

Du rôle de l'intérêt dans nos méthodes d'acquisition de connaissances

Stéphane Plante*

Résumé

*L'épistémologie computationnelle de Kevin Kelly cherche à décrire le processus d'acquisition des connaissances de manière purement formelle, sans tenir compte des motivations des agents. Les travaux de Thomas Gilovitch et de Philip Kitcher démontrent cependant que ces motivations subjectives peuvent profondément modifier notre compréhension du processus d'acquisition des connaissances et remettre en question la validité des connaissances acquises, quelles que soient les méthodes employées. Il est donc nécessaire d'étudier ces motivations à l'aide d'une théorie des valeurs épistémiques. C'est ce que se propose Alvin Goldman, dans son livre *Knowledge in a Social World*, avec sa théorie des valeurs véritistiques. Son approche vise cependant d'abord et avant tout à préserver la validité de nos connaissances dans un environnement social et réduit à nouveau le rôle des intérêts des agents à un rôle secondaire. Pour Fallis et Kitcher, ces valeurs véritistiques sont insatisfaisantes et seule une pleine pesée des intérêts qui animent les agents permettra une compréhension complète du processus d'acquisition de connaissances. Malgré tout, l'approche de Goldman permet de mettre en lumière comment les biais des agents peuvent être fondamentaux pour expliquer leur apparente objectivité.*

* Étudiant à la maîtrise en philosophie, Université de Montréal.

1 Introduction

Dans son plus récent livre, Vincent Hendricks présente l'idée que l'épistémologie formelle a franchi une étape importante lorsqu'elle est passée d'une conception idéalisée d'un agent inactif à une conception plus active de l'agent. Selon lui,

We need to activate the agent in order to make epistemic logic pertinent to epistemology, computer science, artificial intelligence and cognitive psychology. [...] Inquiring agents are agents who read data, change their minds, interact or have common knowledge, act according to strategies and play games, have memory and act upon it, follow various methodological rules, expand, contract or revise their knowledge base, and so on, all in the pursuit of knowledge. Inquiring agents are *active agents*.¹

L'idée d'un agent purement rationnel et logiquement omniscient est ainsi abandonnée au profit de considérations plus proches des capacités et limites de l'agent.

Malgré cela, certaines limites semblent encore inexplorées et pourraient avoir un impact épistémologique important. Ce texte tentera d'explorer l'une de ces limites : l'intérêt de l'agent. Je tenterai ainsi de déterminer s'il est légitime de faire abstraction du rôle de l'intérêt dans notre évaluation de nos méthodes d'acquisition de la connaissance.

Dans cette optique, je prendrai d'abord appui sur la théorie formelle de l'apprentissage de Kevin Kelly qui, aux dires de Hendricks, réalise naturellement la perspective de l'agent actif.² Cette théorie me permettra de présenter comment les motivations de l'agent peuvent transformer et mettre en doute une approche formelle au départ prometteuse. J'utiliserai ensuite les idées de Phillip Kitcher³ pour

¹ Vincent HENDRICKS, *Mainstream and formal epistemology*, Cambridge, Cambridge University Press, 2006, p. 102.

² *Ibid.*, p. 124.

³ Philip KITCHER, « The Ends of the Sciences », dans B. LEITER (éd.), *The Future for Philosophy*, Oxford, Clarendon Press, 2004, p. 208-229.

montrer pourquoi il est non seulement nécessaire mais primordial pour toute épistémologie de considérer ces intérêts. Le livre de Alvin Goldman, *Knowledge in a Social World*⁴, et les diverses critiques qui en sont nées me serviront finalement de base pour tenter de comprendre comment incorporer l'intérêt de l'agent dans un système décrivant nos pratiques d'acquisition de connaissances et comment la vérité peut demeurer un but fondamental des agents malgré, et peut-être même à cause de leurs motivations variées.

2 Une théorie formelle de l'acquisition de la connaissance

L'épistémologie computationnelle de Kelly étudie les méthodes utilisées par un agent pour résoudre un problème épistémique. L'analyse débute par une division du problème épistémique en cinq points :

A learning problem specifies : (1) what is to be learned, (2) a range of relevant possible environments in which the learner must succeed, (3) the kinds of inputs these environments provide to the learner, (4) what it means to learn over a range of relevantly possible environments, and (5) the sorts of learning strategies that will be entertained as solutions.⁵

Pour bien illustrer cette approche, il est utile de reprendre un exemple simple de Schulte⁶. Supposons que le problème épistémique à résoudre pour un ornithologue est de déterminer si tous les corbeaux sont noirs. Ce problème spécifie d'abord ce qui doit être appris (1) :

⁴ Alvin GOLDMAN, *Knowledge in a social world*, Oxford, Clarendon Press / Oxford University Press, 1999.

⁵ Vincent HENDRICKS, op. cit., p. 115-116. Il cite Kevin KELLY, « Iterated Belief Revision, Reliability, and Inductive Amnesia », *Erkenntnis*, vol. 50, no1, 1999, p. 7-53., mais je n'ai pu trouver cette citation en aucun endroit de ce texte. Les cinq points cités sont néanmoins utiles pour l'analyse des idées de Kelly et je conserve donc cette division.

⁶ Oliver SCHULTE, « Formal Learning Theory », *The Stanford Encyclopedia of Philosophy* (Spring 2006 Edition), Edward ZALTA (éd.), URL=<http://plato.stanford.edu/archives/spr2006/entries/learning-formal/>, 2002.

si oui ou non tous les corbeaux sont noirs. Il spécifie aussi (2) quels mondes possibles sont pertinents pour répondre à la question. Si on examine en effet tous les mondes fictifs que notre imagination nous permet de considérer, certains d'entre eux nous paraîtront clairement impossibles et seront écartés d'emblée (un monde où tous les corbeaux seraient à la fois entièrement noirs et entièrement blancs par exemple). Mais parmi les mondes qui demeurent, les mondes possibles, certains ne sont pas pertinents et peuvent aussi être écartés. Dans notre exemple, le problème épistémique se base sur des présuppositions qui éliminent certaines possibilités d'erreurs. La question présuppose par exemple que les corbeaux existent et qu'il nous est possible de les percevoir sans qu'il nous soit nécessaire de démontrer la fiabilité de notre perception (que nous ne sommes pas trompés par un malin génie). Les données épistémiques que présentent ces mondes possibles pertinents (3) sont alors simplement l'ensemble des observations individuelles de la couleur d'un corbeau. Toute autre information, la couleur des hirondelles par exemple ou la longueur du bec de ces corbeaux, n'a pas d'importance pour ce problème spécifique.

Les quatrième et cinquième points sont particulièrement importants et méritent qu'on s'y attarde. Que signifie apprendre la réponse à une question (point 4) ? Kelly ne se conforme pas aux approches qui présentent la connaissance comme l'atteinte de la certitude. Dans certains cas selon lui, une telle infaillibilité est utopique, car elle présuppose que l'enquête s'arrête et qu'il n'est conséquemment plus possible de changer d'idée. Au contraire, dans plusieurs cas, cette enquête semble sans fin.

The procedural conception suggests an alternative response : dropping the *halting* requirement, for after the procedure halts, it cannot retract its opinion and correct it later if its answer is mistaken. [...] The foundational metaphor of instantaneous, partial, support is replaced,

thereby, with the pragmatic metaphor of fallible but self-correcting *convergence* to a correct answer.⁷

C'est donc par la convergence ou non à une réponse correcte que l'on peut déterminer s'il y a apprentissage.

Kelly introduit ainsi plusieurs critères de convergence permettant l'atteinte d'une réponse juste. Pour reprendre l'exemple de l'ornithologue, *il converge avec certitude* si, après un certain temps, il conjecture une réponse, cesse son investigation et se fie à cette réponse à tous les temps ultérieurs. S'il voit un corbeau blanc par exemple, il conjecture que tous les corbeaux ne sont pas noirs et cesse son investigation. L'ornithologue *converge plutôt à la limite* si, après un certain temps, il conjecture une réponse et se fie à celle-ci à tous les temps ultérieurs sans toutefois cesser son enquête. Sa conjecture pourrait être que tous les corbeaux sont noirs et qu'aucune observation ne l'entraîne jamais à changer d'idée. Un autre exemple de convergence est la convergence après un certain nombre de changements d'idée. Cette forme de convergence s'applique particulièrement bien à l'analyse des sciences. Elle consiste simplement à conjecturer plusieurs hypothèses pour s'ajuster aux données empiriques à mesure que celles-ci sont obtenues. Une première hypothèse est émise puis remplacée par une autre lorsque les observations semblent plus cohérentes avec cette seconde hypothèse. Puis, à nouveau, une nouvelle hypothèse que l'on suppose plus proche de la vérité remplace la seconde et ainsi de suite. Si une méthode trouvait une réponse adéquate après que n hypothèses ont été émises⁸, l'enquête cesserait et l'on considérerait qu'elle converge après n changements d'idée. Alternativement, si l'investigation ne s'arrêterait jamais

⁷ Kevin KELLY, « The logic of success », *The British Journal for the Philosophy of Science*, vol. 51, no 4, 2000, p. 641.

⁸ L'arrêt de l'enquête n'a pas à être arbitraire. À titre de comparaison, la convergence à la certitude peut être perçue comme une convergence sans aucun changement d'idée. L'ornithologue cesse ses recherches lorsqu'il voit un corbeau blanc, c'est-à-dire lorsque l'une des deux seules hypothèses permises par la question (les corbeaux sont soit tous noirs, soit ils ne le sont pas) est rejetée. De même, le rejet de toutes les hypothèses alternatives permet de cesser l'investigation après n changements d'idée.

mais que le mécanisme était le même, la convergence serait plutôt considérée *graduelle*. On considérerait alors qu'on se rapproche toujours plus de la réponse mais sans jamais l'atteindre.

Tous ces critères de convergence constituent une première manière de classer les stratégies de réponse à la question (point 5). Une certaine hiérarchie peut ainsi être établie entre eux comme le propose Hendricks : « The epistemic problems solvable with certainty are also solvable with n-mind change, solvable in the limit and solvable gradually. The converse relation does not hold, however. »⁹ En effet, les problèmes qui trouveront leur solution par une stratégie visant la certitude sont moins nombreux que ceux qui trouveront leur solution à la limite. La figure suivante représente bien cette hiérarchie¹⁰ :

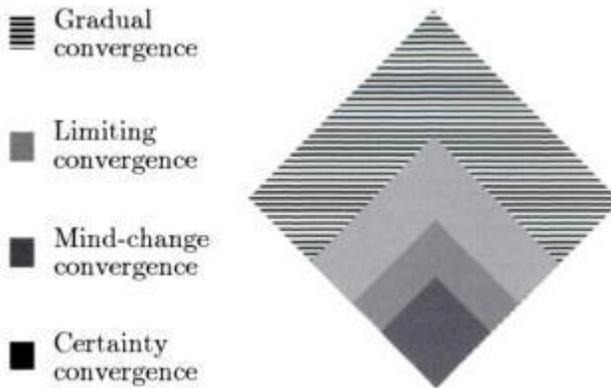


FIGURE 7.2. The nesting of convergence criteria.

Ces notions de convergences sont complétées par des notions de succès qui délimitent elles aussi les stratégies possibles de réponse (point 5). Les méthodes peuvent ainsi être classées selon qu'elles évaluent une hypothèse ou qu'elles en permettent plutôt la découverte. Un exemple de méthode de découverte pourrait être l'inférence à la

⁹ Vincent HENDRICKS, *op. cit.*, p. 120-121.

¹⁰ *Ibid.*, p. 121.

meilleure explication. Un ornithologue ayant vu de nombreux corbeaux pourrait inférer l'hypothèse que tous les corbeaux sont noirs, car c'est là l'hypothèse qui semble le mieux expliquer le fait que tous les corbeaux observés jusqu'ici étaient noirs. Une méthode d'évaluation permettrait quant à elle de tester cette hypothèse en tentant, par exemple, de la réfuter (avec certitude) par un ou des contre-exemples.

Computationally speaking, discovery (converging to a correct answer) and test (determining whether an answer is correct) are simply two different types of problems (think of adding $x + y$ to obtain z as opposed to deciding whether $x + y = z$). If anything, discovery is the more fundamental aim, test being a means for filtering out incorrect answers.¹¹

L'un des résultats importants qui découlent de cette structure permet de démontrer que certains problèmes épistémiques ne sont pas solubles quels que soient les critères de convergence et de succès choisis. On pourrait en fait avoir l'impression qu'il suffit de choisir un critère de succès ou de convergence faible pour s'assurer d'atteindre une réponse. Mais au contraire, certains problèmes sont totalement insolubles. Il est ainsi vain de tenter de répondre à une question sceptique du type « Sommes-nous des cerveaux dans une cuve ? ». Ce type de question nous force en effet à considérer comme pertinente une infinité de possibilités qui mettent en doute toute information épistémique reçue. Cette information n'étant plus concluante, elle ne nous permet plus de répondre à la question. Il s'agit là d'un exemple important d'une limite pour l'agent. Il faut toutefois noter que cette limite est due au caractère fondamentalement contextuel de l'analyse et non à l'agent : la réponse doit cadrer dans le contexte établi par les mondes possibles pertinents. Si rien ne limite les mondes possibles qui doivent être considérés, une infinité de réponses devient possible et aucune information épistémique ne permettra jamais de préférer une réponse à une autre. La question devra demeurer sans réponse.

¹¹ Kevin KELLY, *op. cit.*, 2000, p. 642-643.

Un autre résultat important a aussi été établi par Kelly et Schulte. Un agent cohérent (ou une méthode cohérente) devrait rejeter toute hypothèse qui a été réfutée par l'observation. Or, Kelly et Schulte ont démontré¹² que certains problèmes existent pour lesquels ce pré-requis empêche l'atteinte d'une solution pourtant atteignable, avec certitude, par une autre méthode non-cohérente. Ce résultat semble indiquer qu'une limite de l'agent pourrait être favorable à l'acquisition de certaines connaissances.

3 Le problème de la motivation de l'agent

Mais quel est le rôle précis de l'agent dans cette épistémologie ? Hendricks semble limiter ce rôle au choix et à l'utilisation d'une méthode d'enquête efficace : « If it can be shown *a priori* from the computational epistemological perspective that some method the first-person inquirer is considering adopting is destined to flunk, then there is not much sense in adopting this method, assuming the point of inquiry is to get to the truth. »¹³ C'est pourquoi il remplit bien un rôle actif dans son acquisition de la connaissance ; s'il choisit et utilise efficacement la méthode d'enquête, il obtient de bons résultats qui convergent vers la vérité.

Mais que la méthode soit le résultat d'un choix a des conséquences importantes, car il implique les notions de convergence et de succès. Ainsi, c'est l'agent qui décide de ce qui sera considéré comme étant un succès ou non (une confirmation de l'hypothèse ou la découverte d'une nouvelle hypothèse par exemple), et c'est également lui qui choisit à quel moment s'arrêtera l'enquête (il choisit le critère de convergence). Qui plus est, il est clair que l'agent doit avoir un intérêt dans ce problème épistémique sans quoi il ne tenterait même pas une enquête sur la question. L'agent choisit donc aussi le problème épistémique.

¹²Kevin KELLY et Oliver SCHULTE, « The computable testability of theories making uncomputable predictions », *Erkenntnis*, vol. 4, no. 1, 1995, p. 29-66.

¹³Vincent HENDRICKS, *op. cit.*, p. 128.

Tous ces choix d'une importance capitale sont pourtant absents de la discussion. La présupposition de Hendricks est explicite dans la citation précédente : l'agent doit viser la vérité et ses choix doivent se faire selon ce but ultime. Kelly semble plus ouvert à d'autres buts : « Kuhn's celebrated nihilism about truth may seem to contradict a reconstruction anchored in success, but the logic of success applies to every cognitive goal that transcends the local situation of the scientist. »¹⁴ Il n'explore cependant pas davantage ces autres buts cognitifs.

Dans son livre *How we know what isn't so*¹⁵, Thomas Gilovitch explore l'effet des motivations sur la connaissance et remet ainsi en question le fondement de chacune des cinq propriétés du problème épistémique.

When we prefer to believe something, we may approach the relevant evidence by asking ourselves "what evidence is there to support this belief?" [...] Note that this question is not unbiased : It directs our attention to supportive evidence and away from information that might contradict the desired conclusion. Because it is almost always possible to uncover some supportive evidence, the asymmetrical way we frame the question makes us overly likely to become convinced of what we hope to be true.¹⁶

Ainsi, la question elle-même serait biaisée dès le départ pour mener à un résultat prédéterminé. Ce biais a un effet direct sur le choix du critère de convergence et sur ce qui sera considéré comme un succès :

When the initial evidence supports our preferences, we are generally satisfied and terminate our search ; when the initial evidence is hostile, however, we often dig deeper, hoping to find more comforting information or to

¹⁴ Kevin KELLY, *op. cit.*, 2000, p. 655.

¹⁵ Thomas GILOVICH, *How we know what isn't so: The fallibility of human reason in everyday life*. New York, Free Press, 1991.

¹⁶ *Ibid.*, p. 81. La citation est tirée de Alvin GOLDMAN, *op. cit.*, 1999, p. 235.

uncover reasons to believe that the original evidence was flawed. By taking advantage of “optional stopping” in this way, we dramatically increase our chances of finding satisfactory support for what we wish to be true.¹⁷

Sans qu’il soit nécessaire d’adopter le point de vue radical voulant que toutes nos enquêtes soient fondamentalement biaisées de la sorte, il demeure que la présence d’un biais est toujours possible. À l’agent rationnel idéal de l’épistémologie formelle, on peut donc opposer un agent parfaitement biaisé à toutes les étapes du processus d’investigation. Or, rien dans l’épistémologie computationnelle ne permet de faire la distinction entre les deux ; rien ne permet de distinguer une information acquise de manière objective d’une autre acquise de manière biaisée, ni un agent qui s’arrête parce qu’il a atteint la certitude d’un autre qui s’arrête parce qu’il est simplement satisfait du résultat.

4 Les fins des sciences

Le problème est similaire à celui auquel font face les empiristes logiques. Ceux-ci présentent la science comme une zone libre de valeurs. « They are envisaging a contrast between two ways of reacting to a body of evidence, comparing the sober judge who doesn’t allow personal preferences to interfere with the credulous wishful thinker. »¹⁸ Pour Kitcher, une telle opposition n’a pas lieu d’être.

Pour démontrer ce point, il propose de préciser les différents contextes d’analyse de la science. Les empiristes logiques font en effet une distinction entre le contexte de découverte, qui serait du ressort des historiens, des psychologues et des sociologues, et le contexte de justification, un domaine propre à la philosophie. Cette distinction entre les contextes n’est toutefois pas satisfaisante pour Kitcher qui préfère diviser l’entreprise scientifique en quatre contextes : le contexte d’anticipation, dans lequel les questions à résoudre sont

¹⁷ *Ibid.*, p. 82. La citation est tirée de Alvin GOLDMAN, *op. cit.*, 1999, p. 236.

¹⁸ Philip KITCHER, *op. cit.*, 2004, p. 210.

choisies ; le contexte de poursuite, qui comprend comment enquêter sur ces questions ; le contexte de résolution (le contexte de justification des empiristes logiques), dans lequel certaines propositions sont acceptées ou refusées selon les données épistémiques obtenues ; et le contexte d'application, dans lequel les résultats sont traduits en applications. Ces quatre contextes doivent selon lui être envisagés comme quatre étapes d'un procédé visant certaines fins : les fins des sciences. En ce sens, l'ensemble du processus doit être compris de manière unifiée ; les quatre étapes ne peuvent pas être étudiées séparément, car elles s'influencent les unes les autres. « From this perspective, the approach of logical empiricism must appear myopic. Cutting off the context of justification (roughly my context of resolution) from the surrounding occasions of scientific decision, philosophers of science have believed they could identify the methods governing that context in isolation. »¹⁹ Qui plus est, les jugements de valeur font partie intégrante du processus. À la première étape, les questions choisies seront celles qui auront le plus de chances de mener à un certain bénéfice (qu'il soit technologique, politique, économique ou autre) ; à la seconde étape, de nombreuses contraintes sont imposées sur les méthodes de recherche, pour protéger l'environnement par exemple ; et à la quatrième étape, les applications possibles sont choisies en fonction de leur valeur respective. Il devient donc douteux de croire que de tels jugements de valeur n'affectent pas le contexte de résolution comme le voulaient les empiristes logiques.

Kitcher conclut de ceci que le but des sciences est le seul standard par lequel les méthodes scientifiques peuvent être évaluées. « Effectively, I'm opposing a "deontological" account of rationality, views that take being rational to be a virtue independently of the consequences. If a method were shown to be inept for attaining the goals we set for ourselves, it would be empty to declare it "cognitively virtuous" or to view following it as a "cognitive duty". »²⁰ Similairement, il paraît légitime de généraliser cette discussion à l'acquisition des connais-

¹⁹ *Ibid.*, p. 211.

²⁰ *Id.*, note 12.

sances en général et au problème de l'épistémologie computationnelle. Il apparaît alors qu'il faut d'abord déterminer quels sont les intérêts des agents pour comprendre comment ils choisissent les méthodes qui leur permettront d'acquérir des connaissances. Si leur seul intérêt est la vérité, l'épistémologie computationnelle et l'empirisme logique ne seront pas remis en question. Sinon, ils pourront être jugés insuffisants pour correspondre à la réalité motivationnelle des agents en quête de connaissances.

5 Valeurs véritistiques

Mais comment déterminer quelles fins sont visées par les agents ? Comment déterminer ce qui les motive à rechercher une réponse à certaines questions ? La notion de valeur nous sera ici utile. En effet, ce n'est que si l'agent prête une valeur à un but qu'il tentera de l'atteindre. Un but sans valeur sera ignoré. Comme l'explique Fallis : « we need to understand epistemic values in order to explain or to guide the decisions of real people who want to bring about good epistemic outcomes for themselves and others. »²¹ Ce sont donc les valeurs épistémiques qui nous intéresseront, c'est-à-dire les valeurs qui mèneront à l'acquisition de nouvelles connaissances.

À cet effet, la théorie de la valeur véritistique de Goldman constitue un bon point de départ pour évaluer à quel point la vérité peut intéresser les agents. Goldman explore les circonstances dans lesquelles nos pratiques mènent à une croyance vraie. De telles pratiques, parce qu'elles sont orientées vers la production de vérités, sont dites véritistiques. « The main question for veritistic epistemology is : Which practices have a comparatively favorable impact on knowledge as contrasted with error and ignorance ? »²². Il construit donc une théorie dans laquelle la vérité garde toute son importance.

²¹ Don Fallis, « Epistemic Value Theory and Social Epistemology », *Episteme*, vol. 2, n° 3, 2006, p. 177.

²² Alvin GOLDMAN, *op. cit.*, 1999, p. 5.

Deux classifications sont envisagées. Une première divise les croyances selon trois niveaux : la croyance, l'absence de croyance et le rejet de la croyance. Une deuxième les divise plutôt selon le degré de croyance de l'agent (DB pour *degree of belief*). Ce degré est déterminé par la probabilité subjective que l'agent donne à une proposition. Par exemple, si je crois que la proposition P a 65% de chances d'être vraie, mon DB pour P sera de 0.65 (et de 0.35 pour $\neg P$). La valeur véristique (surnommée V-value) est évaluée selon un modèle de questions et de réponses. Prenons l'exemple d'une question du type « est-ce le cas que P ? » qui peut être abrégée par $Q(P/\neg P)$. Si P est vrai, la croyance que P (ou la croyance que $\neg P$, que l'on suppose être la même chose) mène à une V-value plus grande que le rejet de P. Selon le schéma DB, la V-value est liée au DB de la vérité. Si donc je crois que P a 65% de chances d'être vrai et que P est vrai, ma V-value est de 0.65. Si P est faux, elle sera plutôt de 0.35. C'est toutefois les changements de V-value dans le temps qui nous intéressent. Si donc je passe de la croyance que $\neg P$, ou de l'absence de croyance, à la croyance que P, ma V-value augmentera si P est vrai ou diminuera si P est faux. L'analyse est encore plus précise dans le schéma DB. Si ma croyance en P passe de 0.65 à 0.90, ma V-value a été augmentée de 0.25 si P est vraie ou diminuée de 0.25 si P est faux.

Pour résumer, plus la croyance de l'agent se rapproche de la vérité, plus celle-ci a de valeur. Goldman semble donc en accord avec le projet des empiristes logiques. Mais un point important pourrait rompre cet accord ; pour lui, un état de croyance n'a de valeur que s'il répond à une question qui intéresse le sujet.

Suppose S is ignorant about all of the following matters : What is the 323rd entry in the Wichita, Kansas telephone directory ? Who placed sixth in the women's breast stroke at the 1976 Summer Olympics ? What was the full name of Domenico Scarlatti's maternal grandmother ? Does S's ignorance on all these matters constitute, or even contribute toward, the impoverished V-condition of his credal

corpus ? Does such ignorance imply that his credal state should receive a low V-value, or V-ranking ? If S is totally uninterested in these questions, I am inclined to say that his knowing no answers to them does not count against the V-value of his belief states.²³

L'intérêt de l'agent a donc son importance.²⁴

Mais quel est son rôle dans la valeur d'une croyance ? Goldman perçoit deux possibilités. Une première possibilité consiste à lui donner un rôle dominant. La valeur d'une croyance serait directement proportionnelle au degré d'intérêt de l'agent sur la question. Ainsi, apprendre une vérité sur un sujet très intéressant aurait une plus grande valeur que d'apprendre une autre vérité sur un sujet peu intéressant. Goldman perçoit deux problèmes avec cette possibilité. D'abord, ceci impliquerait que certaines questions triviales mais très intéressantes auraient une aussi grande valeur que d'autres questions plus complexes mais tout aussi intéressantes. Les réponses semblent pourtant avoir une valeur intellectuelle supérieure dans le second cas. Qui plus est, certains intérêts pourraient nous entraîner à favoriser l'ignorance à la connaissance. « This would lead down the path toward abandoning the specialized, veritistic mission of epistemology in favour of a more purely pragmatic enterprise. »²⁵

En conséquence, le rôle de l'intérêt doit être modéré : « ... sensitivity to relative amounts of interest should play a modest role in assessing a practice's epistemic credentials. »²⁶ Mais il n'explique pas

²³ *Ibid.*, p. 88-89.

²⁴ On pourrait objecter qu'il est aussi possible que certaines questions aient une valeur intrinsèque supérieure à d'autres, et ce, indépendamment des intérêts personnels des agents. Ce pourrait être le contexte par exemple qui déterminerait la valeur d'une question. Goldman ne fait aucune mention de cette possibilité et elle ne sera pas développée en détail ici. Notons tout de même qu'elle pose un nouveau problème : Comment juger de la valeur intrinsèque d'une question sans s'emmêler dans ses propres intérêts ou dans ceux de l'agent ? Il semble que même de ce point de vue, il sera nécessaire de considérer les intérêts des agents pour tout au moins s'assurer que la valeur de la question n'est pas entachée d'intérêts.

²⁵ *Ibid.*, p. 95.

²⁶ *Ibid.*, p. 96.

en quoi consiste ce rôle modeste exactement, ce que lui ont reproché plusieurs critiques²⁷. Ce n'est que dans une réponse à une critique de Maffie²⁸ qu'il précise sa pensée en niant finalement que la force de l'intérêt ait une quelconque importance : « Any interest in the question, by any suitable agent, admits the belief into the sphere of V-value. By far the more important issue is how much V-value a belief receives ; and this does not depend on the amount of interest associated with it. »²⁹

C'est donc que l'intérêt ne fait que déterminer si une croyance aura une valeur véritistique ou non, et demeure sans effet sur la grandeur de cette valeur. Ce point pourrait être renforcé par une suggestion de Schmitt³⁰ selon lequel il y aurait deux mesures pour toute croyance, l'une véritistique et indépendante de tout intérêt et l'autre axée sur l'intérêt. La valeur totale d'une pratique ou d'une croyance serait ainsi une fonction de ces deux mesures, mais la valeur véritistique ne dépendrait que de la vérité. Selon Goldman, une telle approche est une alternative plausible à la sienne³¹.

6 Les critiques de Goldman

D'une manière similaire aux empiristes logiques, Goldman semble donc séparer les motivations des agents de la pure rationalité. La nature précise de l'intérêt n'a en effet aucune importance ; l'analyse véritistique peut être lancée par la simple présence d'un intérêt, quel qu'il soit et quel que soit sa force. De fait, même un intérêt potentiel (l'agent considérerait la question intéressante s'il y était confronté)

²⁷ Voir entre autres Philip KITCHER, « Veritistic Value and the Project of Social Epistemology », *Philosophy and Phenomenological Research*, vol. 64, no 1, 2002, p. 191-198 et Don FALLIS, *op. cit.*

²⁸ James MAFFIE, « Alternative epistemologies and the value of truth », *Social Epistemology*, vol. 14, no. 4, 2000, p. 247-257.

²⁹ Alvin GOLDMAN, « Replies to reviews of 'Knowledge in a Social World' », *Social Epistemology*, vol. 14, no. 4, 2000, p. 321.

³⁰ Frederick SCHMITT, « Veritistic value », *Social Epistemology*, vol. 14, no 4, 2000, p. 259-280.

³¹ Alvin GOLDMAN, *op. cit.*, 2000, p. 326.

est suffisant selon lui³². L'intérêt de l'agent occupe donc une place minimale. C'est au contraire la vérité qui occupe la place centrale, car la valeur véritistique est un résultat direct du degré de rapprochement entre la croyance et la vérité.

Mais ceci pose problème. Aucune distinction nette n'est faite entre les intérêts épistémiques, c'est-à-dire les intérêts visant l'atteinte d'une croyance vraie, et les intérêts non-épistémiques, comme un intérêt financier. La valeur véritistique inclut autant l'un que l'autre. En conséquence, « any interest whatsoever becomes an epistemic interest if pursued in conjunction with true belief. »³³ Or, une distinction plus claire entre les deux semble nécessaire pour permettre de comprendre comment des intérêts non-épistémiques peuvent influencer des intérêts purement épistémiques.

La solution devra-t-elle passer par une pesée des intérêts des agents ? Cela semble être une possibilité qu'envisage Kitcher :

Goldman could have proposed that the *pure epistemic* weighting is the one that a subject would produce after a reflective procedure in which she abstracted from the practical implications of potential knowledge ; he could then set the standard for assessing practices to be their production of V-value according to this weighting.³⁴

Similairement, Fallis suggère l'importance de peser les intérêts pour pouvoir les ordonner selon leur force relative et ainsi permettre la prise de décision. Sans une telle analyse des intérêts, certains problèmes ne peuvent être résolus. Supposons par exemple que l'on soit opposé à trois croyances : b1 est une croyance vraie à propos de ce qui ferait du monde un meilleur endroit ; b2 est une croyance vraie à propos du nombre de grains de sable sur une plage particulière ; et b3 est une croyance vraie à propos de la 323^e entrée du bottin téléphonique de Wichita. Il est clair qu'une pratique A qui nous permettrait d'atteindre b1 a une plus grande valeur qu'une autre, B,

³² Alvin GOLDMAN, *op. cit.*, 1999, p. 94-95.

³³ James MAFFIE, *op. cit.*, 2000, p. 253.

³⁴ Philip KITCHER, *op. cit.*, 2002, p. 193.

nous menant à b2 et b3. L'approche véritistique favoriserait effectivement A à B, mais elle le ferait parce que B n'est d'aucun intérêt. Ce n'est pas la bonne raison selon Fallis : « ... the reason that outcome A is epistemically better than outcome B is that it is simply much more *useful* to know b1 than it is to know both b2 and b3. »³⁵ Un autre exemple permet de démontrer ce point. Supposons que b2 et b3 soient remplacés par b5, la croyance vraie de l'heure à laquelle sera diffusé le prochain match des Canadiens de Montréal, et b6, la croyance vraie de l'heure de diffusion de l'émission *Heroes*. Ces deux croyances sont d'un certain intérêt pour certains agents. Conséquemment, pour eux, une pratique C qui mènerait à ces croyances devrait obtenir une certaine V-value. Si on suppose qu'autant la pratique A que la pratique C font passer les agents d'une absence de croyance à une croyance complète, c'est-à-dire que la V-value de l'atteinte de b1, b5 et b6 est la même, il apparaît que l'approche véritistique conclura que la pratique C a plus de valeur que la pratique A !

Pour clarifier, on peut s'imaginer que la classification proposée par Goldman serait le résultat de deux contraintes normatives simples :

NC1 : All other things being equal (epistemically speaking), a particular person not having a particular true belief is never epistemically better than that person having that true belief.³⁶

NC2 : All other things being equal (epistemically speaking), a particular person having a particular false belief is never epistemically better than that person not having that false belief.³⁷

Ces contraintes permettent de classer les croyances selon les trois niveaux présentés plus tôt (croyance vraie, absence de croyance et rejet de la croyance). Elles offrent l'avantage de bloquer une objection de Goldman contre la pensée des intérêts. Tant que NC1 et NC2

³⁵ Don FALLIS, *op. cit.*, p. 182.

³⁶ *Ibid.*, p. 183.

³⁷ *Ibid.*, p. 187.

sont respectées, la « mission véritistique de l'épistémologie » n'est pas en danger.

La classification qu'elles offrent n'est toutefois pas complète. Dès que des croyances d'une même force mais liées à des intérêts divergents entrent en jeu, il n'est plus possible de les hiérarchiser. Pour Goldman, elles ont toutes la même valeur. Qui plus est, on peut remettre en doute sa conception de l'erreur selon laquelle celle-ci aurait une valeur égale et opposée à la vérité. Ainsi une pratique D qui mènerait à 9 croyances vraies et à 2 erreurs serait certainement préférable à une pratique E résultant en 5 croyances vraies et 3 erreurs. Mais serait-elle préférable à la pratique F qui conduit à 7 croyances vraies mais à seulement 1 erreur ? Ceci semble dépendre des conséquences de l'erreur, c'est-à-dire de considérations non épistémiques. Un médecin par exemple doit éviter l'erreur à tout prix lorsqu'il pose son diagnostic. « ... different people may legitimately assign different relative values to having true beliefs and avoiding error. Also, the same people may legitimately assign different relative values under different circumstances. »³⁸

« Thus, non-epistemic interests can help to determine which complete ordering of outcomes should guide the decisions of a person who seeks good epistemic outcomes. »³⁹ Une théorie des valeurs complète doit conséquemment permettre de bien différencier les croyances par une pesée des intérêts. « Goldman's theory does not really go far enough in this direction. The epistemic value of having a particular true belief should actually be weighted by how interested we are in the topic. »⁴⁰

7 Réponses au problème de la motivation des agents

La discussion a jusqu'ici rapproché Goldman des empiristes logiques. Celui-ci semble en effet donner, pour reprendre l'expression

³⁸ *Ibid.*, p. 189.

³⁹ *Id.*

⁴⁰ *Ibid.*, p. 191.

de Kitcher, une valeur presque déontologique à la vérité. Son approche sociale le confronte malgré tout au problème de la motivation des agents auquel il consacre une bonne partie de ses discussions pour diverses pratiques.

Cohérent dans son approche, il ne cherche pas à évaluer la force exacte des intérêts divergents qui sont en jeu. Au contraire, seuls les résultats véritistiques des pratiques l'intéressent. C'est donc la propension de ces pratiques à mener à la vérité qui est évaluée. Mais l'évaluation des motivations occupe malgré tout, de manière indirecte, une certaine place. C'est toutefois à l'agent placé dans le contexte de la pratique de faire cette évaluation. Ce point est particulièrement frappant dans l'analyse de la pratique du témoignage⁴¹. Lorsque vient le temps pour l'agent de déterminer s'il peut ou non se fier à un témoignage, il lui est nécessaire d'évaluer certaines probabilités dont, entre autres, la probabilité conditionnelle que le témoin témoigne que X si X s'est produit et la probabilité conditionnelle que le témoin témoigne que X si X ne s'est pas produit. Ces probabilités doivent être établies par l'agent, de manière subjective, et Goldman ne présente aucun critère permettant de l'aider dans cette tâche. Or comment évaluer ces probabilités conditionnelles sinon en évaluant les raisons que le témoin aurait de mentir, c'est-à-dire ses motivations ? C'est donc l'agent qui a la responsabilité de faire l'évaluation des motivations du témoin. Si ces probabilités sont bien évaluées, l'équation proposée (une inférence basée sur le théorème de Bayes) assurera à l'agent l'utilisant un résultat véritistique favorable. L'utilisation de cette équation est donc une bonne pratique véritistique et cette affirmation peut être faite indépendamment d'une évaluation directe des motivations des agents. Une évaluation indirecte est malgré tout présente.

Similairement, les scientifiques semblent avoir développé une méthodologie efficace pour juger et éliminer les motivations de leurs pairs. À nouveau, Goldman n'a pas les outils conceptuels nécessaires pour aborder cette question directement. Son analyse prend dès lors

⁴¹ Alvin GOLDMAN, *op. cit.*, 1999, chapitre 4, p. 103 à 130.

une tournure plus sociologique, psychologique ou historique⁴² (comme le prescrivaient les empiristes logiques). Il examine donc le problème de la motivation des agents en science en citant plusieurs recherches en psychologie.

Kunda, par exemple, insiste sur le fait que les agents peuvent être motivés à atteindre une conclusion correcte plutôt qu'une conclusion prédéterminée par un biais.

People motivated to arrive at a particular conclusion attempt to be rational and to construct a justification of their desired conclusion that would persuade a dispassionate observer. They draw the desired conclusion only if they can muster up the evidence necessary to support it. . . People will come to believe what they want to believe only to the extent that reason permits.⁴³

La pression des pairs est un bon exemple d'une motivation qui inciterait les scientifiques à être objectifs. Outre le fait qu'ils côtoient ces pairs et qu'un manque d'objectivité pourrait être socialement mal perçu (intérêt social), ceux-ci sont également ceux qui déterminent si le travail du chercheur mérite une publication, une promotion ou toute autre forme de reconnaissance professionnelle (intérêt professionnel). De plus, la reconnaissance de la qualité de leur travail est également ce qui permettra à ces scientifiques d'obtenir les subventions nécessaires à la poursuite de leurs recherches (intérêt financier). En conséquence, les scientifiques doivent constamment justifier leurs assertions devant leurs pairs s'ils veulent progresser.

Goldman confirme ce résultat dans une étude qu'il avait faite en collaboration avec Shaked⁴⁴. Il y construit un modèle mathématique permettant d'analyser comment des scientifiques choisissent quelles

⁴² *Ibid.*, p. 230 à 238.

⁴³ Ziva KUNDA, « The case for motivated reasoning », *Psychological Bulletin*, vol. 108, n°3, 1990, p. 482-483. cité par Alvin noextrasenglishGoldman, *op. cit.*, 1999, p. 236.

⁴⁴ Alvin GOLDMAN et Moshe SHAKED, « An economic model of scientific activity and truth acquisition », *Philosophical Studies*, vol. 63, no 1, 1991, p. 31-55.

expériences réaliser en prenant pour mesure l'utilité estimée maximale pour leur succès professionnel⁴⁵. Il démontre ainsi que, sous certaines conditions, des agents fortement biaisés par leurs intérêts professionnels peuvent atteindre des résultats presque aussi bons que ceux d'agents idéalement rationnels. « Thus, even if scientists are motivationally "sullied" (in Kitcher's phrase), the veritistic consequences of their research can approximate those of motivationally "pure" scientists. »⁴⁶

S'inspirant de Miriam Solomon⁴⁷, Goldman présente aussi la possibilité qu'un biais puisse être favorable à l'acquisition de croyances vraies. Solomon présente deux types de biais. L'exagération est une tendance à donner une importance exagérée à certaines données empiriques. La persévérance des croyances est quant à elle la tendance à conserver nos vieilles croyances même devant des observations tendant à les contredire. Ces deux biais étaient clairement présents chez les chercheurs qui défendaient l'idée de la dérive des continents avant son acceptation dans les années 1960. L'exagération les affectait en ceci que les données épistémiques qu'ils défendaient étaient celles qu'ils avaient eux-mêmes découvertes. Et cette même exagération les persuadait de persévérer dans leur croyance malgré les critiques. « Solomon contends that the net upshot for the scientific community was that more diverse research avenues were pursued and this distribution of research effort ultimately paid the community well. »⁴⁸

La conclusion à tirer pour notre discussion est importante. Aucune de ces réponses ne rejette la possibilité qu'un biais soit présent. Au contraire, il semble que l'agent doive être motivé à remettre en question ses biais pour pouvoir agir de manière « objective ». Ceci

⁴⁵ Notons qu'ici encore l'évaluation des motivations est laissée dans les mains des agents impliqués.

⁴⁶ Alvin GOLDMAN, « The Need for Social Epistemology », dans B. LEITER (Éd.), *The Future for Philosophy*, Oxford (NY), Clarendon Press, 2004, p. 197.

⁴⁷ Miriam SOLOMON, « Scientific Rationality and Human Reasoning », *Philosophy of Science*, vol. 59, no. 3, 1992, p. 439-455.

⁴⁸ Alvin GOLDMAN, *op. cit.*, 1999, p. 234.

confirme une critique générale de Maffie à propos des conclusions de Goldman : « ... all truth-seeking enterprises are inescapably infected with nonveritistic interests (narrowly construed) and therefore not objective in the traditional sense of objectivity as disinterestedness. »⁴⁹ L'objectivité n'étant plus qu'approximée, l'idéal d'un pur agent objectif n'a plus sa place.

8 Conclusion

Kelly, rappelons-le, semble avoir anticipé la critique en admettant la possibilité que le but de nos investigations soit autre que la vérité. De fait, son épistémologie computationnelle semble être aussi efficace à rendre compte du processus d'apprentissage d'un agent biaisé que de celui d'un agent rationnel.

Pour comprendre le processus d'apprentissage complet cependant, il faut, à l'instar de Kitcher, réaliser que l'acquisition des connaissances est inscrite dans un contexte plus large visant certaines fins.

Ces fins demeurent incomprises dans l'approche de Goldman, qui reste convaincu que la vérité est le seul but que doit considérer une analyse épistémologique. Il est conséquemment incapable d'expliquer notre réaction à certains conflits d'intérêts divergents. Pour Fallis, les fins sont au contraire d'une importance capitale et sont déterminantes dans tous les choix épistémiques des agents. D'où la nécessité de bien les comprendre.

L'approche de Goldman n'en démontre pas moins l'importance de l'intérêt des agents pour comprendre leur démarche épistémique. Dans une perspective sociale, les agents sont en effet constamment invités à évaluer les motivations des autres pour s'assurer d'un progrès véritistique. Les sciences utilisent quant à elles ces motivations pour promouvoir une certaine forme d'objectivité.

Une description du processus d'acquisition des connaissances qui ne tiendrait pas compte du rôle des intérêts des agents serait incapable de comprendre ces points importants et n'offrirait donc qu'un

⁴⁹ James MAFFIE, *op. cit.*, p. 253.

portrait partiel du véritable processus. Toute description de l'activité épistémique qui se veut complète se doit donc de prendre en compte le rôle que jouent les différents intérêts des agents.

BIBLIOGRAPHIE

- FALLIS, Don, « Epistemic Value Theory and Social Epistemology », dans *Episteme*, vol. 2, no 3, 2006, p. 177-188.
- GILOVICH, Thomas, *How we know what isn't so : The fallibility of human reason in everyday life*, New York., Free Press, 1991.
- GOLDMAN, Alvin et SHAKED, Moshe, « An economic model of scientific activity and truth acquisition », dans *Philosophical Studies*, vol. 63, no 1, 1991, p. 31-55.
- GOLDMAN, Alvin, *Knowledge in a social world*, New York, Clarendon Press / Oxford University Press, 1999.
- GOLDMAN, Alvin, « Replies to reviews of 'Knowledge in a Social World' », dans *Social Epistemology*, vol. 14, no 4, 2000, p. 317-333.
- GOLDMAN, Alvin, « The Need for Social Epistemology », dans B. LEITER (Éd.), *The Future for Philosophy*, Oxford (NY), Clarendon Press, 2004, 182-207.
- HENDRICKS, Vincent, *Mainstream and formal epistemology*, New York, Cambridge University Press, 2006.
- KELLY, Kevin et Olivier SCHULTE, « The computable testability of theories making uncomputable predictions », dans *Erkenntnis*, vol. 43, no 1, 1995, p. 29-66.
- KELLY, Kevin, « Iterated Belief Revision, Reliability, and Inductive Amnesia », dans *Erkenntnis*, vol. 50, no 1, 1999, p. 7-53.
- KELLY, Kevin, « The logic of success », dans *The British Journal for the Philosophy of Science*, vol. 51, no 4, 2000, p. 639-666.
- KITCHER, Philip, « Veritistic Value and the Project of Social Epistemology », dans *Philosophy and Phenomenological Research*, vol. 64, no 1, 2002, p. 191-198.
- KITCHER, Philip, « The Ends of the Sciences », dans B. LEITER (Éd.), *The Future for Philosophy*, Oxford (NY), Clarendon Press, 2004, 208-229.

- KUNDA, Ziva, « The case for motivated reasoning », dans *Psychological Bulletin*, vol. 108, no 3, 1990, p. 480-498.
- MAFFIE, James, « Alternative epistemologies and the value of truth », dans *Social Epistemology*, vol. 14, no 4, 2000, p. 247-257.
- SCHMITT, Frederick, « Veritistic value », dans *Social Epistemology*, vol. 14, no 4, 2000, p. 259-280.
- SCHULTE, Oliver, « Formal Learning Theory », dans ZALTA, E. N (Éd.), *Stanford Encyclopedia of Philosophy (Spring 2006 Edition)*, URL = <http://plato.stanford.edu/archives/spr2006/entries/learning-formal/>, 2002.
- SOLOMON, Miriam, « Scientific Rationality and Human Reasoning », dans *Philosophy of Science*, vol. 59, no 3, 1992, p. 439-455.