

Vers la naturalisation de l'homo *œconomicus* ?

Jacky Fayolle

9 juillet 2018

«...l'homme moderne, tel que l'a façonné l'évolution, est fondamentalement un apprenti »

Daniel Andler, *La silhouette de l'humain*, 2016

Le développement foisonnant des neurosciences dans les dernières décennies fraie la voie à un naturalisme étendu, lieu de convergence possible des sciences humaines et sociales avec les sciences naturelles. Des économistes s'efforcent de fonder dans cette convergence une conception revisitée de la rationalité économique individuelle.

Apprécier la portée de cette convergence et de ce renouvellement, et relever les problèmes qu'elle soulève, suppose de revenir sur certains apports des neurosciences, d'apprécier l'impact de ces apports sur la tradition de pensée naturaliste et d'examiner la manière dont les économistes s'y prennent pour articuler leur réflexion à celle des neurosciences. Ce texte reprend successivement ces trois points et n'a que la modeste ambition d'avancer certains éléments de discussion.

1. Le cerveau, machine numérique ?

L'homme est-il naturellement un animal calculatoire ? La généalogie du raisonnement statistique peut en effet être vue sous un double angle :

- Pour des spécialistes des neurosciences, le cerveau humain est naturellement prédisposé à ce qu'on pourrait appeler une intuition numérique et statistique.

- Pour des historiens de la pensée, la statistique moderne et ses fondements probabilistes sont le produit d'une lente et progressive maturation historique qui a connu des moments-clés d'accélération et de mutation.

En 2014, la Cité des Sciences et de l'Industrie à Paris (Universcience) inaugure une exposition grand public, fort sérieuse et attractive, accompagné d'un ouvrage collectif où le maître à penser de l'exposition Stanislas Dehaene, spécialiste en neurosciences et professeur au Collège de France, signe un chapitre intitulé « Le cerveau statisticien »¹.

Les neurosciences contemporaines, fortes de nouvelles techniques d'imagerie cérébrale et d'expérimentation, explorent l'interaction entre les capacités prédéterminées du cerveau, doté d'une organisation précoce, et sa plasticité évolutive, propice à l'apprentissage. Dans la dotation initiale, la capacité à une modélisation statistique implicite semble en effet bien présente, si l'on prend au mot, et non comme une métaphore, l'expression de Stanislas Dehaene : « ... notre cerveau accumule en permanence, dans chacun de ses circuits, des statistiques des signaux qui le traversent. Ces statistiques lui servent à construire un modèle interne qui tente de décrire le monde extérieur. Apprendre, c'est choisir le modèle qui explique le mieux les données reçues par les organes des sens. En retour, ce modèle facilite l'interprétation de données sensorielles nouvelles. Lorsque des sensations ambiguës parviennent au cerveau, il en reconstruit l'interprétation la plus probable en combinant les indices sensoriels présents à un moment donné avec le modèle tiré de son expérience passée. Ainsi, sans que nous en ayons conscience, notre cerveau agit comme un statisticien qui collecte une quantité massive de données et en tire des déductions avec une logique presque parfaite » (SD, p.49). « Savoir calculer avec des probabilités est fondamental pour le cerveau, car celui-ci vit dans un monde d'incertitudes et de données partielles, dans lequel il lui faut prendre des décisions en minimisant les risques... Pour percevoir le monde extérieur, notre cerveau réaliserait

¹ C3RV34U, sous la direction de Stanislas Dehaene, Editions de La Martinière, 2014 (référence notée SD dans le texte).

donc en permanence des *inférences statistiques inconscientes* » (SD, p.50). Le cerveau mobilise les informations apportées par les sens « en tenant compte de leur niveau d'incertitude exactement comme le prédisent les équations du révérend Bayes » (SD, p.50). « Les recherches sur le cerveau statisticien suggèrent... qu'il est pourvu dès le départ de capacités abstraites pour le nombre et la probabilité et d'une machinerie corticale fonctionnelle qui compile déjà des statistiques détaillées de son environnement » (SD, p.58).

L'enfant dispose tôt de capacités mentales lui « permettant ... de tirer des inférences statistiques sur la présence, les intentions et les croyances des personnes qui l'entourent » (SD, p.59). L'organisation fonctionnelle complexe du cerveau remonte au fœtus et va de pair avec la longueur de la « maturation cérébrale » postérieure à la naissance, jusqu'à la fin de la puberté². Le cerveau statisticien qui révisé ses prédictions en fonction de l'expérience est précoce, mais la planification et le contrôle des actions, la réflexion et « l'apprentissage explicite » viennent plus tard. L'apprentissage des mots est aussi d'ordre statistique, via l'appréhension probabiliste de la succession des syllabes formant des mots. La structure cérébrale prédétermine les possibilités d'apprentissage, comme celui de la lecture qui mobilise un circuit cérébral précis. Et l'apprentissage scolaire fait retour sur l'anatomie et le fonctionnement du cerveau.

Le cerveau fonctionnerait ainsi depuis longtemps sur le mode du *machine learning* auquel accède les systèmes techniques contemporains, lorsqu'ils sont en capacité d'assimiler dans leur comportement les modèles construits à partir des données auxquelles ils ont accès : le comportement d'un tel système n'est donc plus prédéterminé mais contingent à cet apprentissage algorithmique. La nature du cerveau serait déjà de cet ordre-là mais, comme le dit assez curieusement Dehaene, « sans que nous en ayons conscience » : la conscience est-elle donc encore autre chose ? D'ailleurs, « ...toutes les grandes sociétés de logiciels comme Google ou Microsoft tentent de reproduire les algorithmes d'apprentissage statistique du cerveau humain » (p. 63). Le cerveau, statisticien inconscient ou intuitif ? Et même un peu économiste, puisque procédant par minimisation des risques ? La nature humaine bénéficierait d'une infra-conscience numérique et algorithmique.

D'ailleurs, pour Véronique Izard et Julien Marie, « ce sens des nombres est un système cognitif ancien, qui serait apparu très tôt au cours de l'évolution, et aurait été transmis ensuite au fil des espèces »³. « ... contrairement à ce qu'on a cru pendant longtemps, le nombre fait partie des toutes premières connaissances que l'être humain possède sur le monde » (SD, p.111). Le « calcul approximatif » est une capacité universelle et précoce de l'esprit humain. Pour autant, les difficultés de l'apprentissage explicite ne sont pas évitées, car « le comptage implique des nombres exacts qui vont au-delà de nos facultés intuitives » (SD, p.117). Ne serait-il pas plus correct alors de parler simplement d'intuition numérique approximative plutôt que de savoir statistico-mathématique inné, a fortiori s'il n'accède pas vraiment à la conscience ?

Entre l'homme et la machine s'instaure un jeu de miroirs qui prend l'allure d'un trompe-l'œil épistémologique. L'homme conçoit et construit des machines et mobilise la connaissance de leur fonctionnement pour modéliser celui de son propre cerveau ! C'est sans doute une stratégie efficace de modélisation, qui procède par analogie, sans qu'il soit garanti qu'elle livre une connaissance correcte et définitive du fonctionnement cérébral : un modèle est un outil de simulation provisoire du fonctionnement d'un système technique ou naturel. Le développement des techniques d'intelligence artificielle fournira sûrement d'autres de ces outils ! L'emprunt d'une technique de modélisation à une autre discipline pour la mobiliser, plus ou moins opportunément, dans une discipline distincte n'est pas rare : l'analyse économique a emprunté à la physique, à la biologie, à la

² Cf. le chapitre « AU TOUT DÉPART LE C3RV34U DU BEBE », par Ghislaine Dehaene-Lambertz et Jessica Dubois (pp. 77-104).

³ Chapitre « L'ENFANT, LE C3RV34U ET LES MATHÉMATIQUES », par Véronique Izard et Julien Marie (107-119).

météorologie... La mobilisation, par les neurosciences, de la cybernétique, de l'informatique, etc... pour penser le fonctionnement du cerveau relève au premier degré de la même démarche mais soulève une difficulté supplémentaire quant à sa pertinence ou sa justesse : celle de rabattre le fonctionnement cérébral sur l'état connu, mais daté, de certaines connaissances techniques. Le [modèle phrénologique](#) , sorte de cartographie mécaniste du cerveau, reste dans la mémoire scientifique pour rappeler les limites et les erreurs de cette démarche, qu'on pourrait qualifier d'analogique.

Jean-Pierre Changeux, qui a grandement contribué au développement des neurosciences cognitives et auquel Stanislas Dehaene rend hommage dans son introduction à *C3RV34U* en rappelant son ouvrage magistral de 1983, *L'Homme neuronal*⁴, se défiait d'ailleurs d'une analogie trop rapide entre le cerveau et la machine, en l'occurrence l'ordinateur : « Dans tout ordinateur construit par l'homme à ce jour, on distingue la bande magnétique-programme de la machine construite en "dur". Le cerveau humain, lui, ne peut se concevoir seulement comme exécutant un quelconque programme introduit par les organes des sens. Un des traits caractéristiques de la machine cérébrale est d'abord que le codage interne fait intervenir *à la fois*, nous l'avons vu, un codage topologique de connexions décrit par un graphe neuronique, et un codage d'impulsions électriques ou de signaux chimiques. Ici, la distinction classique "hardware-software" ne tient pas. D'autre part, il est évident que le cerveau de l'homme est capable de développer des stratégies de manière autonome. Anticipant les événements à venir, il construit ses propres programmes. Cette faculté d'auto-organisation constitue un des traits les plus saillants de la machine cérébrale humaine, dont le produit suprême est la pensée » (JPC, p.172). La conscience relève d'un bouclage dynamique entre le cortex, doté de ses fonctions localisées, et le tronc cérébral, doté de fonctions de base, et ce bouclage assure la régulation globale des fonctions cérébrales : « Le système de régulations mis en place évalue, apprécie résonances et dissonances entre concepts et percepts. Il devient mécanisme de perception des objets mentaux, de "hardware-software" de leur enchaînement. Les divers groupes de neurones de la formation réticulée s'avertissent mutuellement de leur action. Ils forment un "système" de voies hiérarchiques et parallèles en contact permanent et *réci-proque* avec les autres structures de l'encéphale. Une *intégration* entre centres se met alors en place. Du jeu de ces régulations emboîtées naît la conscience » (JPC, p.212). La complexité des régulations en question n'autorise pas la réduction définitive à un modèle trop machiniste. Cette tentation est courante chez les physiciens confrontés aux limites de la représentation de la nature par leur cerveau, lorsqu'elle mobilise des échelles très différentes de l'expérience humaine courante.

Antonio Damasio, neuroscientifique américain, rejoint le propos de Jean-Pierre Changeux par sa propre démarche, basée sur l'analyse des propriétés de régulation homéostatique des organismes vivants depuis l'origine de la vie⁵. Cette régulation mobilise les informations échangées au sein de l'organisme via la gamme des affects, des émotions les plus élémentaires aux sentiments les plus complexes. Le cerveau ne serait rien sans ces échanges avec le corps auquel il appartient : « L'esprit est le produit d'interactions entre le corps et le cerveau : il ne s'agit pas d'un phénomène purement cérébral » (AD, p.282)⁶. Cette conception conduit Damasio à récuser l'idée du « cerveau-ordinateur » et « l'interprétation algorithmique de l'humanité » : « ... leurs opérations [des circuits neuronaux] ne correspondent pas à un simple schéma *on / off* et ne peuvent être expliquées par cette simple conception numérique » (p.96). Comme chez Changeux, la conscience résulte de l'intégration

⁴ Jean-Pierre Changeux, *L'homme neuronal*, Fayard, 1983 (référence notée JPC dans le texte).

⁵ Antonio Damasio, *L'ordre étrange des choses. La vie, les sentiments et la fabrique de la culture*, Odile Jacob, 2017 (référence notée AD dans le texte).

⁶ Philosophiquement, c'est une filiation spinoziste plutôt que cartésienne : « Ce dualisme [corps-cerveau] a la vie dure : il est né à Athènes, a été érigé en principe par Descartes, a résisté aux féroces assauts de Spinoza et a été intensément exploité par les sciences informatiques » (AD, p.340).

complexe de fonctions corporelles et cérébrales, qui permet la subjectivité de l'individu dans la perception de soi-même et du monde et qui mémorise aussi ses expériences accumulées.

L'intelligence humaine a tellement bénéficié de l'évolution que l'homme est paradoxalement tenté de croire son cerveau autonome du corps et de l'assimiler aux machines qu'il crée. Plutôt que de copier la conception du cerveau sur celle d'une machine destinée à être dépassée un jour prochain, Jean-Pierre Changeux inversait d'ailleurs la démarche, dans un dialogue avec le mathématicien Alain Connes⁷: mobiliser la connaissance neuro-scientifique pour comprendre, dans un souci de bonne épistémologie contemporaine, le travail mental qui produit les raisonnements et découvertes mathématiques, les tours et détours « neuronaux » de ce travail, au cours duquel la méditation et l'intuition conditionnent la capacité à se convaincre de la validité globale d'une démonstration, à lui donner sens : la rationalité laborieuse, pas à pas, n'est pas auto-suffisante pour être convaincante, elle a besoin d'un état de confiance sous-jacent (ce que confirme Alain Connes dans la partie « L'illumination » de leur ouvrage commun et ce qu'illustre aussi très concrètement Cédric Villani dans sa belle introspection *Théorème vivant*⁸ à propos des travaux qui lui ont valu la médaille Fields). C'est le regard du démonstrateur sur la démonstration qui lui plénement sens. L'intelligence se joue dans cet emboîtement des niveaux de rationalité.

Sceptique face à la conception, défendue par Alain Connes, d'une réalité mathématique extra-cérébrale, Jean-Pierre Changeux ne dissocie pas le raisonnement mathématique à la fois des potentialités cérébrales et de l'évolution culturelle : au cours du développement individuel, de l'embryon à l'adulte, l'épigenèse connecte les neurones, élimine les redondances synaptiques et réorganise le cerveau « par stabilisation sélective des neurones et synapses » (JPC, p.327). « Apprendre c'est stabiliser des combinaisons synaptiques pré-établies. C'est aussi *éliminer* les autres » (JPC, p.329). Le processus d'apprentissage, qui produit l'organisation adulte du cerveau, est ouvert aux influences scolaires et culturelles relevant de l'environnement social : « L'intervention d'une épigenèse active par stabilisation sélective introduit une diversité nouvelle dans une organisation qui, sans cela, deviendrait redondante. Une ouverture sur le monde extérieur compense le relâchement d'un déterminisme purement interne. L'interaction avec l'environnement contribue désormais au déploiement d'une organisation neurale toujours plus complexe en dépit d'une mince évolution du patrimoine génétique. Cette structuration sélective de l'encéphale par l'environnement se renouvelle à chaque génération. Elle s'effectue dans des délais exceptionnellement brefs par rapport aux temps géologiques au cours desquels le génome évolue. L'épigenèse par stabilisation sélective économise du temps. Le darwinisme des synapses prend le relais du darwinisme des gènes » (JPC, p. 359). La socialisation de l'individu participe à son épigenèse individuelle et son contenu est évidemment dépendant du contexte historique et culturel. Le savoir mathématique lui-même n'est pas prédéterminé et s'inscrit dans ce contexte, résultat, pour partie, de l'histoire longue des sciences et des mutations des paradigmes scientifiques. Comme le dit Antonio Damasio, « par bien des côtés, l'histoire des cultures humaines est le récit de notre résistance aux algorithmes naturels grâce à des inventions que ceux-ci n'avaient pas programmées » (AD, p.289).

Pour prendre un exemple clé portant sur l'appréhension statistique de la réalité, l'induction statistique a été le produit d'une révolution épistémologique, qui s'est nouée dans des circonstances précises à la jointure des XVII^e et XVIII^e siècles et qui a permis aux contemporains de s'émanciper d'une fidélité dogmatique aux écrits et dires d'autorité pour s'ouvrir à une lecture expérimentale du monde⁹. Le recours à l'observation s'est appuyé sur le calcul des probabilités pour donner rigueur à la preuve empirique apportée par le raisonnement inductif mobilisant des faits élémentaires mais partiels. La preuve incorpore donc le doute raisonné, comme l'œuvre de Hume en témoigne : c'est le

⁷ Jean-Pierre Changeux et Alain Connes, *Matière à pensée*, Odile Jacob, 1989.

⁸ Cédric Villani, *Théorème vivant*, Grasset, 2012.

⁹ Voir Ian Hacking, *L'émergence de la probabilité*, Seuil, 2002.

refus de l'autosuffisance théorique censée démontrer les causalités à partir de principes premiers au sein d'un monde abstrait et mathématisable, sur le mode platonicien. L'induction probabiliste est aussi un résultat de l'évolution culturelle. Si les neuroscientifiques peuvent aujourd'hui qualifier le cerveau humain de cerveau statisticien ou bayésien, capable d'inférences statistiques informelles, c'est aussi parce que cette mutation culturelle est intervenue et que l'on sait désormais définir l'induction statistique. A supposer que la capacité cérébrale préexiste à cette mutation, elle n'était en tout cas pas qualifiable comme elle peut l'être aujourd'hui.

2. Le nouvel âge du naturalisme

Daniel Andler fournit le large cadrage nécessaire au bon traitement de cette discussion lorsqu'il s'interroge, dans une synthèse de référence, sur la place du naturalisme à l'heure des sciences cognitives¹⁰. Le naturalisme est une attitude fondée sur la conviction de l'unité de l'homme et du monde, sans référence transcendante ou métaphysique échappant par essence aux capacités d'appréhension de la réalité par la connaissance. C'est un refus de la « pensée du séparé », pour reprendre l'expression du philosophe américain John Dewey, figure d'un courant dit pragmatique qui s'efforce de penser ensemble la connaissance et l'action¹¹. L'évolution biologique rattache l'homme à la nature. Le cerveau cause la pensée qui produit la connaissance. Le naturalisme intégral est un système clos de causalités naturelles, à prétention exhaustive, jusqu'à développer une « science naturelle du social » (DA, p.48). Ce système endogénéise l'intentionnalité, produit de l'évolution et de l'apprentissage : la normativité, y compris sa dimension subjective et émotionnelle, participe aux faits prenant place dans les enchaînements de cause à effet, qu'elle soit d'ordre systémique ou relève des idéaux consciemment poursuivis par les agents. Et on peut repérer les biais comportementaux, les déviations et les erreurs des individus par rapport à la normativité rationnelle, face à des situations courantes : les théories de la décision se tiennent à la frontière de l'économie, de la logique, de la psychologie. Pour les opposants au naturalisme, le scientisme naturaliste « pourrait conduire à un monde déterministe, mécaniste, inhospitalier pour l'homme », jusqu'à « le dépouiller de son humanité » (DA, p.27). Il déborderait du discours scientifique pour intervenir sur la normativité des valeurs et du droit et prétendrait définir la rationalité normale en dehors des raisonnements pratiquement adoptés par les individus.

Pour autant, la tradition de pensée naturaliste s'est activement transformée et diversifiée, en fonction des contextes historiques et géographiques où elle a pris racine pour développer sa « conception séculière du monde ». Le pragmatisme américain s'écarte du positivisme scientifique à la française d'Auguste Comte mais conçoit la théorie de la connaissance comme « une province de la psychologie scientifique » (DA, p.40). Le naturalisme évolué n'est pas réductible à un scientisme, il est ouvert aux différentes modalités de la connaissance, il cultive les méthodologies pragmatistes et empiriques qui font de l'acte de connaissance un processus naturel et donc objectivable, mais qui a ses propres limites, lesquelles ne sont pas seulement d'ordre épistémique mais concernent les droits humains (par exemple, à propos de l'expérimentation sur autrui). On peut pratiquer le programme naturaliste comme méthodologie scientifique et être conscient de ces limites, qui participent de la réalité du monde, notamment parce que l'exercice de la science n'est pas pur mais pris dans un réseau d'intérêts particuliers et institutionnels. Les sciences cognitives contemporaines ouvrent un nouveau champ fécond dans le programme naturaliste, en outrepassant les découpages établis entre sciences de la nature et sciences humaines.

Au milieu du XX^e siècle, le développement de la cybernétique ouvre une voie vers la congruence entre les fonctionnements de l'esprit humain et de la machine. « L'ambition de la cybernétique était

¹⁰ Voir Daniel Andler, *La silhouette de l'humain, Quelle place pour le naturalisme dans le monde d'aujourd'hui ?*, nrf essais, Gallimard, 2016 ((référence notée DA dans le texte).

¹¹ Voir John Dewey, *La quête de certitude: Une étude de la relation entre connaissance et action*, nrf, Gallimard 2014.

de construire une théorie unifiée des machines complexes (capables d'auto-organisation et d'autonomie), de l'esprit et du cerveau humain (les deux étant à rapprocher, sans nécessairement se confondre) et de la société, à partir des concepts de contrôle et de communication » (DA, p.109). Or, si l'esprit versatile n'est pas d'emblée réductible à un algorithme, Alan Turing opère la jonction : « Le néo-mécanisme turingien combine deux idées anciennes : celle d'interaction causale au sein d'un système matériel et celle d'un enchaînement rationnel au sein d'un système symbolique... Ce parallélisme entre cause et raison est ce qui donne à la notion abstraite d'information son efficacité » (DA, p.110). Le coup d'envoi est donné à l'émergence d'une « théorie computationnelle de l'esprit » : « les sciences cognitives placent au cœur de leur ontologie l'information... et visent à rendre compte de la cognition comme d'un phénomène relevant de deux descriptions, l'une matérielle ou physique, l'autre informationnelle » (DA, p.111). Les états mentaux obéissent à une syntaxe formelle et leur transformation mobilise le support matériel cérébral.

Le développement des sciences cognitives s'inscrit dans une visée poly-naturaliste : « Les sciences cognitives abordent leur objet sous trois angles distincts : informationnel, neuroscientifique, évolutionnaire. Chacun ouvre une voie vers sa naturalisation... L'analyse informationnelle révèle la structure fonctionnelle des processus mentaux, les neurosciences montrent comment cette structure est réalisée dans le tissu neural, et la biologie évolutive explique comment les structures neurales qui existent chez *Homo sapiens* et confèrent à l'espèce ses capacités cognitives ont émergé » (DA, p.126). Dans une version qualifiée d'homogène, la théorie computationnelle de l'esprit (TCE) assimile l'esprit humain à « une vaste collection de programmes résidant côte à côte dans la mémoire de l'individu ou, ce qui revient au même, à un unique programme nourri d'une colossale banque de données dans laquelle il puise selon les besoins et la tâche du moment » (DA, p.132). A cette version homogène de la TCE s'opposent les théoriciens insistant sur la modularité de l'esprit humain, qui dispose de modules innés résultant de l'évolution et de modules acquis par l'apprentissage. « Certains modules, à l'image du mécanisme d'acquisition du langage, sont *présents dès l'origine* du développement de tout être humain » (DA, p.141). Cette position de principe ne suffit pas à résoudre l'opposition entre l'innéisme, qui croit à la présence innée de concepts fondamentaux (les structures fonctionnelles élémentaires du calcul et du langage inscrites sur le support cérébral) et de l'empirisme qui réduit la dotation initiale à une capacité générale d'apprentissage et renvoie le reste à l'apprentissage pratique, en insistant sur la plasticité cérébrale. « Les sciences cognitives naturalisent pièce par pièce des secteurs entiers de la sphère mentale, mais la naturalisation de l'esprit reste radicalement incomplète » (DA, p.174). La difficile naturalisation de l'intentionnalité et de la conscience, attributs fondamentaux de l'esprit, passe par la dé/re-construction de ces catégories héritées de la philosophie. La cognition humaine relève d'un couplage entre l'individu, son corps et l'environnement, si bien que le patrimoine cérébral de l'individu est pour partie externalisé dans des outils externes et des relations collectives.

Depuis les années 1950, le développement des neurosciences, sciences du cerveau et des phénomènes cérébraux, a bénéficié de l'impulsion cybernétique en travaillant « l'hypothèse structurante du cerveau comme organe de traitement de l'information » (DA, p.182). Le champ aujourd'hui reconnu des neurosciences cognitives est allé bien au-delà de cette impulsion originelle. Les neurosciences sont passées à l'observation, via l'imagerie cérébrale, quand bien même l'interprétation de cette dernière ne va pas de soi : l'oscillation entre le localisationnisme des fonctions cérébrales élémentaires et le connectivisme des fonctions complexes en témoigne (le candide a l'impression que la référence phrénologique, devenue *brain mapping*, hante l'usage de la neuro-imagerie et que l'interprétation de ses observations doit faire effort conceptuel pour s'en émanciper). Ce développement des neurosciences pénètre et perturbe les champs disciplinaires préétablis, jusqu'à faire jonction avec la théorie de la décision et la neuro-économie comportementale. Les neurosciences modifient voire abolissent la vision naïve de l'introspection naturelle et bouleversent les approches psychologiques. Leurs apports sont incontestables : la compréhension renouvelée de la vision, la découverte des neurones-miroirs comme support de la

compréhension d'autrui, par exemple. Pour autant, la maîtrise de la complexité du cerveau et de son fonctionnement reste un défi scientifique, comme en témoigne le propos toujours actuel, précédemment rappelé, de Jean-Pierre Changeux ou celui, plus récent, d'Antonio Damasio.

« L'hypothèse du cerveau bayésien » attire l'attention: « à partir de données incomplètes, le système recherche la conjecture la plus plausible : quelle est la structure de l'environnement ayant la plus grande chance d'introduire le stimulus que le système interprète ? » (DA, p. 235). Le système cérébral livre une reconstruction probabiliste de la réalité, plutôt qu'une simple perception immédiate. Le cerveau finit par substituer ses représentations aux perceptions, même si les premières mobilisent les secondes. L'idée était, on l'a vu, présente dès *L'homme neuronal* de Changeux: « Une parenté se dessine entre le percept, l'image et le concept, et en suggère la même matérialité neurale... L'hypothèse adoptée ici est que percept, image de mémoire et concept constituent des formes ou des états divers d'unités matérielles de représentation mentale, que nous regrouperons sous le terme d' "objets mentaux" » (JPC, p.179). Le processus d'apprentissage ne va pas sans aléa et incorpore une révision bayésienne de l'information sensorielle, codée comme une distribution de probabilité. La représentation du cerveau devient ainsi singulièrement complexe et l'intelligibilité de l'esprit s'avère rebelle à une introspection de sens commun. C'est loin d'être un processus achevé, voire achevable. Changeux proposait un modèle des objets mentaux fondé « sur l'identification d'unités mentales à des états d'activités physiques d'ensembles de neurones », en ajoutant aussitôt: « Le risque principal de cette entreprise est de rester trop simple, de ne pas rendre compte de l'ensemble des processus mentaux, d'être partielle » (JPC, p.225). Le modèle est toujours transitoire, l'observation et l'expérimentation circonscrites, leur interprétation délicate.

L'effort de compréhension de l'évolution cérébrale et cognitive, au fil du temps, aiguise la question du rapport entre l'individuel et le social. La reconnaissance de l'évolution biologique est un acquis scientifique et le cerveau ne peut en être tenu à l'écart, même si la biologie évolutionniste fait l'objet de remaniements fréquents, car l'expérimentation directe est rare dans l'administration de la preuve. L'évolution darwinienne s'inscrit dans l'héritage génétique comme « instinct cognitif », lequel fonctionne comme un algorithme face à des situations courantes (l'évitement d'un prédateur, d'un obstacle, d'un projectile...). Mais il est difficile de comprendre l'émergence rapide, à partir du néolithique, de *l'homo sapiens* moderne sur la base des lenteurs de l'évolution biologique antérieure (cent à cent cinquante mille ans pour que *l'anatomically modern Homo* donne naissance à l'homme moderne du néolithique): la formation et l'intériorisation d'une culture collective jouerait le rôle de dopant évolutionniste pour les comportements complexes, le « cerveau social » est porté par les rapports de parenté et d'alliance au sein des groupes et entre eux. Le penchant spontané des sciences cognitives et du naturalisme dont elles sont porteuses va cependant plutôt du côté de l'individualisme méthodologique, à l'encontre du culturalisme considérant que « la dimension sociale est constitutive de la cognition individuelle » (DA, p.300).

Comment envisager la synthèse ? « La cognition et la socialité humaines se sont détachées du règne animal en un même mouvement, en sorte que la cognition est la propriété de l'homme en société et que la société est organisée en sorte de rendre la culture possible, laquelle repose à son tour sur la transmission culturelle et démultiplie les capacités collectives individuelles » (DA, p. 306). Pour Antonio Damasio, la socialité participe aux stratégies comportementales qui contribuent conjointement à la régulation homéostatique et aux évolutions culturelles. L'homme agit sur les conditions de son évolution : « L'environnement et la population concernée sont ainsi engagés dans un processus coévolutif » (DA, p. 307). « Ce qu'il s'agit d'imaginer est un processus intégrant les deux moteurs, génétique-cognitif et culturel-progressif » (DA, p.309). L'évolution naturelle sur le mode darwinien (des mutations aléatoires soumises à un filtrage adaptatif) est surdéterminée par les interactions entre groupes et individus qui décident, plus ou moins intentionnellement, de l'évolution culturelle. L'accumulation d'un patrimoine culturel collectif et transmissible, ainsi que la coopération à large échelle, au-delà des petits groupes parentaux, permettent à l'intelligence

individuelle de n'être pas conditionnée par le seul espace-temps de l'évolution biologique. Le déverrouillage frontal et les mutations des gènes de communication ont libéré le couple émotion / intellect, favorisé la socialisation et généré une plasticité du cerveau humain apte à l'apprentissage rapide et cumulatif, bien au-delà de la géologie du génome. La taille du groupe qui coopère est facteur de développement endogène, en son sein, des capacités culturelles qui s'impriment dans les cerveaux individuels. Si le naturalisme veut incorporer une conception de la socialité, il ne peut s'en tenir à la normativité faible, pour laquelle les *habitus* individuels ne sont jamais que des routines algorithmiques incorporées par les individus. La normativité forte intègre dans la formation de ces *habitus* le rapport à autrui, dans le contexte de la rencontre circonstanciée avec lui, et incorpore la délibération, intime ou collective, dans la définition des actions intentionnelles : cette délibération est nécessaire à l'appréhension de la singularité d'une situation.

Au bout du compte, Daniel Andler prône un « naturalisme critique », ouvert et pragmatique, qui, sans forcer la convergence des disciplines scientifiques, pratique la confrontation raisonnée et productive de leurs approches et contribue ainsi au fondement d'un humanisme robuste, conscient de l'unité humaine, de son ancrage naturel et de la pluralité des approches de la socialité, qui explorent « l'intentionnalité individuelle et collective à l'œuvre dans l'histoire humaine » (DA, p.331)

3. L'*homo œconomicus* recyclé ou réinventé ?

Prenant ses distances à l'égard de l'*homo œconomicus* à la rationalité utilitaire, mais capable d'exploiter parfaitement l'information disponible sur son environnement présent et futur, le *mainstream* des économistes s'est sensiblement renouvelé depuis le siècle dernier, à la recherche de la jonction avec les neurosciences pour une conception plus réaliste du sujet économique, où l'expérimentation et l'économie comportementale prennent une place accrue. Une personnalité comme Daniel Kahneman, psychologue et économiste, ou psychologue passé à l'économie, lauréat du Prix Nobel d'économie en 2002, est typique de cette jonction. L'axiomatique de l'*homo œconomicus* s'est imposée, durant tout le XX^e siècle, comme la scolastique de référence du courant dominant de la pensée économique, dit néo-classique, inauguré par les pères fondateurs Jevons, Walras, Pareto : résiliente face aux critiques d'une litanie de grands rebelles (Veblen, Keynes, Simon, etc.), elle a procuré une forte puissance analytique, source de nombreux et ramifiés développements, au prix de l'attrition de sa capacité de compréhension de la diversité des comportements humains et de l'affaiblissement de ses interactions avec les autres sciences sociales. Bien des économistes admettaient le caractère réducteur de cette axiomatique mais considéraient que c'était là le prix à payer pour développer des modèles analytiques livrant des conclusions claires et recevables, l'écart à la réalité étant considéré de l'ordre de l'approximation.

Le tournant pris n'est donc pas mince. C'est en particulier celui de l'économie comportementale, sensible à la prégnance des choix non-rationnels et des biais cognitifs ou émotionnels au regard du modèle standard de l'*homo œconomicus* et ouverte à l'expérimentation en laboratoire sur le modèle de la psychologie. La jonction avec les neurosciences offre donc la possibilité d'une naturalisation réaliste de la science économique. Les développements passionnants, par les auteurs actifs à cette jonction, ne manquent pas : la coopération, l'altruisme, les normes sociales, les émotions, etc. comme produit de l'évolution humaine et facteur de régulation des rapports économiques.

Pour autant, le champ des neurosciences étant lui-même foisonnant et pluriel, à la recherche inachevée de sa cohérence, cette jonction ne procure pas d'emblée aux économistes transfuges une axiomatique cohérente et définitivement délivrée des étroitesse de l'*homo œconomicus*. Naturellement, pourrait-on dire ironiquement, ces économistes transportent avec eux l'individualisme méthodologique dont ils sont bardés depuis longtemps et rencontrent préférentiellement les approches neuro-scientifiques dont ils se sentent parents. C'est le cas de la psychologie évolutionniste classique, désignée par Daniel Andler comme l'école de Santa Barbara :

l'homme est le produit de l'évolution biologique et la société émerge des interactions entre individus, à la différence d'une approche davantage holiste privilégiant l'hypothèse du « cerveau social », pour laquelle la dimension sociale est constitutive de la cognition individuelle. Ce naturalisme individualiste convient à nombre d'économistes engagés dans l'économie comportementale et expérimentale, sans les inciter à vraiment aborder la compréhension de la « socialité profonde » propre à l'être humain et donc à renouveler l'articulation aux autres sciences sociales. Or, comme le dit Daniel Andler, « la théorie naturaliste de l'individu semble elle-même en difficulté, comme si, privé de sa dimension sociale, l'individu modélisé était trop éloigné de l'homme réel pour donner lieu à une représentation scientifique plausible » (DA, p.330). La tradition naturaliste s'est suffisamment diversifiée pour n'être pas réductible à cette seule approche et porter des conceptions de la « socialité profonde » et de la « normativité forte ».

Il y a donc le risque que les économistes se contentent d'une version basique du naturalisme, qui les arrange plus qu'elle ne les remette en question, en leur permettant un recyclage naturaliste du choix rationnel. Dans cette version, qualifiée de « monde rêvé » du naturaliste par Daniel Andler (DA, p. 338), l'agent intelligent n'est certes plus *l'homo œconomicus* prétendant à l'omniscience, il fait des erreurs de mesure et est surpris par les aléas, mais il adopte le modèle de « l'intendant », ainsi paraboliquement baptisé par Daniel Andler : « il dispose d'un répertoire de règles générales qui lui permettent de faire face de la meilleure façon possible à toute situation » (DA, p.339). Il adopte un mode calculatoire fait de l'application d'un répertoire de règles à la diversité des situations. Le *new mainstream* économique est tenté de se contenter d'un naturalisme réducteur, qui réduit le contexte à un ensemble d'informations à traiter et le comportement à un algorithme traitant ces informations et répondant aux incitations qui en découlent. La rationalité instrumentale de l'agent rationnel est sauvée. Mais le contexte réel est difficilement réductible à un monde gérable par un répertoire de règles rationnelles s'appliquant à un ensemble exhaustif de situations imaginables, via le traitement de l'information sur ces situations et la prise en compte des préférences individuelles pour paramétrer l'application de la règle. Dans une approche pragmatique, comme celle pratiquée par John Dewey, l'action participe à la connaissance et compose avec l'incertitude : les règles et les situations interagissent fondamentalement, car les singularités de la réalité sont inéliminables et pas toujours réductibles à des aléas probabilisables. L'exemple de la pratique médicale est probant à cet égard. Une situation inédite peut être rapportée à une situation déjà rencontrée et participer à un apprentissage expérimental, sans être réductible à la rationalité calculatoire. Au bout du compte, « le comportement approprié... n'est pas ce que mesure la distance objective à un idéal objectif, déterminé par le théoricien et accessible par une intelligence parfaite ; c'est ce qui est jugé tel au terme d'un ensemble de raisons. Que tel comportement soit approprié n'est pas un *fait* de nature. C'est un jugement susceptible d'être à tout moment contesté et réexaminé » (DA, p.356). La rationalité est délibérative tout autant que calculatoire, et en tout cas faillible et donc révisable.

Dans son *Economie du bien commun*¹², Jean Tirole s'attache notamment à l'économie des comportements ordinaires des individus, à l'intersection de leurs préférences et des incitations qui s'exercent sur eux. La réalité humaine quotidienne est faite de biais cognitifs, de procédés heuristiques, de « raisonnements raccourcis », de croyances endogènes qui permettent les décisions et actions quotidiennes mais peuvent comporter des désajustements par rapport à la réalité prévalente ou par rapport à la visée de l'intérêt collectif. Cette approche d'inspiration naturaliste renvoie à l'obsolescence nombre de débats et d'idées entre « économistes anciens ». Elle peut faire l'objet d'une initiation « pragmatique,... à la fois intuitive et rigoureuse » (JT, p.47). Marie-Claire Villeval, spécialiste de l'économie comportementale, adopte le même type de démarche¹³ : il faut prendre en compte explicitement « les forces psychologiques et les biais des individus dans leurs

¹² *Economie du bien commun*, Jean Tirole, PUF, 2015 (référence notée JT dans le texte).

¹³ Voir son ouvrage pédagogique : Marie-Claire Villeval, *L'économie comportementale du marché du travail*, Presses de Sciences Po, 2016 ((référence notée MCV dans le texte).

modes de décision » (MCV, p.7) pour envisager les bonnes politiques publiques. « ...les actions peuvent obéir à une rationalité autre que celle fondée sur la maximisation des seuls gains monétaires individuels, et elle accorde une véritable place aux biais cognitifs, aux biais de jugement et aux émotions » (MCV, p.9). On sera d'accord pour avoir une modélisation aussi pertinente que possible des préférences, des modes de décision et des jugements des individus, mais on se demande par rapport à quelle norme humaine prédéfinie leurs biais sont appréciés comme tels : quel est l'arrière fond psychologique d'un comportement rationnel non biaisé et le théoricien en est-il suffisamment maître pour se sentir autorisé à définir la norme, comme si, lui, n'avait pas de biais cognitif ? Par exemple, « Les demandeurs d'emploi n'adoptent pas le comportement annoncé par la théorie... » (MCV, p.19). La référence utilisée pour juger de la rationalité est essentielle pour savoir ce qu'on peut tirer et généraliser solidement de l'économie comportementale expérimentale : s'il s'agit de probabilités objectives de retour à l'emploi (mais estimées par qui et comment ?) que le demandeur d'emploi mesure mal, c'est autre chose qu'une pure représentation théorique, sortie du cerveau de l'économiste théoricien. L'économie comportementale apporte beaucoup de résultats recevables, par exemple sur le rôle des incitations non monétaires, qui retrouvent d'ailleurs les bonnes pratiques du management, sans que ça dispense d'interroger son cadre épistémologique, qui peut conditionner la portée de ces résultats.

L'élaboration rigoureuse d'hypothèses et de modèles théoriques est une étape obligée pour orienter le traitement de données toujours plus foisonnantes, a fortiori si le chercheur souhaite dépasser une approche descriptive des corrélations entre phénomènes observés et identifier leurs causalités explicatives, sur lesquelles les agents privés et publics puissent prendre appui pour agir. Une bonne description statistique plante le décor en ordonnant les phénomènes observés, mais pour reprendre une métaphore de Julien Gracq¹⁴, si elle ne veut pas rester « dévoilement quiétiste de l'objet » mais « constituer les préliminaires d'une dramaturgie », elle doit être relayée par l'effort théorique pour définir les causalités envisageables et faire retour sur les données ordonnées en vue de tester ces causalités. En ce sens-là, l'adage de Jean Tirole est recevable, s'il n'est pas compris comme un mépris de l'étape de mise en forme et de description statistiques : « Sans modèle à tester, les données ne révèlent pas grand-chose d'utilisable pour la politique économique » (JT, p.148). Mais le modèle théorique n'est pas unique : c'est un compromis, parmi d'autres possibles, entre des hypothèses cruciales « raisonnablement réalistes » et des hypothèses secondaires « simplificatrices », moins décisives pour les résultats. La modélisation discipline l'exercice analytique de la pensée et permet d'afficher des conclusions résultant de démonstrations explicites mais falsifiables par le recours aux tests empiriques, qu'ils relèvent de l'économétrie classique ou de protocoles expérimentaux. Cette démarche épistémologique, qui se réclame de la référence à Karl Popper, est commune à nombre d'économistes relevant du *mainstream*. Elle fut, par exemple, très clairement développée par Edmond Malinvaud¹⁵. Il prônait « l'alternance dialectique entre l'observation et la modélisation » (EM, p.46), cette dernière étant conçue comme une spécification instrumentale des concepts et des relations théoriques. Mais il affichait son scepticisme envers une identification purement inductive des causalités, telle qu'elle fut alors prônée par certaines approches empiriques des séries temporelles. Il défendait l'autonomie du moment théorique afin d'éviter une modélisation qui colle trop bien aux données par un paramétrage astucieux : « Une telle pratique n'est évidemment pas conforme aux principes de la méthodologie économétrique puisque celle-ci considère le modèle comme spécifié indépendamment des données » (EM, p.367). Malinvaud, modeste, considérait l'axiomatique de la rationalité individuelle, facteur de décision optimale, comme une approximation recevable des calculs économiques courants des agents basant leur intuition sur l'expérience acquise (EM, p.89) : cette axiomatique facilite les raisonnements analytiques du théoricien plus qu'elle ne prétend énoncer la vérité stricte des comportements réels.

¹⁴ Julien Gracq, *En lisant, en écrivant*, José Corti, 1980.

¹⁵ Edmond Malinvaud, *Voies de la recherche macroéconomique*, Odile Jacob, 1991 (référence notée EM dans le texte).

L'adhésion à cette démarche ne devrait donc pas autoriser le théoricien à se croire dépositaire du référentiel de la rationalité qui permet de juger, à cette aune, du comportement des agents, de leurs biais cognitifs et de leur irrationalité pratique : il y a là un glissement de la pensée qui ne va pas de soi. Avec ce glissement, on passe du pluralisme des modèles envisageables et testables à l'idée d'un monde platonicien unique dont la perfection rationnelle, élucidée par le théoricien, précède la réalité dégradée faite d'imperfections et de biais : la réalité relève d'une chute de la perfection platonicienne. Si *l'homo psychologicus* est « un acteur pas toujours rationnel » (JT, p.198), le théoricien prétend détenir les clés de cette rationalité. Par exemple, dans le domaine du marché du travail et de ses réformes, si les salariés anticipaient rationnellement les bienfaits des réformes supposés démontrés par les experts, ils les accueilleraient plus favorablement : ce genre d'affirmation vire à la pétition de principe, sauf à supposer une sur-rationalité de l'expert et à renoncer à comprendre les motivations spécifiques des salariés. L'économiste prend en compte, par souci de réalisme, les facteurs psychologiques, sociaux, juridiques, etc. qui influencent le comportement pratique de l'individu mais ces facteurs sont réduits à leur impact sur la composition rationnelle des préférences et des incitations. « Dans leur interprétation du droit, les économistes voient avant tout un ensemble d'incitations » (JT, p.198). C'est sans doute un peu court pour recevoir l'assentiment des juristes. Jean Tirole envisage favorablement la convergence entre la science économique et la biologie évolutionniste, comme moteur d' « une réunification progressive des sciences sociales », sur « le modèle de la convergence qui existait jusqu'à la fin du XIX^e siècle » (JT, p.202). Le cerveau calculateur et utilitariste des économistes serait-il le produit de la sélection naturelle et cet « *homo darwinus* » (JT, p. 201) la version enfin naturalisée d'*homo œconomicus* ?

Ce que les neurosciences nous promettent, c'est un apport inédit et substantiel à une meilleure compréhension des tours et détours des raisonnements et des comportements humains, inscrits dans les interactions entre le corps, le cerveau et autrui. Il serait dommage que les économistes soucieux de convergence avec les neurosciences inhibent cette promesse par l'importation d'une conception trop préconçue de la rationalité, héritée d'économistes aussi anciens que d'autres, et par la réduction de la prise en compte des neurosciences à un naturalisme numérique étroit dont se distancient eux-mêmes certains neuroscientifiques. Les apports incontestables de l'économie comportementale s'en trouveraient bridés, faute de réellement s'émanciper d'une approche scolastique de la rationalité économique. Au bout du compte, c'est l'axiomatique même de l'*homo œconomicus* qui devrait se trouver radicalement révisée par une attention ouverte aux développements des neurosciences.