

## ENTRETIEN AVEC JON ELSTER



Jon Elster  
Professeur au Collège de France  
titulaire de la chaire de  
Rationalité et sciences sociales  
depuis 2006.

On évoque souvent le caractère pluridisciplinaire de vos travaux et la diversité de vos centres d'intérêt pour vous définir comme un penseur inclassable. Comment vous présenter ?

Lorsqu'on me demande quelle est ma spécialité, je réponds que je fais « ceci-et-cela ». Ce qui m'intéresse, ce sont des problèmes. Pour y faire face, nous devons utiliser les outils appropriés, indépendamment des disciplines, dont la pertinence est plus bureaucratique que scientifique. Je m'intéresse surtout aux décisions individuelles et collectives, et à tout ce qui précède la décision : la formation des préférences, la formation des croyances, les émotions, et aussi les mécanismes d'interaction dans les décisions collectives. C'est sans doute le fil conducteur de ma réflexion.

Il y a aussi dans beaucoup de mes travaux une sorte de composante normative. Non pas que je propose moi-même une théorie de la justice ou du bien, mais la manière dont j'envisage les choix individuels est telle que les conceptions normatives que se font les acteurs eux-mêmes interviennent comme des déterminants

causaux de leurs choix et de leurs décisions. Cette idée d'étude empirique de la justice, ou des motivations de justice ou d'équité m'a beaucoup intéressé. Il y a des problèmes d'équité pour l'allocation des ressources rares par les institutions, lorsqu'il s'agit par exemple de décider qui doit recevoir un rein pour la transplantation ou d'autres questions de ce genre<sup>(1)</sup>. Un autre exemple concerne ce qu'on appelle la justice de transition<sup>(2)</sup>. Si, en 1815, il avait fallu allouer des réparations aux victimes des spoliations de la Révolution, qui aurait dû être prioritaire ? Les gens qui étaient restés en France à lutter pour le roi, par exemple en Vendée, et dont les propriétés avaient été détruites, ou ceux qui avaient émigré et dont les biens avaient été confisqués ? Ou bien ceux dont les besoins étaient les plus importants ? Ce ne sont là que quelques exemples des questions d'équité qui se posent dans les situations de justice de transition. Dans ces décisions collectives – allocation d'organes pour la transplantation ou réparation aux victimes de conflits – les conceptions normatives des agents politiques ou des fonctionnaires sont déterminantes. C'est par ce biais que mes recherches ont une composante normative.

Vous écrivez, à propos de la justice de transition, qu'il s'agit d'abord de régler les comptes du passé avant de préparer l'avenir.

Dans les phases de transition, les deux démarches sont souvent simultanées – l'une tournée vers le passé, l'autre vers l'avenir. Par exemple, on commence le travail d'écriture d'une nouvelle Constitution, et on commence en même temps le travail de représailles et de réparation. Certains pensent que la tâche rétrospective est la plus importante et qu'il faut faire table rase du passé, éliminer tout ce qui est vieux et corrompu afin de pouvoir commencer à construire du neuf. D'autres – dans les pays de l'Est par exemple, mais c'est aussi ma position – jugent que dans le cas de régimes qui sont restés en place pendant des générations et qui ont fini par rendre tout le monde plus ou moins coupable, il devient impossible d'épurer la société. Dans la France de 1945, c'était encore possible dans la mesure où le nombre de collaborateurs était relativement restreint. Dans les pays de l'Est, on peut dire qu'il y a en quelque sorte une complicité de la quasi-totalité de la population. Dans

1. Jon Elster, *Local Justice: How Institutions Allocate Scarce Goods and Necessary Burdens*, Russell Sage Foundation Publications, 1993.

Voir aussi J. Elster et N. Herpin (eds.), *L'éthique des choix médicaux*, Poitiers, Actes Sud 1992

2. Jon Elster, *Closing the Books: Transitional Justice in Historical Perspective*, Cambridge University Press, 2004.

un petit essai sur *Le pouvoir des impuissants*, Vaclav Havel décrit très bien ces petits actes de complicité de tous les jours, qui ont l'air insignifiants, mais contribuent à brouiller les choses. Il dépeint le petit marchand de légumes qui affiche dans sa vitrine un panneau indiquant « prolétaires de tous les pays, unissez-vous ». Par cet acte imposé par le régime, qui ne correspond sans doute pas du tout à une position personnelle, il se rend néanmoins complice du régime. En fin de compte, on obtient un tableau en nuances de gris, et non en noir et blanc. Il n'y a pas de distinction tranchée entre les bons et les mauvais, les résistants et les collaborateurs. Bien sûr, les deux existent, mais il y a surtout une masse énorme de collaborateurs du quotidien. Dans un article de 1992, j'ai écrit que dans de telles situations, rechercher tous les coupables est impossible, en rechercher quelques-uns est arbitraire : il faut donc y renoncer entièrement et se tourner vers le futur. C'est la solution adoptée par l'Espagne en 1978. On a pris la décision consensuelle d'ignorer le régime de Franco et les atrocités commises par les deux parties au cours de la guerre civile, et de mettre également de côté la question des réparations, qui aurait rendu impossible la marche en avant.

Mais en écrivant cela, j'avais peut-être sous-estimé le rôle qu'allait jouer l'ancienne Nomenklatura dans les sociétés des pays de l'Est, par exemple. J'observe également qu'en Espagne, un début de revendication semble s'esquisser aujourd'hui. En Argentine et au Chili, ainsi qu'en Pologne, vingt ans après la chute des dictateurs, la justice de transition est soudainement réapparue. Lorsque justice n'a pas été faite, la demande peut ressurgir longtemps après. On l'a constaté aussi en France avec l'affaire Papon.

De même, c'est un miracle qu'il n'y ait pas eu de violence collective en

Afrique du Sud, alors que la grande masse de la population noire n'a obtenu ni terre ni justice. On peut imaginer qu'il y a là une sorte de bombe à retardement qui pourrait exploser d'un jour à l'autre. Ces problèmes ont une durée de vie très longue. J'ai lu, à propos des guerres de Vendée, que même après la Première Guerre mondiale, il y avait en France des terres « maudites », que personne ne voulait acheter à cause des contestations qui ont suivi la Révolution. On s'en souvient pendant des siècles.

Des problèmes de ce genre se posent dans beaucoup de pays, de l'Irak à l'Argentine. Au début des années 90, j'ai beaucoup voyagé dans les pays de l'Est où les juristes travaillaient à la rédaction de nouvelles Constitutions. Avec quelques collègues de l'université de Chicago, nous avons animé des séminaires de constitutionnalisme comparé pour discuter des grands modèles, de leurs succès et de leurs échecs. Plutôt que de donner des conseils, nous cherchions à établir un dialogue – mais il est difficile d'en mesurer les effets.

[Dans vos travaux, vous mettez en relation des éléments venus de la littérature ou des humanités et des théories scientifiques très actuelles. Peut-on décrire votre activité comme une sorte d'anthropologie philosophique ?](#)

Dans ma leçon inaugurale, j'ai dit que la rationalité, comme valeur transhistorique, transculturelle, etc., relevait d'une anthropologie philosophique et non d'une anthropologie empirique. C'est un grand mot pour dire simplement « analyse conceptuelle ». En réalité, j'ai trouvé mes hypothèses chez les moralistes français du XVII<sup>e</sup> siècle et j'ai cherché à les vérifier chez les psychologues et les économistes du XX<sup>e</sup> siècle. C'est un va-et-vient un peu étrange, déconcertant peut-être. J'ai sur ce point une perspective

d'une grande simplicité. Cela fait 25 siècles que les gens essaient de comprendre le comportement humain ou la nature humaine – disons depuis le temps d'Aristote ou de Platon. Pourquoi le dernier siècle ou la dernière décennie seraient-ils privilégiés ou plus intéressants ? Y aurait-il plus de génies ou de grands penseurs ? Il n'y a aucune raison de le penser, et de fait c'est faux. Il suffit de lire Montaigne, Aristote, La Rochefoucauld, Tocqueville, Proust, pour ne citer qu'eux : ils débordent d'hypothèses.

[Pourriez-vous présenter la théorie du choix rationnel, qui est au cœur de vos travaux ?](#)

La théorie du choix rationnel est quelque chose de très simple. On pourrait l'illustrer par un proverbe norvégien qui dit « ne traverse pas la rivière pour chercher de l'eau ». Il traduit une sorte d'efficacité instrumentale qui est de tous les pays et de tous les temps. On entend dire parfois que la rationalité serait une valeur occidentale, ou moderne, ou masculine, ou je ne sais quoi d'autre. C'est absurde. On ne peut pas ne pas vouloir être rationnel. Si l'on veut atteindre une fin, on veut forcément l'atteindre de la manière la plus efficace. Cela fait partie de l'idée même de « vouloir atteindre une fin ». En ce sens, c'est tout à fait banal. En fait, le point difficile, dans tout développement de la théorie du choix rationnel, concerne les croyances. Dans beaucoup de cas, avant d'agir, rationnellement ou non, il faut former une croyance. Comment peut-on former une croyance rationnelle ? C'est parfois très difficile : dans le cas du réchauffement climatique, il y a sans doute des centaines de modèles. Comment les parcourir tous et se former une croyance permettant de prendre des décisions ? Faut-il prendre la moyenne ? Le scénario du pire ? Peut-on assigner des probabilités à chaque modèle et prendre la

moyenne pondérée ? Personne n'a de réponse. C'est un problème épistémologique très profond, et c'est souvent un obstacle au choix rationnel : comment agir rationnellement en l'absence de croyances bien fondées ? Il y a là une source d'indétermination dans toutes les situations de choix complexes, comme les choix politiques.

Une autre source d'indétermination provient de l'interaction stratégique entre les agents. Lorsqu'avant d'agir, chaque agent doit se faire une opinion sur ce que feront les autres, tout en sachant que les autres sont en train de former leur opinion en se posant la même question à propos de lui, il peut y avoir dans certains cas des situations inextricables. Prenons un exemple courant, le jeu du peureux ou de la poule mouillée (*game of chicken*). Deux voitures foncent l'une vers l'autre, le peureux est le premier qui s'écarte. Dans ce cas, il n'y a pas de choix rationnel : pour chacun des deux, il est optimal de dévier si et seulement si l'autre ne le fait pas. C'est une situation d'indétermination. Il y a souvent des situations de ce genre dans la réalité, par exemple pour la construction de biens à usage public. Si deux armateurs ont intérêt à construire un phare, qui bénéficie à tout le monde et dont la construction unilatérale serait profitable à chacun des deux, même si le deuxième pouvait s'en servir gratuitement, la difficulté est de savoir qui va le construire. En effet, il est dans l'intérêt de l'un de le construire si et seulement si l'autre ne le fait pas. Dans ces situations d'interactions stratégiques présentant des équilibres multiples, quand il n'y a pas de critère permettant la convergence tacite des agents vers l'un de ces équilibres, on ne peut pas former de croyance rationnelle et il n'y a pas d'action rationnelle. Le concept n'est pas défini. Non que les gens ne soient pas rationnels : simplement, la rationa-

lité ne suffit pas pour leur dicter ce qu'il faut faire.

Outre ces problèmes conceptuels, il y a aussi les problèmes concrets : souvent, même lorsqu'ils sont en mesure de former des croyances rationnelles, les gens ne le font pas. Dans ce cas, l'irrationnel ne vient pas de l'indétermination du concept, mais de l'irrationalité du comportement. C'est un phénomène massif. C'est le sujet de mon cours de l'année 2007-2008, après le cours précédent sur le désintéressement. Ce sont deux volets d'une critique de la pensée économique, qui elle, est fondée sur la rationalité et l'intérêt.

**Dans votre leçon inaugurale, vous vous interrogez sur la rationalité de l'acte de voter.**

Ma chaire est intitulée « rationalité et sciences sociales ». Les sciences sociales étudient le vote, et il est légitime de se demander si l'acte de voter peut être considéré comme rationnel. On répond souvent par la négative avec l'argument que la probabilité qu'un vote soit décisif est moindre que le risque de mourir dans un accident de la circulation en se rendant aux urnes. Les gens se comportent comme s'ils ne se souciaient pas des conséquences de leurs actions. On observe pourtant des taux de participation élevés lorsqu'il n'y a pas de candidat largement favori et que le scrutin s'annonce serré, ou lorsque l'enjeu des élections est important, ce qui suggère qu'ils s'en soucient quand même.

**Mais si l'on adopte un point de vue plus global, par exemple une perspective évolutionniste, on dira que c'est un acte inutile et coûteux pour l'individu, mais néanmoins rationnel parce qu'il a une utilité pour le groupe social. La théorie du choix rationnel peut-elle prendre en compte ce type d'analyse ?**

À vrai dire, la théorie du choix rationnel et la théorie de l'évolution n'ont pas beaucoup de rapports. Et cependant, il faut croire que l'évolution, de manière générale, a produit une tendance vers le comportement rationnel, puisque dans une situation donnée, un animal rationnel, qui utiliserait des ressources rares de façon plus efficace, serait évidemment favorisé par la sélection naturelle : c'est en quelque sorte évident. Comment l'évolution n'aurait-elle pas produit la rationalité ?

D'autre part, il semble aussi que par 3 ou 4 mécanismes distincts, l'évolution ou la sélection naturelle a aussi produit la disposition au comportement altruiste, que ce soit l'altruisme réciproque (*tit for tat* ou gagnant-gagnant), la sélection de parentèle (*kin selection*) ou même la sélection de groupe, qui connaît actuellement une sorte de regain de faveur. Mais il est toujours délicat de faire appel à ces mécanismes et les explications évolutionnistes ont parfois tendance à brûler les étapes, ce qui les expose à des spéculations hasardeuses. Bien sûr, en principe, tout cela est un résultat de l'évolution. Pascal disait, à propos de Descartes : « il faut dire en gros 'cela se fait par figure et mouvement', car cela est vrai, mais de dire quelles et composer la machine, cela est ridicule, car cela est inutile et incertain et pénible<sup>(3)</sup> ». En fin de compte, c'est peut-être Descartes qui a eu raison, mais quatre siècles plus tard.

Je pense en tout cas qu'aucun des mécanismes que j'ai mentionnés n'est susceptible d'expliquer la tendance à voter pour la simple raison que généralement l'acte de voter n'est pas observé par les autres. C'est un acte discret et anonyme. La plupart du temps, en ville, personne ne sait si vous êtes allés voter. Or les mécanismes que j'ai évoqués supposent que les actes des uns soient connus

3. Pascal, *Pensées*, 118.

par les autres de sorte qu'ils puissent être encouragés ou sanctionnés le cas échéant. Ce sont des questions complexes qu'il serait trop long de traiter ici. J'ajouterai que même si l'on ne peut pas dire que le vote soit un acte rationnel, le votant pourra en retirer un plaisir psychique. C'est ce qu'on peut appeler l'effet Valmont<sup>(4)</sup> : donner pour le plaisir intime et délicieux que l'on retire de l'acte de donner.

Ou alors, pour expliquer le vote doublement anonyme – personne ne sait si vous votez ni comment vous votez – on peut aussi faire appel à une sorte de pensée magique. L'agent se dit : je suis un membre typique de mon groupe de référence – les professeurs d'université, par exemple – donc si je vote, il est probable que mes semblables voteront également. Par conséquent je vais voter. C'est une confusion entre la valeur diagnostique et l'efficacité causale. Si je vote, cela révèle que je suis dans une certaine disposition, mais en même temps, je pense que les autres, ayant la même disposition, en feront autant. Dans mon dernier livre<sup>(5)</sup>, je cite des textes de Proust qui font appel à cette pensée magique, notamment un passage où le narrateur dit à propos de Saint Loup qu'il est fidèle à sa maîtresse avec l'idée superstitieuse que s'il lui est fidèle, de son côté elle le sera également. C'est peut-être en suivant le même genre de raisonnement superstitieux que l'on vote.

**Il est clair que la rationalité n'est pas le seul moteur du comportement. On peut noter à ce propos qu'il y a de plus en plus de travaux sur les émotions, notamment du côté des sciences cognitives.**

Dans la tradition où je me situe, l'économie du comportement (*beha-*

*vioural economics*), on fait appel aux expériences psychologiques pour montrer comment se font vraiment les choix – rarement rationnels –, et l'on parle effectivement beaucoup des émotions actuellement. Il y a des expériences très impressionnantes, notamment dans l'équipe de Ernst Fehr, à Zurich, qui montrent que sous le coup de l'émotion, surtout de l'indignation ou du ressentiment, les gens sont prêts à faire des sacrifices matériels considérables pour punir quelqu'un qui leur a joué un mauvais tour. Il est intéressant de remarquer qu'il s'agit essentiellement d'émotions négatives – la colère, l'indignation, l'envie, le ressentiment –, tandis que l'amour, la joie, la sympathie, la pitié, ne sont presque pas étudiées à ma connaissance. Cela correspond aussi au fait que, parmi les moteurs de l'action humaine, le sentiment d'injustice est beaucoup plus puissant que le désir de réaliser la justice.

**Cela nous ramène à la question des limites de la rationalité. Peut-on parler de degrés de rationalité, selon les points de vue et les critères considérés ?**

Du point de vue de l'agent, la rationalité est une notion entièrement subjective. Dire que l'action est rationnelle, c'est dire que du point de vue de l'observateur, c'était l'action qui s'imposait comme le meilleur moyen de réaliser ses désirs étant donné ses croyances et étant donné que ses croyances étaient rationnelles. De ce point de vue, il n'y a pas vraiment de degré dans la rationalité. En revanche, l'agent peut tenir compte de ses faiblesses et les anticiper. Je parlerais dans ce cas de *second best rationality*. Quand on se sait incapable au moment de l'action de se comporter de manière rationnelle, alors on peut au préalable

prendre un engagement qui élimine les sources de l'irrationalité ou qui impose une contrainte, comme Ulysse<sup>(6)</sup> demandant à ses compagnons de l'attacher au mât pour ne pas céder au chant des sirènes. En ce sens on peut parler de degrés de rationalité. On peut distinguer d'une part la rationalité parfaite d'un agent qui serait toujours parfaitement maître de lui-même, transparent à lui-même à chaque instant, et d'autre part des agents qui n'ont pas toujours le contrôle d'eux-mêmes, mais qui en sont conscients, et qui sont même capables d'anticiper leur comportement et de se corriger eux-mêmes. J'ai proposé toute une liste de mécanismes permettant d'expliquer comment l'agent peut prendre en compte ses propres tendances à se comporter de manière irrationnelle<sup>(7)</sup>.

**Diriez-vous que l'émotion est irrationnelle ?**

Il est clair que l'émotion peut avoir des effets dont certains subvertissent la rationalité de l'agent, notamment en rendant les croyances moins rationnelles. Prendre ses désirs pour des réalités n'est pas très rationnel, mais sous le coup de l'émotion, nous le faisons souvent. Pour autant, les désirs qui surgissent des émotions ne sont ni rationnels ni irrationnels – je rejoins sur ce point l'analyse de Hume. Bien sûr, on peut avoir des désirs stupides qui rendent la vie misérable, mais c'est autre chose. Paradoxalement, le toxicomane peut être rationnel. De son point de vue, dans une perspective temporelle assez courte, liée à la nécessité immédiate non pas tant de jouir de la drogue que d'éviter le manque, son comportement peut être très « rationnel ». C'est une décision terminologique : on peut utiliser le mot rationnel comme on veut. Pour

4. Choderlos de Laclos, *Les liaisons dangereuses*, lettre XXI.

5. Jon Elster, *Agir contre soi*, O. Jacob, 2007.

6. Jon Elster, *Le laboureur et ses enfants. Deux essais sur les limites de la rationalité*. Paris, Ed. de Minuit, 1987.

7. Jon Elster, *Explaining Social Behavior*, Cambridge University Press 2007, Ch. 13.

ma part, je l'utilise à des fins explicatives, pour comprendre une action à partir des données propres à l'acteur lui-même : ce qui chez l'acteur – les désirs, les croyances, les émotions – précède et cause l'action. Du point de vue explicatif, il y a une structure causale qui permet de dire que l'action est rationnelle – cela n'a rien à voir avec le jugement d'un observateur qui pourrait dire que l'agent s'est comporté de manière stupide, autodestructrice, etc. C'est lié au fait qu'il se soucie surtout de son avenir immédiat.

Cette question de l'horizon temporel de l'agent est importante. Il est difficile de persuader les gens de tenir compte davantage des conséquences lointaines de leurs actes, par exemple quand on cherche à modifier des comportements, comme dans le cas des campagnes de prévention contre le tabac ou l'alcool.

**Comment faire pour être rationnel dans ces situations, aujourd'hui fréquentes, où l'on trouve la conjonction d'une grande incertitude, de l'urgence de la décision et de l'importance des enjeux ? Dans le cas du réchauffement climatique, par exemple, ou encore de l'ESB ou du sang contaminé.**

Incertain, enjeux et urgence : on les rencontre aussi dans la question du terrorisme. Je pense qu'il faut résister à l'urgence. Aux États-Unis, après le 11 septembre, il fallait décider de la conduite à tenir. Il y avait bien sûr beaucoup de colère et de peur, et ces émotions tendent à induire l'urgence, que je définis comme le désir d'agir immédiatement plutôt que de différer l'action. Quand on agit de manière immédiate, on ne prend pas le temps de recueillir toutes les informations nécessaires. Les informations qu'on néglige sont surtout celles qui portent sur les effets indirects lointains et potentiellement pervers des actions qu'on entreprend dans le présent. Sous la pression de

l'urgence, on néglige les conséquences lointaines. Par conséquent, les lois votées par le Congrès américain après le 11 septembre ont probablement créé plus de terroristes qu'elles n'ont permis d'en arrêter, justement parce qu'on n'a pas pris le temps d'étudier les conséquences indirectes, lointaines et peut-être perverses des mesures adoptées. Dans le processus parlementaire et législatif normal, on prend un certain temps – de l'ordre de deux ans – pour préparer une loi, dans des comités et des groupes de travail. Il s'agit surtout de prévenir et de contrôler d'éventuels effets indirects et pervers. Très souvent, on agit comme si les agents sociaux n'allaient pas adapter leur comportement aux lois. On a compris aujourd'hui que les lois changent les comportements des agents : il faut donc essayer de prévoir ces phénomènes. C'est difficile, mais si l'on ne s'en donne pas le temps, on risque d'adopter des lois qui auront des effets différents ou contraires de ceux que l'on recherchait.

Pour le réchauffement climatique, la situation est un peu différente. On est dans l'incertitude, les enjeux sont énormes, mais il n'y a pas la même urgence émotionnelle que dans le cas des chocs brutaux que produit le terrorisme. Il n'y a pas non plus de comportement de panique, comme lors des attentats du 11 septembre. La menace paraît plus lointaine.

Ce sont des problèmes extrêmement profonds. On invoque souvent dans ce cas le principe de précaution, mais s'il était pris à la lettre, il aurait des implications absurdes dans la vie de tous les jours. Absurde et d'ailleurs contradictoire, un peu comme le pari de Pascal, auquel, selon Diderot, un imam pourrait répondre en pariant sur un autre dieu que Pascal.

En ce qui concerne le réchauffement climatique, supposons que l'on accepte l'hypothèse selon laquelle on

court à la catastrophe si rien n'est fait. Il faut donc prendre des mesures. Mais pour être efficaces, ces mesures doivent être radicales. Elles pourraient conduire à réduire drastiquement le niveau de vie des pays en voie d'industrialisation, et engendrer des conflits violents, susceptibles de tourner à la guerre nucléaire, avec le risque d'entraîner la disparition de l'humanité par d'autres voies que la catastrophe climatique. Lorsqu'il y a un scénario catastrophe pour chaque option, on est dans l'embarras.

L'incertitude rend difficile la formulation de croyances bien fondées. Certes, il existe une théorie rationnelle du choix en situation d'incertitude mais elle ne nous apprend pas grand-chose. Elle dit que dans une situation d'incertitude définie d'une certaine manière, on peut rationnellement tenir compte soit des meilleures conséquences de chaque option, soit des pires, mais pas des conséquences intermédiaires. Cela justifie le scénario du pire, mais aussi l'hypothèse la plus favorable. Il faut choisir entre le pire et le meilleur : ce n'est pas une question de rationalité, mais de tempérament. Pessimisme et optimisme n'ont rien à voir avec la rationalité.

Face à des problèmes de ce genre, où il existe un scénario catastrophe plausible pour toutes les options, il est difficile de trouver une réponse rationnelle. Ces questions devraient être évoquées lors du colloque sur la mondialisation qui sera organisé en décembre prochain dans le cadre de l'Institut du monde contemporain. ■

*Entretien : Marc Kirsch*

#### Enseignement de Jon Elster 2007-2008 :

Cours : *L'irrationalité*  
le jeudi à 15h00 (à compter du 10 janvier 2008)  
Séminaire : *La raison en politique*  
le lundi à 17h00 (à compter du 7 janvier 2008)