

En 1804, la théorie du phlogistique était morte. Trente ans plus tôt, elle avait été la théorie préférée de presque tous les chimistes d'Europe. Si ce sont la raison et les preuves qui ont décidé de l'issue de la révolution chimique, alors il semble qu'il ait dû y avoir un moment entre 1774 et 1804 où la balance des preuves a finalement penché contre la théorie du phlogistique et en faveur de la « nouvelle chimie » d'Antoine-Laurent Lavoisier.

Imaginons que le degré objectif de confirmation de la théorie du phlogistique juste avant midi le 23 avril 1787 ait été de 0,51, et celui de la nouvelle chimie de 0,49. À midi, Lavoisier réalisa une expérience importante, et les degrés de confirmation passèrent respectivement à 0,49 et 0,51. En tenant compte d'un décalage temporel dans la dissémination de l'information critique, nous pouvons estimer qu'il y a eu un intervalle relativement court après le 23 avril 1787 à midi avant lequel tous les chimistes rationnels étaient adeptes de la théorie du phlogistique et après lequel ils étaient tous partisans de Lavoisier.

Est-ce que ce scénario consistant en une opinion initialement uniforme, un soudain changement de camp et un nouveau consensus est le signe d'une croissance rationnelle de la connaissance scientifique ? Si vous aviez été un monarque-philosophe, soucieux que vos sujets-scientifiques répartissent leurs efforts afin que la communauté finisse par atteindre la vérité, vous auriez (à raison) considéré que cette affectation des ressources (les scientifiques eux-mêmes) n'était pas une bonne affaire. Avec un si mince écart de preuves entre les deux théories, vous auriez préféré que certains scientifiques n'aient pas été aussi clairvoyants dans leur perception du mérite des théories, de sorte que l'unanimité de leurs décisions ne soit pas atteinte si tôt. Une communauté de chimistes se comportant de la façon décrite dans mon histoire initiale est une communauté mal gérée – une communauté irrationnelle, si vous préférez.

Mon histoire a pour but de soulever un problème important, bien que négligé¹,

* Je suis reconnaissant envers les nombreuses personnes qui ont entendu ou lu des versions précédentes de cet article et m'ont donné de précieux conseils, et en particulier à John Beatty, Gerald Doppelt, Isaac Levi, David Lewis, et Elizabeth Lloyd. J'ai une grande dette envers les idées et les écrits de Thomas Kuhn. Des remerciements particuliers vont à Stephen Stich pour ses nombreux commentaires détaillés et ses suggestions constructives.

1 Le problème est posé par Thomas Kuhn dans « Objectivity, Value Judgment and Theory Choice » [in *The Essential Tension* (Chicago : Chicago University Press, 1977), p. 320-339 ; trad. fra. *La Tension Essentielle* (Paris : Gallimard, 1990), « Objectivité, jugement de valeur et choix d'une théorie », p. 424-449]. Les propositions de Husain Sarkar dans *A Theory of Method* (Berkeley : California University Press, 1983), et mes propres remarques

concernant la croissance de la science. Est-il possible qu'il y ait un désaccord entre les exigences de la rationalité individuelle et celles de la rationalité collective (ou de la communauté) ? Pourrait-il s'avérer que des chercheurs qui suivent les nobles principes de la rationalité individuelle promeuvent mal les projets épistémiques de la communauté à laquelle ils appartiennent ? Se pourrait-il que ceux ayant des motifs plus terre-à-terre contribuent en fait davantage aux efforts épistémiques de leur communauté ? Y a-t-il des conditions dans lesquelles, étant donné nos buts en tant que communauté épistémique, nous devrions vouloir maintenir une diversité cognitive ? Y a-t-il quelque chose que nous faisons, ou pouvons faire, à ce sujet ?

I

Peut-être la situation malencontreuse que je décris est-elle seulement un artefact résultant de présuppositions erronées. Les brèves remarques suivantes ont pour objet de répondre à certaines inquiétudes évidentes et d'indiquer que les problèmes que j'ai soulevés survivent aux modifications du cadre que j'ai employé pour les présenter.

Toute suggestion optimiste selon laquelle le genre de différence subtile que j'ai imaginée est irréaliste parce que les décisions scientifiques sont toujours bien tranchées est démentie par les études récentes de l'histoire des sciences. De plus, même là où les degrés de confirmation diffèrent grandement, des désaccords entre la rationalité individuelle et collective peuvent encore survenir. Une objection plus prometteuse à mon histoire est que nous ne pouvons raisonnablement pas parler de valeurs numériques représentant la confirmation empirique d'une théorie. Dans une version plus modérée, les degrés de confirmation peuvent être représentés par des sous-intervalles connectés de $[0 ; 1]$. Mais il est facile de construire des versions de mon histoire initiale dans lesquelles nous préférons qu'une minorité de scientifiques continuent d'embrasser une théorie dont le degré objectif de confirmation est donné par le sous-intervalle $[0 ; 0,5]$.

Un problème plus sérieux apparaît lorsqu'on nie l'existence de mesures objectives de confirmation. Les travaux minutieux d'histoire des sciences ne nous révèlent-ils pas qu'il y a de nombreux cas où des personnes, aussi raisonnables les unes que les autres, peuvent être en désaccord à propos des mérites de théories rivales, peut-être parce qu'elles ont des idées différentes de l'importance de différents problèmes, ou des critères appropriés pour résoudre ces problèmes ? Dans cet essai, je ne souhaite pas prendre parti sur cette question controversée. Je prétends seulement qu'il y a des moments où nous voulons maintenir la diversité cognitive même dans des situations où il serait raisonnable pour tout le monde de s'accorder sur le fait que l'une des deux

très sommaires au ch. 6 de *Abusing Science* (Cambridge : MIT, 1982), sont pratiquement les seuls traitements ultérieurs de la possibilité de désaccords entre la rationalité individuelle et collective. J'ai beaucoup appris du traitement du sujet par Sarkar, bien qu'il s'intéresse aux alternatives concernant la méthodologie, plutôt qu'aux différences de théories, de programmes de recherches, ou de méthodes.

théories est inférieure à sa rivale, et que nous pourrions être reconnaissants envers la minorité bornée qui continue à défendre des idées problématiques.

Dans les années 1790, seuls une poignée de chimistes continuaient à envisager la possibilité que la théorie du phlogistique puisse revivre. Comme Thomas Kuhn le remarque², il fut probablement déraisonnable de la part de Joseph Priestley de persévérer aussi longtemps qu'il le fit. Cependant, du point de vue de la communauté des chimistes, il n'était pas une mauvaise chose que Priestley et quelques autres aient donné à la théorie du phlogistique une toute dernière chance. Plus récemment, avec l'histoire de la théorie de la dérive des continents d'Alfred Wegener, nous pouvons avoir une idée de la tournure différente qu'auraient pu prendre les événements. Dans les années 1920 et 1930, les thèses de Wegener semblaient rencontrer des difficultés insurmontables, car il y avait des démonstrations géophysiques apparemment rigoureuses selon lesquelles les forces nécessaires pour déplacer les continents auraient dû être bien plus grandes que ce qui était possible. En dépit de cela, quelques géologues, et tout particulièrement Alexander du Toit, continuèrent à défendre les idées de Wegener. Je suggère que cette répartition des efforts cognitifs était préférable à la situation dans laquelle même la petite minorité aurait abandonné la dérive des continents³.

Était-il aussi raisonnable d'être pour la dérive que contre elle dans les années 1920 et 1930 ? Ayant appris à goûter les changements complexes des critères de confirmation qui surviennent en histoire des sciences, peut-être répondez-vous « oui ». Mais alors vous faites face à un problème de maintien de la diversité cognitive du même type que celui avec lequel j'ai commencé : du point de vue de la communauté, il aurait été préférable que les géologues aient été répartis de façon plus égale. D'un autre côté, si vous acceptez l'idée que les arguments géophysiques révélaient véritablement l'in vraisemblance de la théorie de Wegener, alors la répartition réelle de l'effort cognitif paraît être meilleur – même si, bien sûr, certains agents épistémiques, comme du Toit, sont considérés comme faisant un choix irrationnel.

Considérons une dernière suggestion permettant d'éviter les problèmes concernant la division du travail cognitif. Nous pouvons certainement distinguer plusieurs attitudes que les scientifiques adoptent envers les théories, les hypothèses, les programmes de recherche, etc. En particulier, nous pouvons distinguer la *croissance* en une théorie de la *poursuite* de recherches ayant pour but d'appliquer ou d'étendre cette théorie⁴. Une fois cette distinction reconnue, ne pouvons-nous pas accepter une

2 *The Structure of Scientific Revolutions* (Chicago : University Press, 1962 ; seconde édition 1970), p. 159 ; trad. fra. *La Structure des Révolutions Scientifiques* (Paris : Flammarion, 1983), p. 218.

3 Pour une vue d'ensemble de la carrière de la théorie de Wegener, voir Anthony Hallam, *A Revolution in the Earth Sciences* (New York : Oxford, 1973). La persistance de du Toit à défendre la théorie est manifeste dans son livre, *Our Wandering Continents* (Edinburgh : Oliver & Boyd, 1937).

4 Cf. Larry Laudan, *Progress and Its Problems* (Berkeley : California University Press,

solution simple à mon problème ? Alors qu'il pourrait être rationnel pour chacun des scientifiques de croire en la théorie qui est la mieux confirmée par les preuves disponibles, il pourrait ne pas être rationnel pour chacun d'eux de poursuivre des recherches avec cette théorie, et ce qui intéresse la communauté est la répartition des poursuites, non des croyances.

Cette suggestion permettant d'éviter le désaccord entre la rationalité individuelle et collective suppose que l'on adopte deux principes de rationalité individuelle, un pour la croyance et un autre pour la poursuite. L'idée qu'il soit rationnel pour une personne de croire en la théorie la mieux confirmée semble cependant supposer que le but de cette personne est d'obtenir des croyances vraies (ou quelque autre état épistémique désirable, comme l'acceptation de théories empiriquement adéquates). Dans ce cas, il semble pourtant que la personne devrait aussi poursuivre ses recherches avec la théorie la mieux confirmée, puisque suivre une doctrine qui est susceptible d'être fautive risque d'engendrer davantage d'erreurs (ou de conduire davantage à l'état épistémique indésirable). C'est seulement lorsque nous situons l'individu dans une société d'autres agents épistémiques – comme je vais tenter de le faire dans les sections suivantes – qu'il commence à sembler rationnel qu'une personne puisse choisir de travailler sur des idées qu'elle-même et ses collègues considèrent comme étant épistémiquement inférieures.

II

Imaginez, par conséquent, que vous soyez un monarque-philosophe ayant la prérogative de diriger le cours de la recherche scientifique. Vous espérez atteindre certains buts épistémiques – construire une histoire du monde vraie et complète, formuler une théorie empiriquement adéquate, éliminer l'erreur, résoudre autant de problèmes que possible, ou quoi que ce soit d'autre. Vous avez un œil infallible pour détecter les mérites objectifs des théories et un contrôle total de la main d'œuvre scientifique. Quelle règle de division du travail cognitif devriez-vous adopter ?

Peut-être décidez-vous de permettre à vos sujets d'être individuellement rationnels, en les autorisant à croire en la théorie T dès lors que le mérite épistémique objectif de T est supérieur à celui de ses rivales (en exigeant éventuellement que le mérite de T soit supérieur à un certain seuil). Mais, ainsi que nous l'avons vu, cela est une mauvaise stratégie, qui risque de conduire à une opinion uniforme alors que vous voudriez ne pas vous fermer de portes.

Même sans la fiction du monarque-philosophe, nous pouvons toujours considérer le problème de la stratégie optimale de la communauté en vue d'atteindre des buts épistémiques. En continuant d'être vague sur ce que sont ces buts épistémiques, je formulerai le problème en distinguant deux types d'intentions épistémiques que les scientifiques peuvent individuellement avoir. X peut avoir l'intention que X atteigne certains buts épistémiques (autant que possible) : ainsi, par exemple, X peut avoir

1977), p. 108-114.

l'intention que X acquière autant de croyances vraies que possible. Ceci est l'intention épistémique *personnelle* de X . X peut aussi avoir l'intention que la communauté à laquelle X appartient, la communauté des scientifiques passés, présents et futurs, atteigne un but épistémique (autant que possible). Cela est l'intention épistémique *impersonnelle* de X . La question « quelle est la stratégie rationnelle de la communauté ? » peut être reformulée ainsi : « comment les scientifiques décideraient-ils rationnellement de coordonner leurs efforts si leurs décisions étaient régies par leurs intentions épistémiques impersonnelles ? ». La fiction du monarque-philosophe dramatise l'idée que cette décision pourrait nécessiter la subordination des intentions épistémiques personnelles. Si X s'engage dans un projet de la communauté avec l'objectif que la communauté dans son ensemble atteigne un certain but épistémique – un but que tous les collègues de X veulent aussi atteindre – alors il se pourrait que X doive prendre des décisions qui ne coïncident pas avec celles d'une scientifique individuellement rationnelle. X devrait accepter par avance qu'il puisse être parfois nécessaire qu'un ou plusieurs membres de la communauté adoptent pour leurs recherches une théorie inférieure (ou même y croient), et que ce rôle pourrait incomber à X .⁵

Je peux désormais donner une description générale de la classe des problèmes qui concernent la division optimale du travail cognitif. Supposons qu'il existe un ensemble S de scientifiques, chacun d'entre eux ayant le choix entre les éléments d'un ensemble R d'objets cognitifs rivaux. (R peut être un ensemble de théories, de programmes de recherche, de méthodes rivales pour aborder un problème, etc.) Chaque scientifique a une intention épistémique impersonnelle que la communauté issue de S atteigne un certain état épistémique visé ; supposons que l'intention de la scientifique i soit que la communauté atteigne G_i (les G_i pouvant être différents). Pour chaque scientifique, il existe une fonction d'évaluation, dont le domaine est R . Ces fonctions, qui peuvent être distinctes, évaluent les mérites épistémiques des éléments de R , et je ferai l'hypothèse que la rationalité épistémique individuelle exige que chaque scientifique adopte une certaine attitude cognitive (convenant à la catégorie d'objets appartenant à R) envers l'élément de R qui est le mieux classé selon la fonction d'évaluation de la scientifique. (Donc, si les éléments de R sont des théories, l'attitude cognitive peut être la croyance, et l'exigence de rationalité peut être que les scientifiques croient en la théorie de R qui est la meilleure selon leur fonction d'évaluation.) La répartition d'attitudes IR (individuellement rationnelle) est la répartition engendrée parmi les membres de S à partir des fonctions d'évaluation en accord avec l'exigence de rationalité. La répartition OC (optimale pour la communauté) relative à i est la répartition d'attitudes parmi les membres de S qui maximiserait la probabilité d'atteindre G_i . Il existe un désaccord OC-

5 Les scientifiques rationnels de façon altruiste sont ceux qui sont prêts à mener des recherches avec une théorie qu'ils considèrent comme inférieure lorsque, ce faisant, ils promeuvent la réalisation des buts de leurs intentions épistémiques impersonnelles, et de leurs collègues. Cela érige clairement un idéal bien plus extrême que celui qui est critiqué par les historiens et les sociologues des sciences.

IR lorsqu'il existe une répartition d'attitudes parmi les membres de S qui, pour chaque i , conduit à une probabilité plus grande d'atteindre G_i qu'avec la répartition IR.

Ma formulation générale autorise des différences d'intentions impersonnelles des scientifiques et des différences quant à leur évaluation du mérite épistémique, ce qui permet d'insister sur le fait (cf. section I) que les problèmes concernant la division du travail cognitif surviennent même dans l'hypothèse où il y a des changements de critères d'évaluation et des changements de buts. Les problèmes sont toutefois plus simples à poser et plus faciles à résoudre si nous supposons l'uniformité sur ces deux points, c'est-à-dire si nous supposons que les G_i sont tous les mêmes et qu'il y a une seule fonction d'évaluation (« objective ») pour tous les scientifiques. De même que j'ai commencé en faisant implicitement cette hypothèse simplificatrice, je continuerai de l'adopter dans la suite de cet essai. Le problème n'en serait que plus compliqué si cette hypothèse devait être abandonnée.

Étant donné notre simplification, nous pouvons parler d'un état G visé par la communauté et d'une répartition OC qui n'est pas relative à une scientifique i en particulier. Un désaccord OC-IR sera une situation dans laquelle la répartition OC est différente de la répartition IR. Je veux maintenant abandonner l'abstraction des derniers paragraphes pour considérer un type particulier de problème de division du travail cognitif, afin d'identifier certaines conditions dans lesquelles on peut s'attendre à ce que surviennent des désaccords OC-IR. Bien que mon exemple initial concernait le choix entre des théories, les deux prochaines sections exploreront la possibilité de désaccords OC-IR dans le cas, plus facile à résoudre, du choix entre des méthodes de résolution de problèmes. Je reviendrai au cas du choix entre des théories (qui est algébriquement plus complexe) dans la section V.

III

Il y avait autrefois une molécule très importante (MTI). De nombreuses personnes de la communauté chimiste voulaient connaître la structure de cette MTI. Deux méthodes étaient disponibles pour percer le mystère de cette structure. La méthode I supposait d'utiliser de la cristallographie par rayons X et d'analyser les photographies obtenues pour éliminer certains motifs de liaisons possibles. La méthode II supposait de deviner au hasard et de bricoler des modèles. Tout le monde estimait que les chances qu'un individu découvre la structure de la MTI en utilisant la méthode I étaient bien plus grandes que les chances qu'il la découvre en utilisant la méthode II. Puisque tous les membres de la communauté étaient absolument rationnels, tous les chimistes adoptèrent la méthode I. Ils travaillent encore sur le problème.

Le but de la communauté est de percer le mystère de la structure de la MTI aussi rapidement que possible. Supposons qu'à chaque méthode soit associée une fonction de probabilité $p(n)$ qui représente la chance que la méthode fournisse une réponse si n travailleurs y sont affectés. Supposons aussi qu'il soit possible de reconnaître si une réponse proposée est vraie ou fausse. Imaginons également que les relations entre les

fonctions de probabilité représentent leur comportement sur tout intervalle de temps considéré – donc pour tout intervalle de temps t , les probabilités que la méthode I donne une réponse dans le temps t et que la méthode II donne une réponse dans t sont dans le même rapport que les fonctions p .⁶ N travailleurs sont disponibles pour être répartis sur les deux méthodes. La répartition OC est donnée par n travailleurs utilisant la méthode I et $N - n$ utilisant la méthode II, afin de maximiser la probabilité que la structure de la MTI soit découverte, c'est-à-dire afin de maximiser

$$p_1(n) + p_2(N - n) - Prob(\text{les deux méthodes donnent une réponse}).$$

Par soucis de simplification, je supposerai que la probabilité que les deux méthodes donnent la bonne réponse est nulle. (L'effet de cette hypothèse non-triviale est seulement de simplifier les calculs algébriques. Des conclusions qualitativement similaires peuvent être obtenues, bien que plus difficilement, si elle n'est pas faite.)

La solution du problème dépend évidemment de la forme des fonctions $p(n)$. Je les appellerai des *fonctions de retour*, puisqu'elles mesurent la probabilité que le but soit atteint, en retour d'un investissement de n travailleurs. Je considérerai que ces fonctions sont sujettes aux contraintes suivantes : elles doivent augmenter de façon monotone avec n , elles doivent valoir zéro lorsque n vaut zéro, et elles doivent tendre asymptotiquement vers une certaine valeur p lorsque n tend vers l'infini (p représente les promesses intrinsèques de la méthode, sa probabilité de succès sans tenir compte du fait que les efforts humains sont limités)⁷. Étant donné l'hypothèse simplificatrice selon laquelle les deux méthodes ne donneront pas ensemble la bonne réponse, nous savons que les valeurs des asymptotes, p_1 et p_2 , ont une somme inférieure à 1. Ces contraintes autorisent un large choix de fonctions. Pour être plus explicite, il est tout à fait possible que la forme des fonctions soit différente pour les deux méthodes. (Imaginons qu'une méthode réponde bien plus rapidement que l'autre aux efforts des travailleurs.)

Je considérerai deux possibilités pour ces fonctions⁸. Supposons tout d'abord que

-
- 6 Ces fonctions mesurent la probabilité qu'une méthode donne une réponse correcte pour une affectation de n travailleurs, étant donné que le monde est tel que la communauté se le représente, avec sa connaissance actuelle de la molécule. Ainsi, par exemple, si la MTI est peu connue et si les deux méthodes ont été adoptées par des nombres similaires de travailleurs pour étudier une grande classe de molécules, la méthode I obtenant plus fréquemment du succès que la méthode II, quel que soit le nombre de travailleurs affectés, $p_1(n) > p_2(n)$. Je n'approfondirai pas ici la question de savoir exactement comment interpréter les probabilités. Autant que je sache, il est possible de les concevoir selon n'importe laquelle des façons actuellement populaires.
- 7 Je pensais initialement que ces contraintes s'appliqueraient dans tous les cas. Mais, comme Stephen Stich me l'a fait remarquer, il peut arriver que trop de cuisiniers gâtent la sauce. Imaginons, par exemple, que la méthode implique l'observation d'organismes sensibles et qu'une foule dérange le comportement normal des organismes.
- 8 Les deux classes de fonctions que je considère représentent deux principaux cas de figure : ou bien le taux de retour est rapide au début, et ralentit ensuite en s'approchant de l'asymptote, ou bien le taux est initialement lent, accélère lorsqu'une masse critique de travailleurs est atteinte, et ralentit lorsqu'on approche de la saturation.

$p_i(n) = p_i \cdot (1 - e^{-kn})^\dagger$. Dans ce cas, $p_1(n) + p_2(N - n)$ est maximum lorsque

$$n = (kN + \ln p_1 - \ln p_2)/2k.$$

Notons que, même lorsque la méthode I est intrinsèquement plus prometteuse que la méthode II ($p_1 > p_2$), il existe une gamme de conditions (lorsque $\ln p_1 - \ln p_2 < kN$) pour lesquelles la répartition OC consiste à diviser la communauté. Intuitivement, une authentique division du travail cognitif sera préférable pour la communauté si la main d'œuvre disponible est nombreuse (N est grand), ou si les méthodes répondent rapidement à l'effort fourni (k n'est pas trop petit), ou si la différence entre les promesses intrinsèques des méthodes n'est pas trop grande (p_1 et p_2 sont suffisamment proches). L'inégalité ci-dessus représente la façon dont les compromis doivent être faits.

Les fonctions considérées dans le paragraphe précédent ont la propriété que le taux d'augmentation de $p(n)$ est maximum lorsque n est petit. Il se pourrait que cette idée ne soit pas réaliste du tout. Peut-être que les chances d'obtenir une réponse en suivant une méthode donnée augmentent très lentement au début, puis augmentent rapidement une fois qu'une masse critique de travailleurs est atteinte, et enfin augmentent très lentement lorsqu'on approche de la saturation. Nous pouvons décrire ce comportement en imitant l'équation de croissance logistique en biologie des populations, et en faisant l'hypothèse que les p_i sont donnés par

$$p_i(n) = p_i \cdot (3n^2 - 2n^3/kN)/k^2N^2 \quad (n < kN),$$

$$p_i(n) = p_i \quad (n \geq kN).$$

Si les probabilités sont données par ces fonctions, alors il y a différents cas intéressants, selon la valeur de k . Dès lors que $k < 1/2$, il est possible de réaliser les promesses intrinsèques des deux méthodes, et par conséquent la répartition OC consiste en une division de la main d'œuvre. Si $1/2 < k < 1$, il n'est pas difficile de montrer que la valeur optimale de n est inférieure à kN . Lorsque k est supérieur ou égal à 1, n devrait valoir N (rappelons que la méthode I est meilleure, c'est-à-dire que $p_1 > p_2$).

Voici l'interprétation qualitative que je propose de ces résultats. Comme dans le cas précédent, k est un paramètre critique, représentant la sensibilité des méthodes. Si les méthodes sont si sensibles que les promesses intrinsèques des deux méthodes peuvent être réalisées avec la main d'œuvre disponible, alors il est facile de voir que l'intérêt épistémique de la communauté est que la main d'œuvre soit divisée. Même lorsque k est entre $1/2$ et 1 – de telle sorte qu'il est possible de réaliser les promesses intrinsèques d'une des méthodes mais pas des deux – il peut être préférable de diviser la main d'œuvre afin que les promesses d'aucune méthode ne soient réalisées. Dès lors que la différence entre p_1 et p_2 n'est pas trop grande, il sera préférable d'affecter un nouveau travailleur à la méthode II, si la méthode II a suffisamment d'adeptes pour offrir un grand retour sur investissement, plutôt qu'à la méthode I, si la méthode I est proche de la saturation. Une fois que k atteint 1, il est cependant toujours préférable

† Dans cette traduction, le produit de la variable p_i avec la parenthèse x est noté « $p_i(x)$ », afin de lever l'ambiguïté avec « $p_i(x)$ » qui désigne la fonction p_i évaluée en x . (NdT).

d'affecter toutes les ressources à la méthode dont les promesses intrinsèques sont les plus grandes.

Venons-en à la répartition IR dans ces cas. Selon une certaine compréhension simple de ce qu'est la rationalité épistémique individuelle, les agents rationnels jugent les méthodes selon les qualités intrinsèques de ces méthodes, et non pas selon ce que leurs collègues font. Si nous entendons la rationalité épistémique individuelle en ce sens simple, alors il est facile de voir qu'il peut exister des désaccords entre la répartition IR ($\langle N, 0 \rangle$) et la répartition OC (qui est parfois $\langle n, m \rangle$ avec n et m tous deux non-nuls).

Peut-être devrions-nous comprendre la rationalité épistémique individuelle de façon un peu différente. Supposons que ce soit une exigence de la rationalité épistémique individuelle qu'un agent maximise ses chances d'adopter une méthode qui apporte la réponse. Nous pouvons interpréter cette exigence de deux façons. (1) Nous imaginons que l'agent prend une décision dans l'ignorance complète de ce que les autres membres de la communauté font, de sorte que la tâche est de choisir i afin que $p_i(1)$ soit le plus grand possible. (2) Nous imaginons que l'agent connaît la répartition actuelle $\langle r, s \rangle$, de sorte que la méthode I doit être choisie au cas où $p_i(r+1) > p_2(s+1)$. Avec l'une ou l'autre de ces interprétations, il est facile que des désaccords OC-IR se produisent.

À ce stade, il est évidemment possible de modifier les exigences de la rationalité épistémique individuelle en les redéfinissant de telle sorte que les désaccords s'évanouissent comme par magie : il suffit de déclarer qu'un agent individuellement rationnel est une personne qui choisit d'appartenir à la communauté dans laquelle les chances de découvrir la réponse correcte sont maximales. Je pense que c'est une vertu de l'analyse que j'ai présentée de mettre au grand jour cet idéal altruiste de rationalité – un idéal qui me semble être assez éloigné des concepts de rationalité qui figurent dans les discussions philosophiques traditionnelles. Mais que nous décrétons ou non l'identité entre la rationalité de la communauté et de la rationalité de l'individu, mon intérêt porte sur les propriétés des répartitions OC et sur la possibilité que des personnes réelles et imparfaitement altruistes puissent les approximer. La section suivante va explorer la possibilité selon laquelle le fait de permettre à nos scientifiques de s'écarter des nobles objectifs de la rationalité individuelle (et d'agir avec des motifs plus terre-à-terre) puisse en réalité aider le projet de la communauté.

IV

En plongeant dans des détails algébriques, nous avons quitté une communauté de chimistes qui s'efforçaient sans succès de percer le mystère de la MTI. Je ferai l'hypothèse que la répartition OC les concernant impliquait une authentique division du travail (correspondant à l'un des cas considérés dans la section précédente dans lesquels $p_1 > p_2$, $p_1(1) > p_2(1)$). Ils ne réussirent pas à la réaliser, puisque chacun d'eux suivait un des principes de la rationalité individuelle qui conduisaient à la répartition

$\langle N, 0 \rangle$. De plus, comme le mystère de la structure de la MTI pouvait être percé seulement avec la méthode II, leur incapacité à répartir les risques, en tant que communauté, s'avérait coûteuse.

À l'inverse, dans une nation voisine, la communauté chimiste se composait d'égoïstes sans scrupules. Chacun des membres de cette communauté prenait ses décisions rationnellement, au sens où les actions choisies étaient celles qui maximisaient les chances de réaliser les buts, mais les buts étaient personnels et non épistémiques. Ceux qui choisissaient de travailler sur la MTI le faisaient parce qu'ils pensaient que celui qui découvrirait la structure de la MTI gagnerait un prix très convoité. Faisons l'hypothèse simplificatrice (mais non dénuée de plausibilité) que, si une méthode réussit, alors toutes les personnes qui ont adopté cette méthode ont la même chance de gagner le prix. Comment devrions-nous nous attendre à ce que la communauté hobbesienne répartisse ses efforts ?

Imaginez que la communauté ait atteint une répartition $\langle n, N - n \rangle$. Vous êtes une scientifique travaillant actuellement avec la méthode I, et vous envisagez la possibilité de l'abandonner pour la méthode II. Le changement serait bon pour vous – étant donné mon hypothèse sur vos intérêts et vos aspirations – s'il augmentait la probabilité que vous gagniez le prix. Votre probabilité de gagner est la probabilité que quelqu'un dans votre groupe gagne, divisée par le nombre de membres du groupe. (On peut dire intuitivement qu'en choisissant une méthode vous participez à une loterie qui a une certaine probabilité de faire gagner, probabilité qui dépend du nombre de détenteurs de tickets ; votre chance de gagner est la probabilité que la loterie fasse gagner divisée par le nombre de tickets.) Ainsi, à $\langle n, N - n \rangle$, il sera de l'intérêt d'une scientifique de passer de la méthode I à la méthode II si

$$p_2(N - n + 1)/(N - n + 1) > p_1(n)/n.$$

Pour comprendre comment nos agents hobbesiens imaginaires se répartiront, il nous faut découvrir les équilibres, c'est-à-dire les répartitions dans lesquelles personne n'a intérêt à changer de méthode⁹. Nous dirons que la répartition $\langle n, N - n \rangle$ est *stable vers le bas* si $p_1(n)/n$ est supérieur ou égal à $p_2(N - n + 1)/(N - n + 1)$, et *instable vers le bas* sinon. De même, $\langle n, N - n \rangle$ est *stable vers le haut* si $p_1(n + 1)/(n + 1)$ est inférieur ou égal à $p_2(N - n)/(N - n)$, et *instable vers le haut* sinon. $\langle n, N - n \rangle$ est *bilatéralement stable* si elle est à la fois stable vers le haut et vers le bas.

Si une communauté de scientifiques hobbesiens atteint une répartition qui est bilatéralement stable, alors nous pouvons nous attendre à ce qu'elle y reste. Cependant, *être stable* est une chose, et *être atteignable* en est une autre. Même si une répartition particulière pouvait être conservée une fois atteinte, elle pourrait s'avérer inatteignable par un groupe de scientifiques égoïstes. Pour toute répartition $\langle n, N - n \rangle$ qui est

9 Ma réflexion sur le problème de l'évolution de la répartition des efforts dans les communautés scientifiques a été fortement influencée par la discussion classique de R. A. Fisher concernant l'évolution de la répartition des sexes [cf. *The Genetical Theory of Natural Selection* (New York : Dover, 1958), p. 158-160] et par les idées de John Maynard Smith [notamment *Evolution and the Theory of Games* (New York : Cambridge, 1982)].

bilatéralement stable, nous pouvons définir sa *zone d'attraction* comme l'ensemble des répartitions qui conduisent à $\langle n, N-n \rangle$. Plus précisément, nous dirons que $\langle m, N-m \rangle$ conduit vers le haut à $\langle n, N-n \rangle$ si $m < n$ et si pour tout x tel que $m < x < n$, $\langle x, N-x \rangle$ est instable vers le haut (et de même pour « conduit vers le bas »). $\langle n, N-n \rangle$ est *atteignable* si sa zone d'attraction comprend toutes les répartitions.

Il se pourrait que la communauté hobbesienne fonctionne bien mieux que les esprits aux nobles principes de la section précédente qui échouaient à se répartir le travail. Plus précisément, il existe peut-être une répartition qui est à la fois stable et atteignable et qui offre une plus grande probabilité de réussite pour la communauté que les répartitions IR considérées plus haut. (L'idéal, bien sûr, serait de montrer que la répartition OC est à la fois stable et atteignable.) Les facteurs dont on pense fréquemment qu'ils interfèrent avec la conduite rationnelle de la science – la soif de gloire et de fortune, par exemple – pourraient précisément être ceux qui jouent un rôle constructif dans les projets épistémiques de notre communauté, nous permettant en tant que groupe de faire bien mieux que ce que nous aurions fait si nous nous étions comportés comme des individus épistémiquement rationnels et indépendants. Pour exprimer la morale un peu différemment, les institutions sociales de la science pourraient tirer parti de nos travers personnels en orientant nos efforts vers les buts de la communauté, plutôt que vers les buts épistémiques que nous pourrions individuellement nous donner.

Mais est-ce que cette possibilité existe réellement ? Considérons différents cas. Le plus simple est celui dont les fonctions de retour sont données par

$$p_i(n) = p_i \cdot (1 - e^{-kn}) \text{ avec } k \text{ grand et } p_1 > p_2.^{10}$$

Il existe une répartition bilatéralement stable au voisinage de $\langle n^*, N - n^* \rangle$, avec $n^* = p_1 \cdot N / (p_1 + p_2)$. La répartition est atteignable. De plus, si p_1 est seulement légèrement plus grand que p_2 , la répartition donne une probabilité de succès pour la communauté qui est proche de celle donnée par la répartition OC. Morale : il existe des conditions dans lesquelles les hobbesiens font mieux que leurs cousins épistémiquement purs, et même des conditions dans lesquelles ils sont aussi proches de l'idéal que vous pouvez le souhaiter.

La vie est plus compliquée si les fonctions de retour prennent les formes

$$p_i(n) = p_i \cdot (3n^2 - 2n^3/kN) / k^2 N^2 \text{ pour } n < kN,$$

$$p_i(n) = p_i \text{ pour } n > kN,$$

$$\text{où } p_2 < p_1 \text{ et } k < p_2 / (p_1 + p_2).$$

Dans ces conditions, il est possible de réaliser les promesses intrinsèques des deux méthodes, et n'importe quelle répartition $\langle n, N - n \rangle$ avec $kN < n < N - kN$ est une répartition OC. Il existe une répartition bilatéralement stable $\langle n^*, N - n^* \rangle$, avec $n^* = p_1 N / (p_1 + p_2)$. Dès lors que $p_2 / (p_1 + p_2) > k$, cette répartition bilatéralement stable

10 Il est également nécessaire pour la suite du propos que p_2 ne soit pas trop petit. Ces conditions vagues peuvent être formulées plus précisément en exigeant que $\exp\{-kp_2N / (p_1 + p_2)\} \ll 1$.

sera une répartition OC. Jusqu'ici, tout va bien. Tant que les promesses intrinsèques de la méthode inférieure ne sont pas trop basses, et que les méthodes répondent rapidement à l'affectation des travailleurs, il existera une division optimale du travail cognitif que la communauté peut maintenir – si elle peut l'atteindre.

Mais il y a un hic. La zone d'attraction de la répartition stable inclut toutes les répartitions OC, mais il est tout à fait possible que la communauté se retrouve bloquée dans une répartition sous-optimale, et notamment dans la répartition extrême $\langle N, 0 \rangle$. Intuitivement, si p_2 est trop petit ou N trop grand, un non-conformiste pourrait n'avoir aucun bénéfice à abandonner la méthode I. La bonne nouvelle est qu'il existe des instances de ce type de cas dans lesquelles la communauté hobbesienne fait non-seulement mieux que ses cousins aux nobles principes, mais parvient effectivement à une division du travail cognitif optimale et stable. La mauvaise nouvelle est que, lorsque la communauté est trop grande, les intérêts égoïstes conduisent la communauté au même état sous-optimal que la rationalité individuelle.

Il y a cependant un remède. Le problème des grandes communautés (plus précisément des communautés pour lesquelles kN est trop grand) est qu'un unique déserteur de la méthode I ne parviendra pas à fournir une contribution suffisante avec la méthode II pour que cette méthode devienne rentable. Ce qu'il faut, c'est que plusieurs personnes changent de bord ensemble. Imaginons par conséquent que la communauté soit divisée en fiefs (les laboratoires) et que, lorsque le chef local (le directeur de laboratoire) décide de changer de bord, la paysannerie locale (les doctorants) le suive. Supposons que chaque laboratoire contienne q membres et que le directeur puisse ainsi faire changer avec lui x membres de la communauté, avec x inférieur ou égal à q . Évidemment, si $q > kN$, un seul laboratoire peut réaliser les promesses intrinsèques de la méthode II, et il est facile de voir qu'il existe des conditions dans lesquelles une répartition OC stable est atteignable. Morale : un certain degré d'autocratie locale – des directeurs de laboratoires qui peuvent contrôler l'allégeance d'un certain nombre de travailleurs – peut permettre à la communauté d'être davantage flexible.

J'ai exploré jusque là certaines conséquences d'une description très schématique de la façon dont les motifs personnels (et les incitations sociales) pourraient opérer dans une communauté scientifique. Évidemment, mon analyse pourrait être améliorée en adoptant des hypothèses plus réalistes et en introduisant des facteurs qui ont jusque-là été omis (par exemple, la possibilité que les deux méthodes puissent réussir, des différences individuelles de talent et d'intérêt, etc.). Je veux cependant terminer cette section en répondant à une objection évidente. Mes conclusions concernant les répartitions attendues pour la communauté ne dépendent-elles pas non seulement de l'hypothèse selon laquelle chaque membre est mû par le désir de remporter le prix, mais aussi de l'hypothèse selon laquelle chacun a suffisamment d'information concernant les probabilités pour atteindre au mieux ce but ? Et cette supposition n'est-elle pas hautement irréaliste ?

Ma réponse est que l'hypothèse selon laquelle les scientifiques identifient les

probabilités de succès des méthodes, étant donné une affectation d'un certain nombre de travailleurs, est une idéalisation du même type que celle qui est traditionnellement faite dans la théorie de la confirmation. De même que nous idéalisons les jugements de tous les jours avec lesquels les scientifiques évaluent les possibilités de succès de théories rivales, de même nous pouvons partir de l'idée ordinaire qu'une méthode est trop représentée (ou pas assez représentée) ou de la conscience que les chances qu'une méthode réussisse sont très basses à moins qu'elle ne soit adoptée par une masse critique de personnes, et supposer que ces jugements rudimentaires sont remplacés par une attribution de probabilités numériques.

Faire appel à l'ambition humaine n'est bien sûr qu'un début. D'autres mécanismes psychologiques pourraient amener les scientifiques à se trouver encore plus proches de la répartition OC qu'ils ne l'auraient été autrement. Non seulement les vices, qu'il s'agisse de l'avidité ou de la fraude, peuvent jouer un rôle constructif, mais des traits de caractère plus sains peuvent aussi contribuer aux buts de la communauté. La persévérance, l'investissement personnel, les loyautés personnelles ou nationales, et la dévotion envers des causes politiques peuvent, à l'occasion, aider à combler un écart OC-IR.

V

J'ai étudié en détail le problème tel qu'il se pose pour les méthodes. Peut-on obtenir des résultats similaires pour notre problème initial, le problème du choix entre des théories ?

Oui. Mais la description précise de cas de choix entre des théories s'avère être délicate. Les problèmes viennent, pour le dire sommairement, de l'existence de deux sources d'incertitude : nous devons prendre en compte la probabilité qu'une théorie puisse améliorer son statut épistémique apparent et aussi la probabilité que, si elle le fait, elle soit plus proche du but épistémique (par exemple la vérité). Néanmoins, si nous sommes résolus à faire des idéalisations importantes, mon histoire initiale peut être exposée plus en détails. Le texte de cette section va présenter une version très simplifiée de l'histoire ; on trouvera dans les notes des indications pointant vers un plus grand réalisme.

Imaginons que, à un certain moment de l'histoire d'une certaine science, nous ayons deux théories rivales et incompatibles, T_1 et T_2 . Étant donné les preuves disponibles, la probabilité que T_i soit vraie est q_i , et cela est reconnu par tous les membres de la communauté scientifique qui s'intéressent à ces théories. Supposons également que $q_1 + q_2 = 1$, que $q_1 > q_2$, mais que q_1 soit approximativement égal à q_2 .¹¹

11 Il y a ici plusieurs idéalisations différentes : je suppose qu'il soit possible d'associer aux théories des probabilités définies sur la base des preuves disponibles, qu'il y ait une reconnaissance universelle des bonnes probabilités, et que l'une des théories soit correcte. Un premier pas vers davantage de réalisme serait d'abandonner cette dernière supposition, en permettant que $q_1 + q_2 = r < 1$.

Le but de la communauté est d'arriver à l'acceptation universelle de la théorie vraie, d'éliminer les problèmes auxquels cette théorie est actuellement confrontée, et de développer la théorie en l'appliquant à la fois à des problèmes théoriques et à des questions pratiques. En poursuivant ce but, la communauté peut suivre l'une des deux stratégies génériques suivantes : (A) affecter tous les scientifiques à T_1 ; (B) affecter n scientifiques à T_1 , $N-n$ à T_2 (avec $0 < n < N$). Je considérerai les mérites de ces stratégies de la perspective d'une étape ultérieure de l'histoire de la communauté – « le jour du Jugement » – au moment auquel nous attribuons des utilités épistémiques aux différentes conséquences¹².

Les résultats possibles le jour du Jugement sont les suivants. Si tous les scientifiques ont réalisé des recherches avec la théorie vraie, alors je supposerai que c'est la meilleure de toutes les possibilités, et que la communauté a accumulé, à travers la résolution de problèmes et autres tâches, l'utilité épistémique u_1 . D'un autre côté, si tout le monde a travaillé avec la théorie fautive, alors je supposerai que les efforts ont été complètement vains, avec l'utilité épistémique $-u_1$. Pour comprendre les conséquences de la division du travail, nous avons besoin d'introduire le concept d'un *état concluant*, une situation dans laquelle les théories T_1 et T_2 ont pu être départagées. Je ferai l'hypothèse que les deux théories disponibles sont actuellement rongées par des anomalies, et qu'il est nécessaire que ces problèmes soient résolus si elles veulent remporter une adhésion sans réserve (honorant ainsi l'idée traditionnelle que les théories naissent réfutées). Nous dirons qu'un état concluant en faveur de l'une des deux théories, disons T_1 , est atteint dans le cas où, le jour du Jugement, T_1 a réussi à surmonter ses problèmes et T_2 n'a pas réussi à le faire, malgré l'opportunité qui lui a été donnée. La controverse entre T_1 et T_2 est résolue si des recherches ont été menées avec les deux théories et si l'une résout ses anomalies actuelles, et pas l'autre.

Faisons maintenant l'hypothèse très optimiste selon laquelle la nature, bien que n'étant pas un livre ouvert, n'est pas non plus hostile : même si des théories correctes peuvent rencontrer des anomalies, des théories qui résolvent toutes leurs anomalies sont correctes. Ainsi, si nous atteignons un état concluant, nous n'avons pas à nous soucier de faux positifs – dans un état concluant, nous résolvons le problème et nous le résolvons correctement. Grâce à cette hypothèse, j'attribuerai aux différents résultats les utilités épistémiques suivantes : si la division du travail résultant de l'une des stratégies (B) conduit à un état concluant, alors nous obtenons l'utilité épistémique u_2 ($0 < u_2 < u_1$) ; sinon, l'utilité épistémique est 0 (nous sommes toujours dans la même situation problématique, même si notre travail peut nous avoir donné une vision plus claire des problèmes auxquels les deux théories concurrentes font face)¹³.

12 Il serait ici plus réaliste d'autoriser la communauté à utiliser les théories alors disponibles afin d'obtenir une théorie-fille qui serait plus proche de la vérité qu'aucune des deux autres.

13 Mes hypothèses sur les u_i peuvent être facilement modifiées pour refléter des conceptions différentes concernant les valeurs de tel ou tel résultat, ou même des situations particulières différentes qui requièrent l'attribution de valeurs différentes. Remplacer des hypothèses concernant les états concluants par des hypothèses plus réalistes est une tâche plus délicate.

L'utilité attendue de la stratégie plus prometteuse (A) (affecter tous les scientifiques à T_1) se calcule facilement. Elle vaut

$$q_1u_1 - q_2u_2.$$

Pour calculer l'utilité attendue de la stratégie (B_n) (choisir la répartition $\langle n, N-n \rangle$), nous avons besoin de reconnaître qu'un état concluant en faveur de T_1 sera atteint dans le cas où (a) T_1 résout ses anomalies actuelles, (b) suffisamment de travailleurs sont affectés à T_2 pour donner une chance à T_2 . Je ferai l'hypothèse que la probabilité que (b) soit satisfait est 1 si $N-n$ est supérieur à une certaine valeur m , et 0 sinon. La probabilité que (a) soit satisfait est la probabilité que T_1 soit vraie multipliée par la probabilité que T_1 réponde aux efforts de n scientifiques. En appelant $p_i^*(n)$ la probabilité qu'une théorie vraie T_1 réponde à l'affectation de n travailleurs en résolvant ses problèmes, nous pouvons écrire l'utilité attendue de (B_n), où $m < n < N-m$ (de sorte qu'une chance soit donnée aux deux théories), comme

$$q_1p_1^*(n)u_2 + q_2p_2^*(N-n)u_2.$$

La division du travail est ainsi préférable s'il existe un n qui satisfasse la contrainte « donner une chance aux deux », $m < n < N-m$, tel que

$$q_1p_1^*(n)u_2 + q_2p_2^*(N-n)u_2 > q_1u_1 - q_2u_2.$$

Je considérerai seulement l'une des fonctions de retour introduites dans la section III (des résultats similaires peuvent être facilement obtenus dans l'autre cas). Supposons que $p_i^*(n) = p_i \cdot (3n^2 - 2n^3/kN)/k^2N^2$ ($n < kN$), $p_i^*(n) = p_i$, sinon. Nous pouvons simplifier la discussion sans l'appauvrir en supposant que $p_1 = p_2 = p$. Supposons, comme dans nos discussions précédentes, que $k < 1/2$, de sorte qu'il soit possible d'affecter les scientifiques à chaque théorie en lui permettant de réaliser ses promesses intrinsèques. Si n appartient à l'intervalle $[kN, (1-k)N]$, alors il semble raisonnable de conclure que la condition selon laquelle les deux théories doivent avoir eu une chance est satisfaite. Aussi, l'inégalité cruciale pour préférer l'une des stratégies (B_n) est

$$pu_2 > (q_1 - q_2)u_1.$$

Ainsi, à moins que les chances maximales qu'une théorie vraie surmonte ses problèmes soient faibles (p est petit) ou que l'utilité d'une action immédiate soit grande (u_1 est grand par rapport à u_2), la répartition OC implique à nouveau une authentique division.

Les motivations non-épistémiques pourraient-elles rapprocher les scientifiques

En principe, on devrait prendre en compte l'astucieuse possibilité qu'une fausse théorie continue d'apparaître plausible, et l'idée d'une simple opposition entre d'un côté la victoire d'une théorie et de l'autre un conflit non résolu devrait être remplacée par l'étude de l'évolution des probabilités attribuées aux théories concurrentes. Par conséquent, nous désirons avoir une estimation des probabilités qui peuvent être attribuées à chaque théorie lorsqu'il existe une division particulière du travail cognitif, et utiliser cela pour exprimer la probabilité de prendre une décision correcte le jour du Jugement. Mes explorations préliminaires de l'algèbre plus complexe qui en résulte suggèrent que, dans un nombre significatif de cas, l'approche plus simple du texte ne nous égarera pas beaucoup.

individuels de la répartition OC ? Supposons que la motivation principale de chaque scientifique soit le désir d'être reconnu par la postérité comme ayant été l'un des premiers défenseurs de la théorie acceptée. Si la communauté est initialement divisée, avec la répartition $\langle n, N - n \rangle$, et si $kN < n < (1 - k)N$, alors il existe une répartition stable atteignable, $\langle n^*, N - n^* \rangle$, où

$$n^* = q_1 N / (q_1 + q_2).$$

Dès lors que $k < q_2$ (une hypothèse très faible, puisque $k < 1/2$ et que q_2 est proche de q_1 [$= 1 - q_2$]), ce sera une répartition OC. Morale : comme pour les modèles précédents, il existe des circonstances particulières, bien que hautement idéalisées, dans lesquelles la répartition IR diverge de la répartition OC et dans lesquelles des motivations extra-épistémiques amènent la communauté vers la répartition OC. Les structures sociales au sein de la communauté scientifique peuvent tirer parti des motivations personnelles des individus au bénéfice des projets épistémiques de la communauté.

VI

Pour finir, je souhaite donner brièvement quelques indications sur la façon dont l'analyse que j'ai commencée ici pourrait être approfondie et étendue. Premièrement, il serait relativement aisé de considérer des structures sociales et des motivations personnelles (des loyautés nationales ou personnelles, par exemple) que j'ai écartées de l'analyse. Deuxièmement, dans mon traitement des cas particuliers, j'ai implicitement supposé que la taille de la main d'œuvre disponible était fixée. Mais cela est certainement déraisonnable. De même que nous pouvons considérer qu'il existe en science des problèmes d'optimisation au sein d'un domaine de recherche, nous pouvons aussi considérer des problèmes d'optimisation entre différents domaines. Même si les intérêts personnels pouvaient aussi aider la communauté des scientifiques à obtenir une répartition raisonnable des scientifiques entre les domaines de recherche, des difficultés de reconversion pourraient interférer dans le processus.

L'optimisation entre domaines n'est pas le seul problème. Il existe un problème de division du travail de plus grande envergure encore. Les buts épistémiques de la communauté n'épuisent pas l'ensemble des buts de la communauté, et nous pouvons demander comment, étant donné tous les buts que nous avons pour nous-mêmes et pour nos semblables, nous devrions allouer les ressources pour la poursuite des buts épistémiques de notre communauté. Étant donné la solution de ce problème d'optimisation, nous connaissons la taille de la main d'œuvre disponible en sciences. Nous pouvons alors demander quelle est la division optimale du travail parmi les domaines scientifiques et en venir enfin à la question qui a été posée de façon préliminaire dans cet essai : quelle est la division optimale du travail au sein d'un domaine scientifique, et de quelles façons les intérêts épistémiques et non-épistémiques nous en rapprochent-ils ou nous en éloignent-ils ? Cette question prend place, au bout du compte, au sein d'un ensemble de problèmes d'optimisation

imbriqués.

Analyser ce qui est optimal ne conduit pas nécessairement à être optimiste. L'une de mes rengaines a été la possibilité que des facteurs psychologiques (et les institutions scientifiques qui exploitent ces facteurs) qu'on a souvent considérés comme étant préjudiciables au progrès cognitif puissent en fait jouer un rôle constructif. Mais il serait extrêmement surprenant que les structures sociales de la science actuelle, qui ont évolué à partir des propositions de personnes qui avaient des buts tout à fait différents pour cette entreprise et qui la pratiquaient dans un milieu social très différent, se voient donner raison par une analyse d'optimalité. Comment devons-nous concevoir au mieux les institutions sociales pour le progrès de la connaissance ? Les philosophes ont ignoré la structure sociale de la science, mais ce qui importe, c'est de la transformer.

PHILIP KITCHER

Université de Californie/San Diego

Traduction : THOMAS BOYER-KASSEM