôt une défense *par le sens commun* (la notion de sens commun devient le motif ou la raison qui assure la crédibilité de telle ou telle croyance). Engel s'appuie essentiellement sur G. E. Moore pour sa conception du sens commun : nulle part (à ma connaissance) il n'utilise le procédé reidien de défendre une thèse de sens commun *parce qu'elle* est de sens commun, et l'on trouve au contraire des déclarations qui vont dans le sens d'un refus de ce type de défense épistémologique¹. Pour ma part, je défendrai que seule la conception « forte » (ou reidienne) du sens commun permet de répondre vraiment au problème sceptique.

Pour voir cela, il faudra évoquer le deuxième point de désaccord, qui est intimement lié, à savoir la question de la réflexivité épistémique, c'est-à-dire la question de savoir si nous « savons que nous savons » ou si nous sommes « justifiés à croire que nous savons ». Engel s'appuie notamment sur les arguments de Timothy Williamson pour rejeter l'exigence de réflexivité. Là où la démarche d'Engel est particulièrement riche, c'est qu'il sait faire place à une forme légitime d'internalisme, alors que le rejet de la réflexivité épistémique est souvent associé à un rejet pur et simple de toute forme d'internalisme. Je suis entièrement d'accord avec Engel lorsqu'il montre que l'externalisme de la connaissance ne suffit pas, et qu'il faut reconnaître une forme d'internalisme ; là où je ne suis pas sûr de pouvoir le suivre en revanche, c'est lorsqu'il soutient que cet internalisme peut être séparé de toute exigence de réflexivité épistémique. C'est pourquoi je défendrai ici qu'il nous faut bien trouver un motif ou une justification ultime pour considérer (réflexivement) nos croyances comme des connaissances, et c'est précisément cette justification (ou « méta-justification ») qui ne peut se trouver, à mon sens, que dans un recours fort (ou reidien) au sens commun.

1. Je pense notamment aux réserves qu'il exprime contre « toute forme de sanctuarisation de la connaissance commune » [Engel 2007, 17] et plus encore au rejet explicite d'une justification de sens commun des croyances « anti-sceptiques » (*je ne suis pas un cerveau dans une cuve*, etc.) : « Il n'est pas nécessaire d'admettre qu'on ait cette justification par défaut, *parce que ce serait une croyance de sens commun* » [Engel 2007, 194, je souligne].

Je commencerai cette discussion par une présentation d'ensemble de l'épistémologie de Pascal Engel avec tout ce qu'elle a à mes yeux d'élégant et de convaincant (section 1), afin de mieux situer le point précis du désaccord. Je montrerai ensuite pourquoi on ne peut pas faire l'économie d'une exigence de réflexivité épistémique (section 2) et je montrerai comment cette exigence soulève un problème sceptique auquel Pascal Engel n'apporte pas, à mon sens, de réponse satisfaisante (section 3). Enfin, je présenterai une défense forte du sens commun comme réponse au problème de la réflexivité épistémique (section 4).

1 Structure générale de l'épistémologie de Pascal Engel

On peut résumer la stratégie épistémologique de Pascal Engel en disant qu'il s'agit d'un néo-mooréanisme qui intègre des éléments d'internalisme. Mais cette caractérisation de néo-mooréanisme n'est pas conceptuellement très explicite et, de fait, elle est parfois employée dans des sens différents. Il me semble plus clair de caractériser la position des épistémologues par rapport au(x) problème(s) qu'ils entendent résoudre. En l'occurrence, le néo-mooréanisme (et celui de Pascal Engel en particulier) est une réponse au problème sceptique, et même plus précisément à une certaine version du problème sceptique qu'est l'argument de clôture (ou argument de l'ignorance). Le livre *Va savoir!* est entièrement structuré autour de cet argument de clôture, mais il aborde néanmoins d'autres arguments sceptiques dont il me semble utile de rappeler ici les réponses pour mieux voir la cohérence d'ensemble de la stratégie engélienne. Je partirai donc de la caractérisation la plus générale du projet d'Engel comme un projet de réponse aux arguments sceptiques. Et pour chaque argument sceptique, j'essaierai de souligner comment Engel donne à chaque fois une réponse concernant la notion (externaliste) de connaissance, et une réponse pour la notion (internaliste) de justification, les deux réponses étant toujours, chez Engel, d'une remarquable cohérence entre elles.

Le premier argument sceptique évoqué dans *Va savoir!* est l'antique argument du trilemme ou de la régression [Engel 2007, 11] : nos connaissances doivent reposer sur d'autres connaissances, de telle sorte qu'elles constituent des chaînes de connaissance (ou de justification) qui sont soit infinies (mais c'est impossible!), soit circulaires (mais alors elles ne justifient rien!), soit terminées par des croyances sans fondement (et donc non justifiées!). À ce problème antique, Engel adopte la réponse qui, à ma connaissance, a toujours été majoritaire parmi les épistémologues et l'est encore aujourd'hui, à savoir le fondationnalisme : certaines connaissances sont « basiques », en ce qu'elles ne reposent pas sur l'inférence à partir d'autres connaissances mais sont délivrées immédiatement par certaines sources épistémiques, telles que la perception, la mémoire, l'introspection, etc. Disons tout de suite qu'Engel applique cette solu-

tion fondationnaliste non seulement à la notion (essentiellement externaliste) de connaissance, mais aussi à la notion de justification : si nous avons des facultés épistémiques (telles que la perception) dont la fiabilité externaliste permet de produire basiquement des connaissances [Engel 2007, chap. 1], nous avons aussi (au niveau internaliste) des justifications qui nous sont données de manière immédiate (par exemple par nos *expériences* perceptives) [Engel 2007, 156–157].

Deuxième argument sceptique, l'argument infaillibiliste : les scénarios sceptiques montrent que nos croyances ne sont pas à l'abri de *toute possibilité d'erreur* : elles *pourraient* (dans un monde possible) être fausses. Mais alors, comment peuvent-elles être des connaissances ? Comment peuvent-elles être justifiées ? Pour répondre à cet argument précis, Engel adopte de nouveau une réponse qui est très majoritaire chez les épistémologues contemporains (mais qui a été notoirement rejetée par Descartes) à savoir une forme de *faillibilisme* – c'est-à-dire le rejet du principe infaillibiliste selon lequel une croyance doit être à l'abri de toute possibilité d'erreur pour être une connaissance ou pour être justifiée. Au niveau de la connaissance (externaliste), le faillibilisme s'exprime comme *fiabilisme* : il suffit qu'une croyance soit *fiable* pour qu'elle soit une connaissance, or il est possible qu'un processus soit fiable sans être infaillible [Engel 2007, 45–51]. La version internaliste du faillibilisme, c'est la reconnaissance que nos justifications sont *révisables* ; et si l'on combine l'idée de révisabilité des justifications avec l'idée, vue plus haut, de justification immédiate ou basique, on obtient la notion contemporaine de « justification *prima facie* » [Engel 2007, 156–157]. La justification *prima facie* est la version internaliste du fondationalisme faillibiliste. Mais cette idée de justification basique et révisable n'est pas encore suffisante pour parler d'un « *néo-mooréanisme* internaliste », selon la caractérisation qu'adopte Engel.

C'est le troisième argument, l'argument de clôture (ou argument de l'ignorance), qui justifie la terminologie de « *néo-mooréanisme* ». C'est aussi à ce niveau de la discussion qu'Engel défend une thèse qui est loin de faire l'unanimité au sein des épistémologues. L'argument de clôture repose (comme le précédent) sur les scénarios sceptiques (du type Malin Génie ou Cerveaux dans une cuve – CUC), mais il s'intéresse à la manière dont nous pouvons (ou non) connaître la proposition anti-sceptique (« je ne suis pas un CUC ») : d'après le principe de clôture, je *dois* connaître cette proposition anti-sceptique pour pouvoir savoir que j'ai deux mains, or il semble que je ne peux pas savoir que je ne suis pas un CUC, donc je ne peux pas savoir que j'ai deux mains. Face à cet argument, il existe essentiellement trois solutions (outre la réaction sceptique elle-même, bien sûr) : la solution anti-clôture, la solution contextualiste et la solution « *néo-mooréenne* ». Sans entrer dans les détails, les deux premières solutions reviennent à admettre la prémisse d'ignorance anti-sceptique, à savoir que (au moins dans un sens pertinent) nous ignorons que nous ne sommes pas des CUC. Et toute la subtilité consiste alors à essayer de voir comment éviter la conclusion sceptique radicale (que nous ignorons avoir deux mains). Engel n'est pas convaincu par ces stratégies subtiles et choisit la

voie des néo-mooréens, c'est-à-dire de cette catégorie d'anti-sceptiques qui sont prêts à affirmer *non seulement* que nous savons que nous avons deux mains, *mais aussi* que nous *savons* que nous ne sommes pas des CUC. C'est-à-dire que Engel admet entièrement le principe (ou l'engagement) de clôture selon lequel nous sommes engagés à savoir les conséquences logiques (reconnues) de nos connaissances. Et il décline cette réponse « néo-mooréenne » à la fois au niveau externaliste de la connaissance et au niveau internaliste de la justification. Au niveau externaliste, il s'appuie pour cela sur les théories de la *sécurité* (notamment de Timothy Williamson) qui permettent de maintenir la clôture de la notion de connaissance [Engel 2007, 104 *sqq.*]. Au niveau internaliste, il réfute les arguments de Crispin Wright contre la « transmission de garantie », ce qui permet de dire que ma justification (perceptive) à croire que « j'ai deux mains » me justifie pleinement (par déduction) à croire que « je ne suis pas un CUC » [Engel 2007, 157–162].

Sur chacune des trois étapes que je viens de présenter, je suis entièrement convaincu par les arguments de Pascal Engel. J'adhère donc sans réserve à son « néo-mooréanisme combinant externalisme de la connaissance et internalisme de la justification ». On peut résumer cette épistémologie au moyen du tableau 1 ci-dessous, qui permet de voir la grande symétrie et l'élégance formelle de la position d'ensemble.

	Pour la connaissance (notion externaliste)	Pour la justification (notion internaliste)
réponse à l'argument du trilemme	foundationalisme (facultés de connaissance basique)	foundationalisme (justifications basiques)
réponse à l'argument de la possibilité d'erreur	fiabilisme des facultés (non infaillibilisme)	défaisabilité des justifications <i>prima facie</i>
réponse à l'argument de clôture (ou de l'ignorance)	clôture, et sécurité des croyances anti-sceptiques (= néo-mooréanisme)	justification des croyances anti-sceptiques par transmission de garantie (= néo-mooréanisme internaliste)

Tableau 1 – Réponses de Pascal Engel aux arguments sceptiques

Pourtant, dans le cours de ces démonstrations (plus précisément dans le cours de la réponse à l'argument de l'ignorance) Engel introduit un autre thème, qu'il semble considérer comme nécessairement lié avec les éléments précédents, et dont je suis beaucoup moins convaincu. C'est le thème du rejet de la réflexivité épistémique. Ce thème est d'abord introduit dans sa version externaliste, comme rejet du principe de « luminosité », ou « principe *KK* » :

(KK) Nécessairement si je sais que p , alors je sais que je sais que p . $Kp \Rightarrow KKp$

Le rejet de ce principe est une manœuvre maintenant classique de l'épistémologie externaliste pour répondre au problème sceptique. L'idée est la suivante :

pour que ma croyance d'avoir deux mains (ou de ne pas être un CUC) soit une connaissance, il suffit qu'elle soit fiable ou sûre *en fait* (ou *d'un point de vue externe*); il n'est pas nécessaire que j'aie en plus la capacité de second ordre de *savoir* qu'elle est fiable (une chose est qu'elle soit une connaissance fiable, autre chose que je *sache* qu'elle est fiable). Peut-être que nous ne sommes pas capables *d'un point de vue internaliste* de nous assurer que nos croyances sont fiables, mais cela ne les empêche pas d'être *en fait* des connaissances fiables. Le sceptique tirerait alors implicitement sa conclusion d'une confusion entre le niveau externaliste (est-ce que ma croyance est fiable?) et le niveau internaliste (puis-je savoir qu'elle est fiable?).

Ce raisonnement typiquement externaliste est présenté par Engel à un moment [Engel 2007, 107–108] où il n'a pas encore dévoilé toutes ses cartes, et en particulier à un moment où il n'a pas encore exposé l'idée (à mon sens très judicieuse) que l'externalisme ne peut suffire à lui seul pour répondre au scepticisme et qu'il faut bien introduire une certaine dose d'internalisme si l'on veut construire un anti-scepticisme pleinement satisfaisant². Étant donné cet ordre de présentation, le lecteur pourrait s'attendre à ce qu'Engel réintroduise une certaine légitimité de la réflexivité épistémique lorsqu'il abandonne ensuite le point de vue strictement externaliste pour faire droit à la part nécessaire d'internalisme. Mais ce n'est pas la voie qu'adopte Engel : conservant le strict parallélisme que nous avons déjà vu entre connaissance et justification (cf. tableau), Engel applique à la justification internaliste le rejet de réflexivité qu'il avait d'abord exposé pour la connaissance externaliste. Les justifications *prima facie* qu'il défend introduisent bien un élément d'internalisme dans la mesure où c'est bien un élément de la vie mentale du sujet (une croyance non inférentielle, une expérience) qui constitue sa justification, mais elles justifient « sans pour autant être réflexives au sens où elles feraient l'objet d'une conscience actuelle que nous les avons » [Engel 2007, 156–157]. Il semble donc qu'Engel choisisse ici la voie d'un internalisme « faible », un internalisme *sans réflexivité*, et qui pourrait donc se permettre de faire au sceptique la même réponse que l'externaliste williamsonien :

peut-être que *réflexivement* je ne peux pas savoir que je sais, ou être justifié à croire que je suis justifié, mais cela n'empêche pas ma croyance (que j'ai deux mains) d'être objectivement connue ou justifiée, indépendamment de l'accès réflexif que j'ai ou non à ce statut épistémique.

C'est cette réponse au sceptique qui me semble insatisfaisante, et sur laquelle je vais revenir plus en détail, car la position de Pascal Engel à ce sujet est plus riche et complexe que ce que je viens d'indiquer. Mais malgré

2. L'argument d'Engel qui me semble décisif à ce sujet est présenté dans les p. 120–123 de *Va savoir!* [Engel 2007], où Engel montre que l'acceptation même du principe de clôture révèle une motivation fondamentalement internaliste. L'autre auteur qui, personnellement, a achevé de me convaincre de la nécessité d'associer une réponse internaliste à la réponse externaliste est [Pritchard 2005].

les nuances à apporter, je défendrai qu'Engel ne fait pas suffisamment droit à l'exigence de réflexivité internaliste – je défendrai qu'il nous faut bien trouver une *raison* (internaliste) de croire à la fiabilité de nos facultés épistémiques, et que cette raison ne peut être que le fait que ces facultés sont des facultés *du sens commun*.

2 L'exigence de réflexivité épistémique

Pour bien poser le problème de la réflexivité épistémique, il faut tout d'abord faire clairement la distinction entre croyances de premier ordre, ou irréfléchies (par exemple : *j'ai deux mains*), et croyances de deuxième ordre, ou réfléchies (par exemple : *ma croyance d'avoir deux mains est justifiée / est une connaissance / est certaine*). Ainsi que l'a souligné Alston, il est essentiel de distinguer ces deux niveaux pour éviter toutes sortes de confusions en épistémologie [Alston 1980]. Avant Alston, le premier philosophe (à ma connaissance) à avoir clairement souligné l'importance de cette distinction de niveaux est le cardinal Mercier :

Cette distinction de deux ordres de connaissances, les unes *directes* ou *spontanées*, les autres *réfléchies* et *librement voulues*, est de la plus haute importance dans la question de la certitude. Faute de la faire, on rabaisse le débat à des proportions insignifiantes et l'on discute à côté de la vraie question.

Réfléchir, c'est prendre pour objet de sa connaissance un acte antérieur de la faculté réfléchissante. Examiner la valeur d'une adhésion spontanée, c'est prendre comme objet de connaissance cette adhésion présupposée. La réflexion est donc la condition *sine qua non* de la position du problème critériologique. [Mercier 1911, 36–37]

Engel est parfaitement conscient de cette distinction. Ce qu'il dit (si l'on s'en tient pour l'instant à la reconstruction simplifiée donnée dans la section précédente), c'est que nos croyances de premier ordre peuvent parfaitement constituer des connaissances (externalistes) et même être justifiées (du point de vue internaliste) sans qu'on ait besoin pour cela d'avoir une connaissance ou une justification réflexive, c'est-à-dire sans que la croyance de deuxième ordre soit elle-même connue ou justifiée. Par exemple, ma croyance que « j'ai deux mains » peut très bien être une connaissance même si je ne sais pas que « ma croyance que j'ai deux mains constitue une connaissance ». Ou au niveau de la justification, ma croyance que « j'ai deux mains » peut très bien être justifiée même si je ne suis pas justifié à avoir la croyance de deuxième ordre « ma croyance que j'ai deux mains est justifiée / est une connaissance ». Donc Engel accepte tout à fait cette distinction des niveaux épistémologiques, et il s'appuie même sur cette distinction pour défendre que la justification de premier ordre est possible sans justification réflexive.

Mais Mercier va plus loin que cela. Il ne défend pas seulement l'importance de la distinction des niveaux : il soutient que le problème sceptique à proprement parler se pose fondamentalement au niveau réflexif :

Or, le scepticisme digne de discussion ne va pas jusqu'à tout contester. Pour le sceptique comme pour nous, la nature est en possession d'assentiments certains.

Le problème surgit tout juste au moment où, réfléchissant sur nos états de conscience, nous nous demandons si la réflexion confirme nos adhésions spontanées ; en d'autres mots, si elle peut, oui ou non, trouver justifiable et justifié l'assentiment naturel de l'âme que nous appelons certitude. [Mercier 1911, 37]

Il en tire ainsi sa formulation du « problème critériologique » (ou, plus traditionnellement, du « problème sceptique du critère ») :

Lorsque l'intelligence réfléchit sur ses adhésions spontanées, est-elle en mesure de discerner celles qui sont légitimes de celles qui ne le sont pas, celles qui sont fondées sur la connaissance de la vérité de celles qui n'offrent pas une pareille garantie ?

[Mercier 1911, 38]

Si l'on suit Mercier, le sceptique ne chercherait pas réellement à remettre en cause les croyances de premier ordre ou irréflechies (j'ai deux mains). Ce que le sceptique remet en cause, ce serait seulement les croyances *réflexives* (je sais que j'ai deux mains). Mais dans ce cas, la stratégie d'Engel, qui consiste à défendre le statut épistémique des croyances de premier ordre simplement en concédant qu'une connaissance réflexive est *impossible*, mais en insistant sur le fait qu'elle n'est pas nécessaire à la justification des croyances de premier ordre – cette stratégie, donc, reviendrait tout simplement à *accepter* le scepticisme, tout en faisant contre mauvaise fortune bon cœur. Sur le fond de la question, celui de la possibilité d'une connaissance réfléchie, l'attitude « anti-réflexivité » de Williamson ou d'Engel reviendrait finalement à donner raison au sceptique.

Je ne suis pas sûr d'être d'accord avec Mercier sur le fait que le sceptique cherche *uniquement* à remettre en cause les croyances réflexives, sans toucher aux croyances de premier ordre du type « j'ai deux mains ». Mais je crois qu'il a raison en revanche de souligner qu'une réponse satisfaisante au scepticisme *doit* défendre nos croyances réflexives, parce que sans l'appui de ces croyances réflexives (je sais que j'ai deux mains) nos croyances de premier ordre elles-mêmes perdraient à mon sens leur justification. C'est ce dernier point que conteste la solution anti-réflexivité de Williamson et Engel, et je vais donc le défendre par un argument développé. Je reviendrai ensuite plus en détail sur la question de savoir si Engel refuse totalement un tel principe.

Le principe de réflexivité que je défends peut être formulé, en première approche, de la manière suivante :

(MC1) Si je ne suis pas justifié à croire que je sais que p , alors je ne suis pas justifié à croire que p . $\sim JKp \Rightarrow \sim Jp$

En particulier, si je ne suis pas justifié à croire que mes croyances perceptives constituent des *connaissances*, je ne suis pas justifié à croire ces croyances perceptives elles-mêmes. Si un tel principe est vrai, alors on ne peut pas répondre au sceptique que l'absence de justification réflexive à propos de la perception ne pose aucun problème de justification pour les croyances perceptives de premier ordre : au contraire, le principe énonce qu'un défaut au niveau de la justification réflexive « se transmet » en défaut de justification de premier ordre.

Pourquoi accepter un tel principe ? Pourquoi exiger une forme de réflexivité pour la justification des croyances ? Il me semble que le meilleur argument a été apporté par Michael Huemer dans sa défense du principe qu'il appelle « méta-cohérence » [Huemer 2011, sec. 3]. Huemer commence par montrer la conséquence absurde ou du moins paradoxale à laquelle mène un refus absolu de tout principe de réflexivité. La conséquence paradoxale, c'est qu'il est possible de croire *en même temps et de manière justifiée* « p » et « je ne sais pas que p », et donc également la conjonction :

(PM) p mais je ne sais pas que p .

Huemer fait remarquer que cette croyance (conjonctive) entre dans la catégorie des « paradoxes mooréens » : elle n'est pas logiquement contradictoire et pourtant il semble extrêmement intuitif de dire qu'elle ne peut être crue de manière rationnelle par un sujet. Moore lui-même, dont les exemples les plus célèbres utilisent plutôt la notion de croyance (p mais je ne crois pas que p), a également envisagé ce paradoxe qui engage la notion de savoir :

Les chiens aboient, mais je ne sais pas qu'ils aboient. [Moore 1962, 277, traduction de l'auteur]

Si l'on admet que les croyances de forme (PM) sont paradoxales et irrationnelles, encore faut-il interpréter *pourquoi* elles le sont. Et c'est cette interprétation qui mène au principe de méta-cohérence. Pour cela, Huemer doit d'abord montrer l'insuffisance des interprétations du paradoxe de Moore qui attribuent l'origine du problème simplement à des normes de l'*assertion* – c'est essentiellement l'interprétation de [Williamson 2000], pour qui la connaissance est la norme de l'assertion. C'est-à-dire que pour pouvoir affirmer p , je serais « tenu » de savoir que p , et donc mon assertion de p contiendrait (pour des raisons conversationnelles) une implicature selon laquelle je sais que p . Huemer fait légitimement observer que ce type d'interprétation ne permet pas de rendre compte du caractère tout aussi paradoxal de la simple *croyance* « p mais je ne sais pas que p », lorsque celle-ci est entretenue silencieusement dans mon esprit et non prononcée.

Pourquoi donc la simple croyance de forme (PM) est-elle paradoxale ou irrationnelle ? Huemer suggère que l'interprétation la plus plausible est un principe de défaiseur épistémique :

(Défaiseur MC) Lorsque S croit que p de manière *prima facie* justifiée à t , mais vient à prendre conscience à $t + 1$ qu'il *ne sait*

pas que p , sa justification *prima facie* pour p est défaite à $t + 1$ par cette information.

On peut donner des exemples très intuitifs de ce principe. Imaginons que je croie que Tibère aimait les lentilles sur le témoignage de la *Vie des douze Césars* de Suétone. Mais j'apprends ensuite que Suétone n'est pas du tout fiable et se plaisait à inventer toutes sortes de détails. Si je continue à croire que Tibère aimait les lentilles, il n'est pas prouvé que ma croyance soit positivement *fausse*, mais étant donné ce que je sais de ma source épistémique, je sais que ce serait un pur hasard que ma croyance soit vraie ; autrement dit, je sais que ma croyance n'est pas une *connaissance*. Je peux garder ma croyance, mais ce ne serait pas justifié : ma justification est défaite.

Ce type de défaiseur de justification a été appelé par [Pollock & Cruz 1999] *undercutting defeater*³ par opposition au *rebutting defeater*. Je parlerai personnellement de défaiseurs fondationnels et de défaiseurs directs. Un défaiseur direct serait par exemple le fait de trouver une lettre authentique de Tibère affirmant qu'il n'aime pas les lentilles. Typiquement, un défaiseur direct vous donne une raison positive d'accepter le contraire de ce que vous étiez initialement justifié à croire, tandis qu'un défaiseur fondationnel vient seulement « couper l'herbe sous le pied » de votre justification antérieure, de telle sorte que vous devriez désormais retenir tout jugement sur la question.

Si donc on accepte qu'une croyance réflexive du type « ma croyance que p n'est pas une connaissance » constitue un défaiseur (fondationnel) de la justification pour p , alors on doit admettre le principe de méta-cohérence suivant :

(MC2) Lorsque S croit que p et réfléchit consciemment à la question de savoir s'il sait que p , alors il ne doit pas être disposé à croire qu'il ne sait pas que p , autrement sa justification initiale à croire que p est défaite.

Sous cette forme, il me semble que le principe de méta-cohérence est extrêmement difficile à refuser (nous verrons d'ailleurs qu'Engel se garde en fait de le refuser).

Mais la version du principe qui nous intéresse ici est en fait une version plus forte, que j'appellerai « méta-cohérence positive », par opposition à (MC2) qui représente seulement une « méta-cohérence négative ». Le principe (MC2) exige simplement (pour rester justifié à croire que p) que je *n'aie pas* la croyance réflexive de « ne pas savoir que p ». Autrement dit, le principe (MC2) laisse encore la possibilité suivante lorsque je réfléchis consciemment à la question de savoir si je sais que p : il se pourrait que je n'aie aucune idée (aucune justification) sur cette question de deuxième ordre. Je crois que p , on me demande si je *sais* que p , et je prends conscience que « je ne sais pas si je le sais ou non ». Que se passe-t-il dans ce scénario où je *doute* simplement de savoir, sans pour autant croire que je ne sais pas ? Huemer soutient que ce simple *doute* constitue encore un défaiseur fondationnel de justification.

3. On trouve également dans la littérature *undermining defeater*.

Prenons un exemple : supposons que la connaissance historique d'un certain événement de la vie d'Henri IV (disons sa liaison avec Fleurette de Nérac) soit fondée sur une lettre signée de la main du roi lui-même. Mais un jour, un historien vient à montrer que l'authenticité de cette lettre peut être mise en doute ; il se *pourrait* qu'elle soit de la main d'un faussaire et ennemi du roi. Si la lettre est de la main du roi, alors la croyance dans sa liaison avec Fleurette de Nérac est une *connaissance* parce qu'elle repose sur une source fiable ; mais si elle est du faussaire, alors elle n'est pas une connaissance. Mais comme nous ne *savons pas* si nous savons ou non, il semble que la seule attitude justifiée soit d'abandonner notre croyance qu'Henri IV a eu une liaison avec Fleurette de Nérac. Il semble *injustifié* de maintenir cette croyance tout en ayant à l'esprit un doute sur le fait qu'elle soit ou non une connaissance. Si l'on accepte que le *doute réflexif* (« peut-être que je ne sais pas que p ») constitue également un défaiseur fondationnel, au même titre que la croyance effective que « je ne sais pas que p », alors il semble que l'agent épistémique qui se demande explicitement si telle croyance est ou non une connaissance n'a plus d'autre issue que de croire *positivement* que sa croyance est une connaissance, faute de quoi il perd sa justification *prima facie*. Autrement dit :

(MC3) Lorsque S croit que p et réfléchit consciemment à la question de savoir s'il sait que p , alors il doit disposer d'une justification à croire qu'il sait que p , autrement sa justification à croire que p est défaite⁴.

Ce principe de méta-cohérence positive est peut-être légèrement moins fort que le principe de méta-cohérence négative, mais il me semble néanmoins extrêmement difficile à contester.

En tout cas, dans la formulation que je viens de donner, ce principe de méta-cohérence échappe aux objections les plus courantes qui sont faites aux principes de réflexivité, à savoir l'objection de l'enfant et l'objection de la régression à l'infini. L'objection de l'enfant consiste à dire que, intuitivement, il semble tout à fait possible à un enfant d'avoir des croyances justifiées avant même que celui-ci ne maîtrise les concepts méta-épistémiques qui lui permettraient d'avoir une justification méta-épistémique explicite ; il est donc possible d'être justifié à croire que p sans avoir une justification méta-épistémique. Le principe (MC3) ne conteste pas la conclusion de cet argument : il est en effet possible à quelqu'un qui *ne réfléchit pas* à la question méta-épistémique (parce qu'il ne le peut pas) d'être justifié dans ses croyances, mais quelqu'un *qui y réfléchit* doit être en mesure d'y apporter une réponse positive pour préserver sa justification. L'objection de la régression à l'infini consiste à dire que si nous devons avoir pour chaque croyance p la croyance Kp , alors il nous faudrait aussi avoir la croyance KKp , la croyance $KKKp$, etc., à l'infini, or notre cerveau n'est pas en mesure de contenir cette infinité de croyances. Là encore,

4. Cette formulation est équivalente à celle donnée par [Huemer 2011, sec. 1] : « Le fait de croire catégoriquement que p nous engage, après réflexion, à accepter que nous savons que p . »

le principe (MC3) ne conteste pas la conclusion : nous n'avons évidemment pas cette infinité de croyances ; ce que dit (MC3), c'est que pour chaque degré où la question méta-épistémique est explicitement posée, nous devons être *disposés* à acquérir une justification (et une croyance) méta-épistémique à ce degré.

Je considérerai donc désormais que le principe de méta-cohérence positive est un principe épistémologique extrêmement fort.

3 L'exigence de réflexivité appliquée au problème sceptique

Voyons à présent comment la méta-cohérence s'applique au problème du scepticisme, et comment Engel en traite.

Ainsi que je l'ai montré dans [Guillon 2016], le principe de méta-cohérence introduit clairement dans la discussion sur le scepticisme une exigence de réflexivité qui va à l'encontre de la stratégie williamsonienne : pour être justifié à croire que j'ai deux mains, je dois être justifié à croire que je *sais* que j'ai deux mains. Autrement dit, je dois être justifié à croire que la perception est une source épistémique fiable. Il semble donc que le principe de méta-cohérence s'oppose à la stratégie « néo-mooréenne » que veut adopter Engel, lorsqu'il refuse à la fois la réflexivité de la connaissance (le principe *KK*) et la réflexivité de la justification (grâce à la justification *prima facie*). On pourrait s'étonner d'une telle stratégie pour plusieurs raisons.

On pourrait tout d'abord faire observer qu'une telle stratégie anti-réflexivité n'est, en fait, pas si « mooréenne » que cela : Moore lui-même, lorsqu'on se penche attentivement sur ce qu'il dit, ne refuse pas toute forme de réflexivité épistémique. En effet, à la question de savoir si les croyances de sens commun constituent des *connaissances*, la réponse de Moore ne consiste pas à rester dans le doute sur leur statut épistémique ; il *affirme* qu'elles sont des connaissances, et qu'il sait qu'elles le sont :

Mais est-ce que je *sais* réellement que toutes les propositions [de sens commun] sont vraies ? N'est-il pas possible que je les croie simplement ? Ou que je sache qu'elles sont très probables ? [...] En réponse à cette question, je crois que je n'ai rien de mieux à dire que le fait qu'il me semble que je les connais en effet, et ce avec certitude. Il est, certes, évident que pour la plupart d'entre elles je ne les connais pas *directement* : c'est-à-dire que je ne les connais que parce que, dans le passé, j'ai connu comme vraies d'*autres* propositions qui servaient d'argument en leur faveur. Si, par exemple, je sais que la terre a existé de nombreuses années avant ma naissance, je ne sais cette chose, assurément, que parce que j'ai su d'autres choses dans le passé qui servaient d'argument en sa faveur. Et je ne sais certainement pas exactement quels

étaient les arguments. Pourtant, il me semble que tout cela n'est pas une bonne raison de douter que je le sais. Nous sommes tous, à mon sens, dans cette étrange position où nous *savons* de nombreuses choses, à l'égard desquelles nous *savons* en outre que nous avons dû avoir des arguments en leur faveur, et pourtant nous ne savons pas *comment* nous les savons, i.e. nous ne savons pas quels étaient les arguments. [Moore 1993, traduction de l'auteur]

Que dit Moore, exactement ? Pourquoi peut-il sembler refuser une forme de réflexivité épistémique ? Parce qu'il soutient que nous ne savons pas *comment* nos croyances de sens commun constituent des connaissances. Nous savons sans savoir *comment* nous savons. Le principe suivant est donc faux :

(KHK) Si je sais, alors je sais *comment* je sais.

Mais ce qu'il ne dit pas, c'est que nous savons sans savoir *que* nous savons : au contraire, pour Moore, nous savons que nous savons (ou en tout cas nous sommes justifiés à croire que nous savons) ; nous avons ce privilège réflexif qui nous permet de satisfaire au principe de méta-cohérence. Autrement dit, Moore ne dit rien ici contre le principe

(KK) Si je sais, alors je sais *que* je sais.

Et, *a fortiori*, il ne dit rien contre le principe de méta-cohérence tel que je l'ai exprimé plus haut. Donc je serais très réticent à considérer comme « néo-mooréenne » une solution au problème sceptique qui consiste à refuser toute forme de réflexivité. L'épistémologie d'Engel est néo-mooréenne lorsqu'elle accepte le principe de clôture, mais pas lorsqu'elle rejette le principe de réflexivité.

Par ailleurs, indépendamment de la question, somme toute historique, de la fidélité à Moore, on peut remarquer qu'il y a une certaine incongruité, ou tout du moins une dissymétrie, à accepter le principe de clôture tout en refusant le principe de réflexivité. Les deux principes en effet sont assez voisins d'un point de vue formel. En effet, l'un comme l'autre consistent à dire que la justification pour une croyance p suppose, pour n'être pas défaite, la justification d'une autre croyance dans notre « espace épistémique » : le principe de clôture exige la justification d'une autre croyance q qui est une conséquence logique (connue) de p , tandis que le principe de méta-cohérence exige la justification d'une croyance de deuxième ordre Kp . Il n'est certes pas *logiquement contradictoire* d'accepter une exigence tout en refusant l'autre, mais il semblerait formellement plus élégant d'accepter la seconde si on accepte la première. En tout cas, il semble que le refus de l'une et de l'autre entraîne le même genre de « paradoxes ». Si l'on refuse la clôture, on se trouve confronté à la possibilité paradoxale d'un agent « inconséquent », qui tel la tortue de Lewis Carroll, est prêt à affirmer le paradoxe suivant : « oui, j'affirme que p , et je vois bien que p implique q , simplement je me refuse à soutenir que q » [Hawthorne 2005, 32]. Si l'on refuse la méta-cohérence, on se trouve confronté à la possibilité paradoxale d'un agent « incohérent », qui est prêt à affirmer « oui, j'affirme que p , seulement je n'ai pas la moindre idée si ma croyance

que p est une connaissance ou non ; pour autant que je sache, il se pourrait aussi bien qu'elle soit absolument hasardeuse ». Il me semble que le genre de motivations rationnelles qui peut amener quelqu'un à rejeter la rationalité de « l'agent inconséquent » devrait aussi amener à rejeter la rationalité de « l'agent incohérent ».

Pour toutes ces raisons, il me semble que le rejet de la réflexivité par Engel est un élément de dissymétrie qui affaiblit sa position néo-mooréenne. Mais Engel rejette-t-il réellement le principe de méta-cohérence (et donc la réflexivité) ? Quand on examine le détail de sa solution, ce n'est pas si clair.

Le fait qu'Engel soit en fait réticent à rejeter absolument la contrainte de méta-cohérence peut se voir tout d'abord dans son traitement de l'épistémologie du témoignage [Engel 2007, 163–174]. Dans ce domaine, l'opposition entre justification réflexive et justification *prima facie* est traditionnellement présentée sous la forme du débat entre réductionnisme et principe de crédulité. D'après le réductionniste, pour qu'une croyance testimoniale soit justifiée, il faut qu'elle ait été formée *sur la base* d'une connaissance préalable que « le témoin est fiable » ; cette croyance méta-épistémique préalable permet ensuite (en conjonction avec une croyance factuelle du type « le témoin a dit que p ») d'*inférer* que p . Le témoignage est alors réduit à une inférence à partir de deux connaissances observationnelles directes (j'ai pu observer par le passé la fiabilité du témoin, et j'ai entendu ses paroles). À l'inverse, le tenant du principe de crédulité souligne que, dans la réalité, nous ne procédons pas à ce genre « d'inférence » dans laquelle la parole d'autrui est réduite à une sorte de « jauge de vérité » plus ou moins fiable ; dans la réalité, il semble que nous acceptons ce qui est affirmé par les témoins de manière basique, non inférentielle, et en *déléguant* les raisons à notre témoin (plutôt qu'en ayant nos propres raisons égotistes consistant dans l'observation en première personne de la fiabilité du témoin-jauge). Dans ce débat, on aurait pu s'attendre à ce qu'Engel défende la position anti-réductionniste, qui se fonde sur une notion de justification *prima facie* ou d'autorisation épistémique. Mais en fait, Engel préfère rejeter l'alternative, et garder des éléments de l'une et de l'autre positions. De l'anti-réductionniste, il conserve le fait qu'une croyance testimoniale n'est pas une inférence sur la base de raisons égotistes : *mes* raisons de croire que p , ce sont les raisons du témoin... ce ne sont pas mes observations préalables « que le témoin est fiable ». Mais cela veut-il dire que je peux me dispenser totalement de croire que le témoin est fiable lorsque j'accepte son témoignage ? Non, et c'est là que le réductionnisme a une part de vérité : pour Engel, nous *avons*, et nous *devons* avoir, des « raisons de croire que la source est fiable », et ces raisons ne sont pas « indépendantes de toute donnée probante », au contraire, elles « reposent sur des données et des croyances d'*arrière-plan* » [Engel 2007, 173]. Le principe de méta-cohérence est donc satisfait : lorsqu'un agent épistémique croit que p sur la base du témoignage de S , il a toujours à sa disposition des raisons qui lui permettront de répondre positivement si on venait à l'interroger sur la fiabilité de sa source épistémique S (bien qu'il n'ait pas nécessairement *préalablement*

pensé à ces raisons et à cette question méta-épistémique au moment où il a formé sa croyance testimoniale). Autrement dit, lorsqu'il est justifié à croire tel témoignage, il a aussi à sa disposition des raisons qui lui permettront (si besoin) de former la croyance que ce témoignage est une source de connaissance.

Dans le cas du témoignage, on voit que ce qui permet de résoudre la contrainte de méta-cohérence, c'est finalement le recours à un aspect du réductionnisme, à savoir le fait de faire appel à une *autre* source épistémique (la perception et la mémoire contextuelles) pour authentifier tel ou tel témoin comme source fiable de connaissance. C'est évidemment la solution la plus classique pour fournir une méta-justification du témoignage en général. Mais ce type de solution, s'il marche bien pour le témoignage, ne s'extrapole pas aisément pour la question de la méta-justification de la perception : comment sais-je que Paul est fiable ? Parce que j'ai *vu* moi-même la vérité de plusieurs de ses affirmations. Comment sais-je que ma vue est fiable ? Parce que j'ai vu la vérité de ses affirmations ? Cette réponse ne va pas ; elle présente à l'évidence une forme de circularité qui pose problème.

Comment Engel traite-t-il donc de la méta-cohérence dans le cas de nos sources perceptives de connaissance ? On peut voir que là encore Engel évite de violer le principe de méta-cohérence, mais il le satisfait d'une manière qui me semble problématique.

La défense de la justification perceptive *prima facie* commence, dans le texte d'Engel, de manière comparable à la défense anti-réductionniste du témoignage : de même que le témoignage n'est pas une inférence à partir des croyances préalables « tel témoin a dit que *p* » et « tel témoin est fiable », de même la croyance perceptive n'est pas non plus une inférence à partir des croyances préalables « j'ai une impression visuelle de table » et « mes impressions visuelles sont fiables ». Je n'ai donc pas besoin d'une méta-justification (ou justification de second ordre) *antécédente* : mes impressions visuelles peuvent conférer à mes croyances une justification *prima facie* sans que j'aie antécédemment une justification réflexive. C'est ici qu'apparaît le thème anti-réflexivité [Engel 2007, 192]. Notez que jusqu'ici, on n'a pas enfreint le principe de méta-cohérence tel que je l'ai énoncé plus haut, car ce dernier n'exige de méta-justification *qu'au* moment où l'agent *prend conscience* de la question de second ordre : le philosophe méta-cohérent admet tout à fait que telle impression, à elle seule, donne une justification *prima facie*, même sans méta-justification, mais il ajoute que cette justification *prima facie* sera *défaite* (par un défaiseur fondationnel) si à la réflexion il m'apparaît que je n'ai aucune raison de me fier à cette impression. Or, la difficulté de l'argument sceptique contre la perception c'est qu'il nous invite précisément à adopter le point de vue réflexif et à nous demander explicitement si nos impressions perceptives sont fiables ou non. Face à cette question explicite, le philosophe méta-cohérent soutient qu'il *faut bien* que nous soyons en mesure de produire une méta-justification de la perception (même si *initialement*, ou *pré-réflexivement*, nous n'en avons pas besoin pour justifier nos croyances de premier ordre).

Donc la question qui intéresse le philosophe méta-cohérent, c'est celle de savoir si Engel exige ou non une méta-justification *postérieure*, *post-réflexion* (et non pas *antécédente*). Or, on peut remarquer que, contrairement à l'objection que j'ai avancée plus haut, Engel semble adopter ici une réponse tout à fait parallèle à sa réponse pour le problème de clôture. Que dit en effet Engel lorsqu'il accepte le principe de clôture? Exige-t-il de nous que nous ayons une justification *antécédente* pour tous les scénarios sceptiques logiquement incompatibles avec le fait d'avoir deux mains (CUC, malin génie, etc.)? Non, et il a tout à fait raison de ne pas exiger cela! Ce qu'exige Engel, de manière très légitime, c'est seulement que nous soyons en mesure *lorsque le scénario nous est explicitement opposé* d'acquérir la croyance justifiée que ce scénario est faux. D'où pourrait donc provenir ma justification à croire que je ne suis pas un CUC, ou une pure âme manipulée? Ici, Engel adopte en fait la réponse « déductiviste » de Peter Klein [Klein 1995] : ma justification à croire la négation des hypothèses sceptiques $\sim HS$, c'est que j'ai une croyance qui implique logiquement leur fausseté, par exemple la croyance (antécédemment justifiée) que j'ai deux mains! Je peux donc *déduire* que je ne suis pas un CUC du fait que j'ai deux mains. Dans cette solution, la justification pour $\sim HS$ est non seulement *chronologiquement postérieure* à la justification (*prima facie*) pour « j'ai deux mains », mais elle est aussi *logiquement dépendante* de cette justification *prima facie*. Ainsi, lorsque le sceptique m'oppose un scénario du type cerveaux dans une cuve, ce scénario ne représente pas pour moi un défaiseur direct, car j'ai les moyens de rejeter ce scénario, par simple déduction à partir de ma croyance (justifiée) d'avoir deux mains. Voilà pour la justification des propositions anti-sceptiques $\sim HS$. Qu'en est-il alors de la justification des propositions méta-épistémiques de type Kp , ou « telle source épistémique est fiable »? Avons-nous une justification *postérieure* à les croire (pour ne pas enfreindre le principe de méta-cohérence)? Engel ne distingue pas explicitement ces propositions méta-épistémiques des propositions anti-sceptiques, mais il est tout à fait possible d'interpréter son texte comme signifiant que nous *avons* une justification postérieure pour ces propositions réflexives (bien que nous n'en ayons pas de justification *antécédente*). Voici en effet le principe que propose Engel pour la justification *prima facie* en général :

(iii) Quand on est justifié *prima facie*, on est *par là même* justifié à croire que l'hypothèse qui déferait p est fausse.

[Engel 2007, 193]

Ce principe s'applique, comme on vient de le voir, aux éventuels défaiseurs directs que sont les hypothèses sceptiques radicales (CUC, malins génies, etc.). Nous sommes justifiés à les rejeter *sur la base* (déductive) de notre justification *prima facie* à croire que « nous avons deux mains » (ou toute autre proposition perceptive anodine). Mais si l'on admet que la proposition « ma vue n'est pas fiable » constituerait *aussi* un défaiseur ainsi que le soutient le principe de méta-cohérence (cette fois un défaiseur fondationnel et non plus direct), alors le principe énoncé par Engel implique une réponse similaire : chaque fois que je suis justifié *prima facie* à croire que p (sur la base de la vue, par exemple),

je suis *par là même* justifié à croire Kp (c'est-à-dire à croire que ma vue est fiable). Ainsi, bien que nous n'ayons pas de méta-justification *antécédente*, nous aurions toujours une méta-justification *résultante* de notre justification *prima facie*. Ceci n'est qu'une interprétation ; il n'est pas impossible qu'Engel souhaite en réalité rejeter le principe de méta-cohérence, mais le fait qu'il insiste tant sur l'absence de méta-justification *antécédente* et *indépendante* semble suggérer qu'il ne s'oppose pas au réquisit d'une méta-justification *postérieure* et *résultante*.

Malheureusement, il me semble que cette manière de satisfaire au principe de méta-cohérence (si c'est bien l'idée d'Engel) pose de sérieux problèmes ; et ce que je défendrai, c'est qu'il nous faut bien une justification *indépendante* (bien que non *antécédente*), ou en tout cas pas une justification résultante sur le modèle du principe (iii) que l'on vient de mentionner. À mon sens, ce principe de justification résultante fonctionne bien lorsqu'il s'agit de déduire des croyances anti-sceptiques de type $\sim HS$, selon le modèle déductiviste de Peter Klein. On voit bien en effet comment la croyance (justifiée) d'avoir deux mains peut justifier (par déduction) la croyance de ne pas être un pur cerveau ou une pure âme (puisqu'alors je n'aurais pas de mains). Mais comment est-ce que la croyance (certes *prima facie* justifiée) d'avoir deux mains pourrait-elle me justifier à croire que « ma vue est fiable » ? Ici, aucune déduction (ni même aucune induction) n'est possible : je pourrais très bien avoir deux mains et que ma vue ne soit pas fiable ; le fait d'avoir deux mains n'augmente même pas la probabilité que ma vue soit fiable. Les deux faits sont totalement distincts et indépendants l'un de l'autre. En étant justifié *prima facie* à avoir une croyance perceptive, je ne vois pas comment je pourrais être « *par là même* » (ou de manière résultante) justifié à croire que ma perception est fiable. Peut-être suis-je *en même temps* justifié à croire que ma perception est fiable, à chaque fois que je suis justifié par la perception à croire un certain contenu, mais même si la méta-justification arrivait toujours *en même temps* que la justification de premier ordre, il me semble qu'elle devrait néanmoins, d'un point de vue logique, reposer sur un *autre motif* que la justification de premier ordre elle-même. Elle doit reposer sur un autre motif de justification tout simplement parce que, encore une fois, elle porte sur un fait métaphysiquement indépendant du fait que justifie (au premier ordre) la perception. C'est la vue qui me justifie immédiatement à croire que j'ai deux mains ; mais ce n'est pas la vue qui me justifie immédiatement à croire que la vue est fiable. Il faut bien ici une autre source de justification.

Ma conclusion, à ce stade, est donc qu'Engel n'a pas donné une théorie satisfaisante de la méta-justification des croyances perceptives. Si le principe de méta-cohérence est vrai, alors il nous *faut* une théorie de la manière dont nous sommes justifiés à tenir nos perceptions pour fiables ; et la théorie selon laquelle la méta-justification résulterait directement de la justification *prima facie* aux croyances perceptives ne marche pas.

4 Méta-justification par le sens commun

À mon sens, le recours chez Reid à une « théorie forte » du sens commun vient précisément d'une tentative pour répondre au problème de la méta-justification (ou au « problème du critère ») tel que nous l'avons posé dans la section précédente : pour défendre que notre perception est méta-justifiée (que nous sommes justifiés à la tenir pour fiable), il nous faut faire explicitement appel au *motif* qu'il s'agit d'une faculté du sens commun. Ainsi, si Engel ne propose pas de défense forte du sens commun, ce serait précisément parce qu'il n'a pas suffisamment identifié la contrainte de méta-cohérence et le fait que pour la satisfaire il nous faut faire appel à d'autres motifs que le simple contenu des perceptions elles-mêmes.

Dans cette dernière section, je proposerai rapidement une théorie forte du sens commun, inspirée de Reid, mais aussi d'autres auteurs moins connus dans la tradition « latine » (ou post-Buffier) du sens commun. L'objectif de cette section est de montrer comment l'appel explicite au sens commun, en tant que tel, permet de répondre à l'exigence de méta-cohérence.

Lorsqu'on essaie de formuler une théorie forte du sens commun, on est immédiatement confronté au fait qu'il existe une multitude de façons d'interpréter une telle théorie – lesquelles interprétations se trouvent pour la plupart représentées chez Reid lui-même, qui n'a pas cherché à les distinguer et à choisir entre elles. Il y a essentiellement deux questions qui posent problème :

- (i) Quelles croyances sont « de sens commun » ? Celles qui sont *naturelles* ? Celles qui sont *communes* ? Celles qui sont *inévitables* ?
- (ii) En quoi le fait d'être « de sens commun » nous donne-t-il une raison de nous fier à telles et telles facultés (et donc un argument contre le sceptique) ? Est-ce pour des raisons *pragmatiques* ? Est-ce parce qu'elles sont à l'abri de tout débat puisque tout le monde les accepte ? Ou derechef parce qu'elles sont inévitables ?

Je commencerai par écarter deux réponses à la question (ii) qui sont explicitement présentes chez Reid.

La première réponse est la réponse « par le rire », que Reid reprend à un texte de Fénelon : lorsque le sceptique met en cause les croyances de sens commun, je n'ai pas à lui répondre autrement que par le rire, ce qui est la réponse appropriée face à quelqu'un qui n'est pas « dans son bon sens » ou « en bonne santé ». Je n'ai pas à entrer en dialogue avec le sceptique parce qu'il a perdu le sens. Dans cette interprétation, le recours au sens commun n'est pas une manière de *donner une justification* pour la confiance que je mets dans mes facultés ; c'est au contraire une manière de montrer que *je n'ai pas à donner de justification*. Par conséquent, c'est une interprétation qui revient à rejeter (ou en tout cas à se détourner de) l'exigence de méta-cohérence. Ce n'est donc pas ce que nous cherchons ici.

La deuxième réponse est présentée par Reid lorsqu'il se pose à lui-même la question des « raisons » pour lesquelles, contrairement au sceptique, il « n'essayera jamais de rejeter [la raison et la perception] » [Reid 1764, VI, 20]. Après avoir rappelé que « la partie saine de l'humanité ne sera pas très inquiète de connaître [ses] raisons », il consent cependant à donner trois éléments de réponse explicite. Je reviendrai un peu plus loin sur la première réponse qu'il donne. Les deuxième et troisième réponses consistent en fait en une réponse pragmatiste : je suis justifié à croire mes sens parce que cela amène à toutes sortes de succès. D'une part (c'est la raison 2), si mes perceptions sont des illusions, elles sont néanmoins des illusions que *la nature* a voulu me donner, et donc je ne dois pas m'attendre (*a priori*) à beaucoup de succès pratiques si j'entreprends de m'opposer aux projets de la nature. D'autre part (c'est la raison 3), j'ai pu *observer a posteriori* que le fait de me fier aux sens n'apportait que des succès pratiques. Cette réponse pragmatiste ne me satisfait pas non plus, car à mon sens les raisons pragmatiques sont au mieux des raisons pratiques d'*accepter* certaines propositions (de « faire comme si ») mais ce ne sont pas des raisons épistémiques de *croire* ces propositions. Or il me semble que le principe de méta-cohérence nous oblige à trouver des raisons épistémiques [*truth-directed*] de croire que notre perception est fiable.

Venons-en maintenant à la première « raison » que donne Reid pour se fier à la perception. Cette réponse célèbre rejoint la question (i), à savoir quelles croyances comptent comme étant « de sens commun ». Pourquoi Reid ne met-il pas en doute la perception ?

Premièrement, parce que ce n'est pas en mon pouvoir : pourquoi donc ferais-je une tentative vaine ? [Reid 1764, IV, 20, traduction de l'auteur]

C'est la fameuse réponse d'*inévitabilité* : les croyances de sens commun sont celles qui sont pour nous inévitables, et *pour cette raison* il est déraisonnable de chercher à les renverser (puisqu'on n'y parviendra pas). Cette réponse ne me paraît pas vraiment satisfaisante non plus. La première difficulté, c'est que si les croyances de sens commun étaient si inévitables que cela, il ne devrait pas exister de sceptiques du tout ; donc il ne semble pas que ces croyances soient *absolument* inévitables. La deuxième difficulté, c'est que l'inévitabilité ne semble pas donner une raison *épistémique* en faveur de la fiabilité de nos sens. En somme, ce serait encore un motif de type pragmatiste (quoiqu'assez différent des deux autres). Je rejeterai donc l'idée d'inévitabilité comme réponse à la question (ii), c'est-à-dire comme *raison* de tenir nos sens pour fiables ; mais je ne la rejeterai pas totalement pour répondre à la question (i). Et c'est même en partant du problème de l'inévitabilité que je vais apporter ma propre réponse à la question (i).

La première question, donc, d'une théorie (forte) du sens commun, est celle de savoir quelles croyances comptent comme étant « de sens commun ». Et il y a principalement trois réponses possibles : les croyances *naturelles* (en latin, les auteurs parlent d'ailleurs de *sensus naturae communis*), les croyances

communes, et les croyances *inévitables*. À mon sens, ces trois caractérisations doivent être retenues et sont importantes, mais elles n'ont pas le même rôle dans « l'ordre des raisons ». La plus importante des trois, celle qui constituera la *raison épistémique* qui nous justifie à tenir la perception pour fiable, sera la caractéristique de naturalité. Mais les deux autres lui seront liées, comme conséquence et comme symptôme indicateur.

Repartons de la caractéristique problématique d'inévitabilité. On vient de dire, en objection à Reid, que les croyances de sens commun ne devaient pas être si inévitables que cela puisqu'elles sont mises en doute par les sceptiques. Comment se tirer de cette difficulté ? Sont-elles inévitables ou pas ? Je crois que le cardinal Mercier a apporté une réponse tout à fait satisfaisante à cette question :

Spontanément, c'est-à-dire sous l'influence exclusive de l'objet à connaître et indépendamment de toute intervention de la volonté libre, l'intelligence a des assentiments déterminés.

Mais la volonté libre a le pouvoir d'attacher l'attention de l'intelligence à la considération de ces adhésions spontanées, à l'effet de voir quelle est la cause de leur fermeté. [Mercier 1911, 36]

Autrement dit, les « adhésions » aux croyances de sens commun sont inévitables (« spontanées ») *tant que nous ne nous posons pas la question réflexive de leur fiabilité*. Mais l'esprit humain a toujours la liberté d'adopter un point de vue réflexif sur ses croyances, et *par là même* de les rendre « résistibles ». Le sceptique sait bien cela : il sait que s'il veut atteindre le doute, il lui faut adopter une attitude réflexive (c'était le *but* même des tropes de Sextus Empiricus que d'entretenir « artificiellement » un doute qui, sans ces secours de la réflexion, retomberait immanquablement dans un assentiment au sens commun). Mais étant donné que le doute est, malgré tout, possible dans ces circonstances réflexives, on ne peut pas tenir les croyances de sens commun pour *absolument* inévitables, et c'est pourquoi il nous faut bien les « motiver » ou les « méta-justifier ».

Ce caractère d'inévitabilité n'est donc que conditionnel ; mais tout conditionnel qu'il soit, il a à mon sens une grande importance, parce que cette inévitabilité peut être perçue (ou « introspectée ») de manière purement internaliste. Je « sens bien » que mon doute sur mes impressions visuelles est « artificiel », et que si je n'y faisais pas réflexion, je croirais tout simplement et inévitablement ce que mes yeux me disent. De cette inévitabilité (conditionnelle) que je sens en moi-même, je peux inférer que la vue fait partie de ma *nature* d'agent épistémique. Je peux inférer que la vue (ou la perception en général) est une faculté *naturelle*. L'inévitabilité (conditionnelle) peut donc servir de *ratio cognoscendi* de la naturalité de nos facultés. Ce raisonnement n'est pas présent chez Reid : pour Reid, c'est une tâche empirique très complexe de déterminer si tel ou tel principe fait partie ou non de notre nature ; cela présuppose des observations extérieures sur les autres hommes

et leur manière de raisonner, sans doute aussi des réflexions sur le langage et ses structures profondes par opposition à ses structures superficielles, etc., [Yaffe & Nichols 2000, sec. 1 : Common sense and ordinary language]. La difficulté de cette théorie empiriste de la naturalité, c'est qu'elle s'appuie évidemment sur les sens : c'est par les sens que j'observe les autres hommes et leur langage. Si donc je m'appuie sur les sens pour déterminer que « les sens sont naturels et donc fiables », on est conduit à une forme de circularité qui semble problématique. À l'inverse, si l'on peut soutenir que l'homme peut connaître le caractère *naturel* de ses impressions sensibles sans s'appuyer sur une observation extérieure de la nature, mais simplement en percevant *en lui-même* l'inévitabilité (conditionnelle) d'adhérer à ces impressions, alors on évite cette forme de circularité. Il y aurait plusieurs objections à faire à cette théorie de l'introspection immédiate du caractère naturel des facultés, mais dans ce bref aperçu d'une théorie du sens commun, je m'appliquerai surtout à donner l'idée d'ensemble de la théorie.

L'agent épistémique perçoit donc en lui-même que certaines adhésions sont inévitables à moins d'y faire réflexion, et de ce fait observé, il déduit que ces impressions (ou les facultés correspondantes) font partie de sa *nature* même d'agent épistémique. Cela signifie-t-il que ces croyances de sens commun (ou les principes sur lesquels elles reposent) sont *communes*, ou universelles ? Janos Imre, un auteur hongrois du début du XIX^e siècle⁵, a défendu que l'on pouvait inférer le caractère universel du sens commun précisément du fait qu'il est *naturel* :

Pour prouver l'existence du sens commun naturel, il n'est pas nécessaire d'examiner les inclinations de tous les hommes un par un, et de conclure ainsi par une induction parfaite, ce qui est absolument impossible ; mais il me suffit de savoir par le témoignage du sens interne que ce sens est de nature ; en effet, une fois que je connais cela, j'infère légitimement que le même sens se rencontre chez les autres qui sont dotés de la même nature que moi ; c'est ainsi en effet que j'attribue à bon droit aux autres hommes le *désir du vrai bonheur*, dont je fais l'expérience en moi-même ; c'est ainsi également, vu que je fais l'expérience en moi-même du sens commun naturel, que j'en infère à juste titre qu'il est présent à l'identique chez les autres hommes, parce que les autres jouissent de la même nature rationnelle que moi. [Imre 1819, 145, traduction de l'auteur]

Cette théorie me semble intéressante en ce qu'elle renverse l'idée reidienne ou empiriste en général selon laquelle l'universalité empiriquement observée serait la *ratio cognoscendi* du caractère naturel de telle ou telle faculté. Au contraire, c'est le caractère naturel du sens commun qui devient la *ratio*

5. Il se situe dans la tradition latine qui, après Buffier, a développé dans des ouvrages universitaires, parfois des sommes pédagogiques, la théorie du *sensus naturae communis* comme source épistémique à part entière.

cognoscendi de son caractère universel. Je ne cherche pas à rejeter totalement l'idée qu'une observation empirique des différentes populations (et de leur structures linguistiques) puisse nous apprendre quelque chose sur le caractère plus ou moins profond de telle croyance ou de telle faculté, mais il me semble important (à cause du problème de circularité mentionné plus haut) de ne pas se reposer *entièrement* sur l'observation empirique pour déterminer que telle ou telle faculté est de sens commun.

Dans la théorie que je propose, les croyances de sens commun ont donc les trois caractéristiques d'être inévitables (sauf réflexion), naturelles, et communes, mais selon l'ordre de la *ratio essendi* c'est parce qu'elles sont naturelles qu'elles sont inévitables et communes, et selon l'ordre de la *ratio cognoscendi*, nous connaissons (par sens intime) qu'elles nous sont inévitables, ce qui nous permet d'inférer qu'elles sont naturelles, et donc dans un deuxième temps qu'elles sont communes.

Dans les paragraphes précédents, j'ai répondu à la question (i) : quelles croyances sont « de sens commun » ? Fondamentalement, ce sont les croyances *naturelles*. Mais je n'ai pas encore répondu à la question (ii) : pourquoi le fait que ces croyances soient *naturelles* constitue-t-il un motif (une raison épistémique) de les tenir pour fiables ? Plusieurs auteurs contemporains, notamment [Plantinga 2000, 148], ont défendu un principe de « présomption de fiabilité » pour nos facultés naturelles. Le principe est essentiellement de la forme suivante :

Présomption de fiabilité

Lorsque je sais qu'une source épistémique *S* fait partie des facultés cognitives naturelles de l'esprit humain, j'ai une justification à croire qu'il s'agit d'une source fiable de connaissances.

Ce principe me semble une réponse indispensable au problème du critère ou au problème de la méta-justification de la perception. Seulement, en lui-même, il ne dit pas *pourquoi* une source épistémique *naturelle* peut (de manière justifiée) être présumée comme source fiable. C'est à cette question que je vais essayer de répondre à présent, en m'appuyant sur un autre auteur de la tradition latine du sens commun, l'italien Salvatore Tongiorgi (fin XIX^e siècle), et sur un principe argumentatif de Thomas Reid que je n'ai pas encore mentionné.

Pour ne pas faire durer le suspense, la réponse de Tongiorgi consiste à dire que je suis basiquement justifié à croire que « mes sources épistémiques naturelles sont fiables », parce que le fait de mettre ce principe en doute serait auto-défaisant (*self-undermining* dans le vocabulaire contemporain). Mais pour bien comprendre l'intérêt de cette réponse, il me semble qu'il peut être utile de faire un petit détour aux origines historiques de la philosophie du sens commun.

La philosophie du sens commun, à mon sens, trouve naissance dans une réflexion approfondie à la suite de Descartes sur le problème du fondationnalisme en épistémologie. La question que les épistémologues fondationnalistes, depuis Aristote, ont eu à résoudre est la suivante : quelle faculté peut

produire des croyances qui soient justifiées de manière basique? La réponse d'Aristote, reprise par la majorité des scolastiques, c'est que l'évidence nous permet d'accéder à des vérités logiques qui, fondamentalement, se réduisent au principe de non-contradiction. Si on me demande pour quelle raison je crois à ce principe, je ne pourrai pas apporter une « raison » inférentielle : j'y crois de manière basique. Mais si on me demande si je peux considérer cette croyance basique comme une *connaissance*, la réponse aristotélicienne est, en substance, la suivante : cette proposition *doit* être considérée comme une connaissance, et comme une connaissance première (une vérité primitive), pour les trois raisons suivantes

[Cette vérité] peut et doit être admise par le philosophe sans démonstration ; cela est manifeste parce que

1° elle n'a besoin d'aucune démonstration

2° elle répugne à toute démonstration

3° même la négation ou le doute [de cette vérité] l'affirme.

[Tongiorgi 1862, § 426, traduction de l'auteur]

Le principe de non-contradiction, d'après Aristote, (1°) n'a pas besoin de démonstration parce qu'il est évident ; et (2°) il répugne à toute démonstration parce que le fait même de commencer une démonstration *présuppose* que l'on a accepté le principe de non-contradiction. Il est dans sa nature de *principe* d'être totalement étranger à un fondement démonstratif. Enfin (3°), Aristote montre, par un argument apagogique, que celui-là même qui voudrait mettre en doute le principe de non-contradiction (ou le nier) ne peut s'empêcher d'en présupposer la validité si du moins il entre en débat contradictoire avec son interlocuteur (s'il se tait, nous dit Aristote, il est comparable à une plante et nous n'avons pas à le réfuter). De ces trois conditions, la troisième me paraît la plus fondamentale : en termes contemporains, on dirait que la négation ou le doute du principe de contradiction sont deux attitudes épistémiques *self-undermining*, auto-défaisantes ou auto-réfutantes ; de telle sorte que la *seule* attitude épistémique légitime et justifiée (car non auto-réfutante) est d'*accepter* le principe de non-contradiction.

Cette réponse du fondationnalisme aristotélicien pour « méta-justifier » les fondements basiques de la connaissance me semble extrêmement satisfaisante intellectuellement : elle fait bien voir que telle proposition ne peut être rien d'autre qu'un principe basique, et que la seule attitude justifiée est de croire ce principe basiquement. Le seul défaut de cette théorie des fondements du fondationnalisme, c'est qu'elle ne fonctionne que pour les vérités logiques (ou « idéales »), réduites au principe de non-contradiction. Le modèle d'Aristote, de fait, était la science déductive mathématique. Mais il est clair que pour les vérités *d'existence*, ou les vérités contingentes, on ne va pas pouvoir s'appuyer sur le seul principe de non-contradiction. Si notre connaissance, dans son ensemble, repose sur des « vérités premières », il est clair qu'on ne peut pas réduire ces vérités premières au seul principe de contradiction (ou aux vérités logiques en général). Sur quel principe pourront donc reposer les vérités contingentes? C'est à ce problème que la philosophie première de Descartes

proposait une solution décisive : il faudrait admettre un *second* principe à la base des vérités d'existence ou contingentes, à savoir la croyance « je pense, j'existe » (et toutes les vérités de pure introspection qui sont associées à ce principe – j'ai présentement une impression de rouge, de froid, etc.). Et ce que montre Descartes, c'est que le principe « je suis, j'existe » est indubitable non seulement parce qu'il est évident mais aussi et surtout *parce que le fait même d'en douter présuppose sa vérité*. Autrement dit, le *cogito* a exactement le type de méta-justification épistémique qui avait permis à Aristote d'asseoir la validité du principe de non-contradiction ! D'un point de vue strictement historique, on sait qu'une version du *cogito* était présente chez saint Augustin, mais Augustin n'avait pas vu que son argument anti-sceptique en faveur de la croyance « j'existe » permettait d'élever cette croyance au rang de principe des vérités contingentes au même titre que le principe de non-contradiction était principe des vérités de raison. Jacob Schmutz [Schmutz 2007, 231] a montré de façon passionnante qu'au Moyen Âge Duns Scot avait déjà envisagé d'ajouter au principe de non-contradiction certaines vérités premières contingentes telles que « je suis éveillé » ; et Pierre d'Ailly est allé plus loin en montrant explicitement que l'argument augustinien en faveur de *ego sum* permettait de placer ce principe sur un pied d'égalité avec le principe de non-contradiction [Schmutz 2007, 242]. Mais quels que soient les mérites de Duns Scot et de Pierre d'Ailly, il semble que Descartes soit le premier à avoir essayé de réellement reconstruire tout le savoir humain simplement sur les bases de l'introspection (avec son principe *ego sum*) et de la raison (avec son principe de non-contradiction). Et il est ainsi le premier à avoir mis clairement en évidence que, lorsqu'on essaie *réellement* de tout reconstruire sur ces bases, on est directement confronté au problème de l'argument sceptique : en effet, l'introspection nous permet d'affirmer une existence contingente (la mienne), mais comment tirer de là (et de la raison) l'existence du monde extérieur, alors même qu'un malin génie pourrait me tromper. On sait comment Descartes prétendait s'en sortir : par la démonstration rationnelle de l'existence et de la bonté de Dieu (ce qui lui permettait de ne postuler finalement aucune autre source *basique* que l'introspection et la raison, puisque la perception pouvait être *réduite* à un usage conjoint de l'introspection et de la raison). On sait aussi à quel point cette solution est difficile à accepter.

À mon sens, la philosophie du sens commun telle qu'elle apparaît d'abord chez Buffier n'est rien d'autre qu'une réponse au problème mis en évidence et non résolu par Descartes : une fois qu'on dispose du principe de non-contradiction pour fonder les vérités idéales, et du principe *ego sum* pour fonder les vérités contingentes *qui me concernent*, comment fonder (au sens fort du fondationnalisme épistémologique) toutes les autres vérités contingentes *qui concernent le monde extérieur* ? Le début de la réponse est évident : il faut bien considérer les sens comme des sources basiques de connaissance (et non plus comme des sources *réductibles* à l'introspection et à la raison comme chez Descartes). Le problème évident qui se pose alors est celui qui nous occupe depuis le début de notre réflexion : mais *pourquoi* considérer

les sens comme sources basiques de connaissance, c'est-à-dire comme sources *fiabiles* (ou comme « critères de vérité » pour reprendre la terminologie des épicuriens et des stoïciens lorsqu'ils se posaient la même question)? Buffier est sans doute le premier à avoir répondu à cette question de la manière suivante : *parce que les sens font partie de notre sens commun, c'est-à-dire de nos facultés naturelles*. Cette réponse revenait à introduire le sens commun comme *troisième principe* de fondement ultime de la justification, aux côtés de l'évidence (non-contradiction) et de l'introspection (*cogito*). Mais on a vu que cette réponse n'était pas encore complète, car on peut encore poser la question : *pourquoi* le fait que les sens soient des facultés naturelles nous justifie-t-il à présumer leur fiabilité? Et c'est là que la réponse de Tongiorgi prend tout son sens : Tongiorgi a essayé de montrer que la croyance « mes facultés naturelles sont fiables » peut et doit être posée comme principe justifié *exactement pour les mêmes raisons qui faisaient du principe de non-contradiction et du cogito des principes*; c'est-à-dire : parce que cette croyance 1°) n'a pas besoin d'être démontrée, 2°) répugne à toute démonstration, et 3°) est encore affirmée (présupposée) par celui qui la nie ou la met en doute. Autrement dit, le principe de « présomption de fiabilité » des facultés naturelles a la propriété épistémiquement très intéressante que sa négation (ou un simple doute) est auto-réfutant.

Telle est la motivation de la « théorie des trois vérités primitives » de Tongiorgi que je cite à présent dans sa formulation originelle :

Bien qu'il y ait de nombreuses choses qui ne peuvent ni ne doivent être démontrées, cependant les vérités primitives, qu'il est nécessaire de supposer au fondement avant toute enquête philosophique sont seulement au nombre de trois, à savoir :

1. *le Fait premier*, qui est l'existence à proprement parler
2. *le Principe premier*, qui est le principe de contradiction, *le même ne peut à la fois être et ne pas être*
3. *la Condition première*, à savoir l'aptitude de la raison à atteindre la vérité. [...]

Comme nos connaissances sont de deux ordres, idéal et réel, il faut trouver dans les deux ordres quelque chose de premier : dans l'ordre idéal le *principe premier*, dans l'ordre réel le *premier fait*. L'un et l'autre ensemble sont nécessaires à la constitution de la science. [...] Mais en plus de ces deux-là, un troisième est requis, qui permet la composition de l'un et l'autre ordres, et grâce auquel les faits peuvent être portés à l'ordre idéal, et les idées appliquées aux choses : en effet, vous n'aurez pas cela si vous ne supposez pas que l'esprit peut avoir des idées conformes aux objets. [Tongiorgi 1862, §427, traduction de l'auteur]

La formulation du principe de présomption de fiabilité est un peu différente de celles que nous avons vues jusqu'à présent : Tongiorgi regroupe sous le

nom de « raison » l'ensemble des facultés épistémiques naturelles de l'homme, et au lieu de parler de « source fiable » ou de « source de connaissance », il désigne la connaissance comme un « terme de succès », à savoir « atteindre la vérité ». Mais fondamentalement, l'idée est bien la même, et s'applique tout spécialement à nos facultés de perception, qui permettent de passer de « l'ordre idéal » aux « objets » auxquels nos idées doivent être conformes. Ce qui est plus intéressant pour nous, ce sont les justifications que donne Tongiorgi pour dire que ce principe 3 est aussi fermement justifié que le principe de contradiction et le principe de mon existence :

[le principe 3 répugne à toute démonstration :]

Si quelqu'un voulait démontrer l'aptitude de la raison à atteindre la vérité, il présupposerait immédiatement que cette vérité particulière peut être connue au moyen de la démonstration. Mais la démonstration est l'œuvre de la raison. Donc il présupposerait que la vérité peut être connue par la raison.

[la négation ou la mise en doute du principe 3 est auto-réfutante :]

Enfin, quiconque met en doute quelque chose ou le nie, connaît au moins ceci comme vrai : qu'il faut en douter ou le nier. Donc il reconnaît que quelque vérité peut être découverte par la raison.

[et contre les sceptiques qui affirment que la raison n'est pas fiable :]

Réponse : Qui a révélé cela aux sceptiques ? Si la raison se trompe toujours, ne faut-il pas penser qu'elle se trompe aussi en ceci qu'eux-mêmes affirment ? Cette unique réponse suffit à rabattre tous les sophismes des sceptiques ; et ils ne peuvent ouvrir la bouche sans tourner contre eux-mêmes leurs armes. [Tongiorgi 1862, § 428–430, traduction de l'auteur]

L'idée de Tongiorgi est donc que le sceptique ne peut mettre en doute (ou *a fortiori* nier) la fiabilité de la raison qu'en faisant usage de la raison elle-même, et donc en montrant par là qu'il croit à la fiabilité de la raison. En cela, sa posture est épistémiquement auto-réfutante et injustifiée. La seule attitude qui puisse être justifiée, lorsqu'on s'interroge sur la fiabilité de nos facultés épistémiques, c'est donc *d'admettre* pour principe cette fiabilité.

Le principe de « présomption de fiabilité » de nos facultés naturelles peut donc être défendu par un argument « apagogique » ou « par auto-réfutation de l'adversaire », qui est un argument proprement épistémologique (un argument où l'on montre que l'admission du principe est la seule attitude *épistémiquement justifiée*) et non pas un argument pragmatique (montrant simplement l'*efficacité* d'une telle supposition). Et c'est à mes yeux la meilleure réponse que l'on puisse apporter au problème sceptique du critère – c'est-à-dire au défi de produire une bonne raison (épistémique) de croire à la fiabilité de nos sens (et de nos autres facultés naturelles).

Bien sûr, on pourrait faire de nombreuses objections à cette solution. Je n'en considérerai qu'une seule, qui me permettra de compléter Tongiorgi par une suggestion que je crois très juste de Thomas Reid.

L'objection est la suivante : l'argument d'auto-réfutation consiste à dire qu'on est bien obligé d'utiliser « la raison » ou « nos facultés naturelles » pour mettre en doute (ou pour nier) la proposition « nos facultés naturelles sont fiables ». C'est peut-être vrai si l'on parle à ce niveau très général, sans distinguer nos facultés naturelles. Mais que se passe-t-il si l'on entre dans le détail : suis-je obligé d'utiliser l'odorat pour mettre en doute le fait que l'odorat est fiable ? Suis-je obligé d'utiliser les sens externes pour construire un argument qui nie la fiabilité des sens externes ? Suis-je obligé d'utiliser la mémoire pour nier la fiabilité de la mémoire ? Il semble que non : il semble au contraire que je peux m'appuyer sur ma faculté d'intuition rationnelle et ma faculté déductive (ou inductive – disons : ma raison inférentielle) pour mettre en doute mes sens externes ou ma mémoire. Peut-être est-il impossible de mettre en doute la raison inférentielle sans utiliser la raison inférentielle. Mais dans ce cas, l'argument d'auto-réfutation ne permettrait de sauver qu'une seule (ou peut-être qu'un petit nombre) de mes facultés naturelles, et non pas *toutes* mes facultés naturelles *en bloc*. L'idée de présumer la fiabilité de *toutes* mes facultés naturelles simplement parce qu'elles sont naturelles repose uniquement sur la formulation générale et imprécise de l'argument. Il est possible de mettre en doute la fiabilité des *sens* et de la *mémoire* et de ne s'en tenir qu'aux certitudes de la *raison* et de *l'introspection*, donc d'être un sceptique solipsiste remettant en question l'existence du monde extérieur.

Thomas Reid, bien qu'il n'ait pas formulé cette objection en ces termes, apporte néanmoins dans ses écrits un élément qui permet, je crois, d'y répondre. L'argument auquel je pense, qui apparaît sous différentes formes dans les écrits de Reid, peut être appelé « l'argument de symétrie ». L'idée générale est qu'il n'est pas justifié de faire une différence entre nos facultés naturelles, en traitant l'une comme fiable et l'autre comme non fiable.

La Raison, dit le sceptique, est le seul juge de la vérité, et vous devriez rejeter toute opinion et toute croyance qui n'est pas fondée sur la raison. Pourquoi, cher Monsieur, devrais-je croire la faculté de Raison plus que celle de la perception ; elles sortent toutes deux du même atelier, et ont été faites par le même artiste ; et s'il me met dans les mains une mauvaise marchandise, qu'est-ce qui l'empêcherait de m'en mettre une autre ? [Reid 1764, VI, 20, traduction de l'auteur]

Reid emploie ce principe de symétrie entre nos facultés pour rejeter ce qu'il appelle le « semi-scepticisme » de Hume, qui fait confiance à la raison, mais pas aux autres facultés, ou à la raison *et* à l'introspection (par laquelle je connais mes « Idées ») mais pas à la perception. Pour Reid, le « semi-scepticisme » est fondamentalement instable, et il s'efforce de le rendre « cohérent » en le poussant à un scepticisme absolu, qui doute de *toutes* les facultés humaines :

Je me permettrai de faire une addition au système sceptique, sans laquelle, à mon sens, il ne peut pas tenir. J'affirme que la croyance dans l'existence des impressions et des idées est aussi peu fondée sur la raison que celle dans l'existence des esprits et des corps. Personne n'a jamais offert, ni pu offrir, aucune raison pour cette croyance. [...] Un sceptique complet et cohérent ne cèdera donc jamais sur ce point ; et tant qu'il y tient, vous ne pourrez jamais l'obliger à céder quoi que ce soit d'autre.

À un tel sceptique je n'ai rien à dire ; mais aux semi-sceptiques, je souhaiterais leur demander pourquoi ils croient à l'existence de leurs impressions et de leurs idées. [Reid 1764, V, 7, traduction de l'auteur]

Dans le texte que je viens de citer, Reid fait comme s'il n'avait « rien à dire » au sceptique complet et cohérent, celui qui rejette l'introspection en même temps que la perception. Plus loin, Reid montre que pour être *pleinement* cohérent, le sceptique ne doit pas seulement rejeter l'introspection mais *aussi* la raison (c'est-à-dire ici la faculté inférentielle), et dans cet autre texte, il ne se montre pas si absolument démuné pour répondre au sceptique absolu qui refuserait de se fier aussi bien à la raison qu'à l'introspection et à la perception. C'est le texte que nous avons déjà cité plus haut :

Peut-être le sceptique acceptera-t-il de douter de la raison, plutôt que de donner aucun crédit à la perception. [...] Pour ma part, je n'essaierai jamais de la rejeter [la raison] ; et bien que la partie saine de l'humanité ne sera pas très inquiète de connaître mes raisons, cependant si elles peuvent être utiles à quelque sceptique, les voici. [Reid 1764, VI, 20, traduction de l'auteur]

S'ensuivent les trois raisons mentionnées plus haut, à savoir les deux raisons pragmatistes et la justification par l'inévitabilité. Comme je l'ai dit, ces trois raisons ne me paraissent pas satisfaisantes épistémologiquement. Mais le raisonnement de symétrie qui précède me semble un très bon argument pour étendre à toutes nos facultés naturelles la présomption de fiabilité établie par l'argument d'auto-réfutation (de Tongiorgi) : si nous sommes obligés de nous fier à certaines facultés naturelles, pourquoi croirions-nous *certaines* facultés de notre nature et pas d'autres ? N'est-ce pas la même nature qui nous donne les unes et les autres ? Si certaines facultés naturelles sont fiables, n'est-ce pas une raison (au moins *prima facie*) de supposer les autres fiables également ? Il est donc légitime de considérer « en bloc » la question de la fiabilité de toutes nos facultés naturelles : sont-elles fiables ou non ? Et à cette question générale, la seule présomption qui ne soit pas auto-réfutante est que nos facultés naturelles sont fiables.

Pour conclure cette section, je résumerai l'argument général par lequel j'ai essayé de répondre au défi sceptique de la méta-cohérence pour la perception :

1. Par introspection, je sens que mes croyances perceptives sont *inévitables* à moins que je fasse consciemment l'effort réflexif de les mettre en question.

2. Cette inévitabilité me permet de conclure que cette faculté de croyance perceptive fait partie de ma *nature* comme agent épistémique. (Sens commun comme *nature*)
3. Si une de mes facultés épistémiques est douteuse, alors les autres n'ont pas de raison d'être fiables, et inversement si j'ai des raisons de supposer la fiabilité d'une de mes facultés naturelles, cela constitue une raison de supposer la fiabilité des autres facultés conférées par la même nature. (Principe de symétrie – Reid)
4. Or je ne peux pas ne pas être justifié à tenir au moins certaines de mes facultés épistémiques pour fiables (d'après 5).
5. En effet, douter de (ou nier) la fiabilité de *toutes* mes facultés épistémiques est épistémiquement auto-réfutant (donc injustifié), de telle sorte que seule *l'affirmation* de la fiabilité de certaines de mes facultés naturelles peut être justifiée. (Principe d'auto-réfutation du doute sur nos facultés – Tongiorgi)
6. Donc je suis justifié (*prima facie*) à tenir chacune de mes facultés épistémiques naturelles pour fiables.
7. Donc (par 2 et 6) je suis justifié (*prima facie*) à tenir la perception pour une source fiable de connaissance.

Et par conséquent, lorsque le sceptique me demande, à propos de telle croyance perceptive basique, si elle est une *connaissance fiable*, je n'ai pas à enfreindre le principe de méta-cohérence : je suis en mesure de lui répondre, de manière pleinement justifiée, qu'elle *est* bel et bien une connaissance fiable, *parce qu'*elle est fondée sur une faculté de sens commun. Le recours explicite (ou « fort ») au sens commun permet de répondre à l'exigence de réflexivité épistémique.

5 Conclusion

Dans cet article, j'ai essayé de montrer ce qui me semble particulièrement convaincant et élégant dans le « tableau » d'épistémologie générale qu'a dressé Pascal Engel. Il me semble que son épistémologie est très séduisante parce qu'elle combine des intuitions externalistes très fortes à propos de la connaissance (la connaissance, en tant que telle, ne requiert rien de plus qu'une fiabilité *de facto* de nos croyances, une fiabilité qu'on peut interpréter plus précisément en termes de « sûreté »), tout en faisant droit aux intuitions internalistes qui résistent lorsqu'on essaie de répondre de manière satisfaisante au défi sceptique (après tout, ce défi *est* un défi internaliste ; il ne met pas fondamentalement en cause le fait que nous « sachions » ceci ou cela dans un sens externaliste, mais plutôt le fait que nous soyons *justifiés* à croire ceci ou cela, alors même qu'il nous est difficile de *nous rendre raison* à nous-mêmes de la légitimité de ces croyances).

Un aspect très élégant de l'épistémologie de Pascal Engel, c'est qu'elle apporte des réponses parallèles aux grands arguments sceptiques pour le domaine de la connaissance externaliste et pour le domaine de la justification internaliste. Pour répondre au problème du trilemme d'Agrippa, Engel montre que la connaissance comme la justification sont toutes deux *fondationnalistes*. Pour répondre à l'argument infaillibiliste, il adopte une solution faillibiliste de la connaissance, et révisabiliste de la justification. Pour répondre à l'argument de clôture enfin, il défend à la fois la clôture de la connaissance, et la transmission de la justification. Sur ce dernier point, Engel défend donc un « néo-mooréanisme » de la connaissance externaliste, et en même temps (ce qui est plus original) un « néo-mooréanisme » internaliste.

Sur tous ces points, je n'ai pu qu'exprimer mon accord et mon admiration, dans la mesure où les arguments de Pascal Engel m'ont largement convaincu, et sont même à l'origine de beaucoup de mes propres thèses en épistémologie.

Mais j'ai néanmoins identifié deux points de désaccord qui concernent le problème de la « réflexivité » et la théorie du sens commun. Engel soutient qu'une solution « néo-mooréenne » contient, en plus du principe de clôture, un principe d'anti-réflexivité, qu'il applique (là encore) aussi bien à la connaissance (négation du principe *KK*) qu'à la justification (pas de justification de second ordre). Or il me semble indispensable, pour répondre vraiment au scepticisme, de fournir une justification réflexive, et il me semble par ailleurs que pour fournir une telle justification réflexive, on est obligé de faire un appel au concept de sens commun beaucoup plus fort que celui qui est fait par Moore et, à sa suite, par Engel. Je conclurai donc ces réflexions en rappelant les « questions » que mon propos a soulevées par rapport à l'épistémologie de Pascal Engel.

1. Pascal Engel est-il d'accord pour accepter un principe de méta-cohérence positive ? C'est-à-dire : lorsqu'on me demande explicitement, à propos de telle croyance qui était *prima facie* justifiée, si elle est une connaissance fiable, je *dois* être en mesure de répondre (de manière justifiée) qu'elle est une connaissance fiable, autrement ma justification *prima facie* est défaite (fondationnellement).
2. Si l'on accepte ce principe de méta-cohérence positive, quel sera le fondement de notre justification à tenir la perception comme une source fiable lorsque la question sera posée par le sceptique ? Et en particulier :
3. Comment le fondement de notre méta-justification (*ma vue est fiable*) pourrait-il *résulter* immédiatement d'une croyance perceptive de premier ordre basiquement justifiée (*j'ai deux mains*), alors qu'il n'y a aucun rapport déductif ni inductif entre les deux faits ? N'est-on pas obligé de trouver un fondement de justification sinon *antécédent* (chronologiquement) du moins *indépendant* (logiquement) du contenu des croyances perceptives elles-mêmes ?
4. Si l'on admet qu'il faut un fondement de méta-justification de la perception qui soit indépendant des contenus perçus eux-mêmes, Pascal Engel

serait-il prêt à accepter un recours fort au sens commun, c'est-à-dire la théorie selon laquelle nous sommes justifiés à considérer la perception comme fiable *au motif* qu'elle est une faculté du sens commun ?

5. Et si un recours fort au sens commun est légitime, Pascal Engel accepterait-il de dire que les facultés de sens commun sont celles que nous connaissons être *naturelles* parce que nous les sentons en nous-mêmes comme *inévitables* (sauf réflexion), et que leur naturalité nous donne une raison de nous y fier *parce que* le fait de mettre en doute nos facultés naturelles serait une attitude auto-réfutante ?

Bibliographie

- ALSTON, William P. [1980], Level-confusions in epistemology, *Midwest Studies In Philosophy*, 5(1), 135–150, doi : 10.1111/j.1475-4975.1980.tb00401.x.
- ENGEL, Pascal [2007], *Va savoir! : De la connaissance en général*, Paris : Hermann.
- [2008], Vérité, croyance et justification : propos d'un béotien dogmatique, dans *Pensées pour le nouveau siècle*, édité par A. Wald Lasowski, Paris : Fayard, 339–362.
- [2012], *Épistémologie pour une marquise*, Paris : Ithaque.
- GUILLON, Jean-Baptiste [2016], The evil demon argument as based on closure plus meta-coherence, *Synthese*, sept., 1–29, doi : 10.1007/s11229-016-1215-9.
- HAWTHORNE, John [2005], The case for closure, dans *Contemporary Readings in Epistemology*, édité par M. Steup & E. Sosa, Malden, Ma : Blackwell, 26–43.
- HUEMER, Michael [2011], The puzzle of metacoherence, *Philosophy and Phenomenological Research*, 82(1), 1–21, doi : 10.1111/j.1933-1592.2010.00388.x.
- IMRE, János [1819], *Amicum foedus rationis cum experientia seu philosophia crisi recentissima deducta per Johannem Imre*, Pest : typis Joannis Thomae Trattner.
- KLEIN, Peter [1995], Skepticism and closure : Why the evil genius argument fails, *Philosophical Topics*, 23(1), 213–236, doi : 10.5840/philtopics19952315.
- LEMONS, Noah Marcelino [2004], *Common Sense : A Contemporary Defense*, New York : Cambridge University Press.

- MERCIER, Désiré-Joseph [1911], *Cours de philosophie*, t. IV : Critériologie générale ou traité général de la certitude, Louvain : Institut Supérieur de Philosophie, 6^e éd.
- MOORE, George Edward [1962], *Commonplace Book, 1919-1953*, London ; New York : Allen & Unwin ; Macmillan.
- [1993], A defence of common sense, dans *Selected Writings*, édité par T. Baldwin, London ; New York : Routledge, 106–133.
- PLANTINGA, Alvin [2000], *Warranted Christian Belief*, New York : Oxford University Press.
- POLLOCK, John L. & CRUZ, Joseph [1999], *Contemporary Theories of Knowledge*, Lanham : Rowman & Littlefield Publishers.
- PRITCHARD, Duncan [2005], *Epistemic Luck*, New York : Oxford University Press.
- REID, Thomas [1764], *Thomas Reid, an Inquiry into the Human Mind on the Principles of Common Mind*, Dublin : Alexander Ewing.
- SCHMUTZ, Jacob [2007], L'existence de l'ego comme premier principe métaphysique avant Descartes, dans *Généalogies Du Sujet. De Saint Anselme à Malebranche*, édité par O. Boulnois, Paris : Vrin, 215–268.
- TONGIORGI, Salvatore [1862], *Institutiones philosophicae*, t. 1 : Logica, Bruxelles : Goemaere.
- WILLIAMSON, Timothy [2000], *Knowledge and Its Limits*, Oxford ; New York : Oxford University Press.
- YAFFE, Gideon & NICHOLS, Ryan [2000], Thomas Reid, *Stanford Encyclopedia of Philosophy*, URL <http://plato.Stanford.Edu/entries/reid>.