

## Modèles de la rationalité animale : interprétation *versus* architecture

Benoit Hardy-Vallée  
Department of Philosophy  
University of Toronto  
170 St. George St., 4th Floor,  
Toronto, ON, M5R 2M8  
Phone: (416) 978-3316  
Fax: (416) 978-8703  
Email: [ben.hardy.vallee@utoronto.ca](mailto:ben.hardy.vallee@utoronto.ca)  
<http://hardyvallee.net>

**Résumé :** Pour plusieurs philosophes, une solution à la question de la rationalité animale (à savoir si, et en quel sens, on peut les considérer comme des agents rationnels) dépend, en bout de ligne, de notre capacité à les interpréter. Dans cet article, je critiquerai cette façon de faire et montrerai comment une approche alternative est possible et souhaitable. Dans un premier temps, je reconstruirai l'argumentaire de cette conception « interprétiviste ». J'argumenterai par la suite à l'effet que cette conception mène à la mise au jour d'espèces nominales, et non naturelles. Bien que cela puisse constituer un but légitime en soi, je soutiendrai qu'une autre conception dite « architecturale » permet une approche plus rigoureuse de la rationalité, apte à identifier des espèces naturelles. Elle consiste à identifier les mécanismes responsable de ce que je nommerai la compétence rationnelle. Je soutiens, par la suite, que l'écologie du comportement supporte une description du comportement animal comme étant économiquement rationnel, puis qu'on peut aujourd'hui identifier, dans leur grandes ligne, les mécanismes neuronaux responsables de la compétence rationnelle. Je suggère, conséquemment, que le concept d'agent rationnel inclue d'autres espèces que la nôtre.

**Mots-clé :** rationalité, naturalisme, animal, décision, espèces naturelles, neuroscience, interprétation

**Nb de mots:** 8389

**Abstract:** For several philosophers, answering the question of animal rationality (that is, if, and in which sense, they can be considered as rational agents) depends, in the end, of our capacity to interpret them. In this article, I will criticize this strategy and will show how an alternative approach is possible and desirable. First, I will try to reconstruct the “interpretivist” conception. I will argue afterward that this conception lead to the identification nominal, instead of natural, kinds. Although that can constitute a legitimate goal, I will argue that another approach known as “architectural” allows a more rigorous approach of rationality, able to identify natural kinds. It consists in identifying mechanisms underlying what I call the rational competence. I contend thereafter that behavioral ecology supports a description of animal behavior as economically rational, then that it is nowadays possible to broadly identify, the neuronal mechanisms of rational competence. I suggest consequently the concept of rational agent includes other species that ours.

**Key words:** rationality, naturalism, animal, decision, species natural, neuroscience, interpretation

**Word count:** 8389

## Modèles de la rationalité animale : interprétation *versus* architecture

*Il n'y a point d'art qui apprenne à découvrir sur le visage les inclinations de l'esprit*

- Macbeth, Acte I, Scène IV

### 1 Introduction

Lorsqu'on offre à des sujets le choix entre deux options, A et B, la première étant incertaine mais très désirable, la seconde étant moins désirable mais assurée, les sujets préfèrent systématiquement la seconde, même si en multipliant la probabilité par l'utilité des options, la première s'avère plus intéressante. Les sujets démontrent ainsi une *aversion au risque*, ce qui tendrait à montrer que les agents ne se conforment pas aux théories normatives. Dans une autre expérience, les sujets doivent cette fois-ci choisir à répétition entre une option A et une option B, lesquelles ne sont pas également payantes. Un agent parfaitement rationnel devrait systématiquement choisir l'option qui s'est révélée la plus payante, alors que dans les faits, les sujets *appariement les probabilités* : si A a été payante dans 70% des cas et B dans 30%, les sujets ont tendance à essayer A dans 70% des cas et B dans 30%, ce qui au final est moins payant que d'essayer A dans 100% des cas. Ces résultats ne sont pas en soi si surprenants : il n'est pas inattendu que des agents réels ne se comportent pas comme le prédisent des théories qui décrivent des agents idéalisés. Sauf que dans ce cas-ci, les premiers sujets sont des rats, les seconds des singes, et que les « choix » étaient des quantités de nourriture ou des gouttes de jus.<sup>1</sup>

Est-il vraiment justifié de dire que ces animaux ont pris des décisions *irrationnelles* ? Pour poser un geste irrationnel, il faut bien être, la plupart du temps, rationnel : un caillou ou une voiture ne peuvent pas être irrationnels parce qu'il n'y a aucun sens à dire qu'ils sont rationnels. Y a-t-il toutefois un sens à dire que des animaux sont rationnels ou irrationnels ? Bien souvent,

pour répondre à cette question, on se demande dans quelle mesure les animaux ont des « théories de l'esprit », des concepts, un langage, des croyances, etc., Ces questions, cependant, touchent plus à la rationalité *théorique* (la rationalité des pensées, croyances et inférences). Il s'agit là d'un certain type de rationalité et on ne peut affirmer *a priori* que la rationalité se réduit à *penser* rationnellement. En économie, l'agent rationnel n'est pas celui qui pense rationnellement, mais qui démontre des *préférences* rationnelles, révélée par le comportement : s'il préfère A à B, il choisira A à B. Le consommateur qui préfère acheter un item A à un item B parce que le premier est moins coûteux démontre une forme de rationalité différente de la capacité à articuler rationnellement des pensées. Si la question de la rationalité des pensées animales est justifiée, la question de la rationalité de leurs action l'est aussi.

Pour plusieurs philosophes, cette question dépend, en bout de ligne, de notre capacité à interpréter les animaux: comme un agent rationnel est une créature à laquelle on peut attribuer des croyances et désirs, un animal rationnel si on applique avec succès « le principe de rationalité » ou encore la « posture intentionnelle ». Dans cet article, je critiquerai cette façon de faire et montrerai comment une approche alternative est possible et souhaitable. Dans un premier temps, je reconstruirai l'argumentaire de cette conception *intéprétiviste* (partie 2). J'argumenterai par la suite à l'effet que cette conception mène à la mise au jour d'espèces nominales, et non naturelles (partie 3). Bien que cela puisse constituer un but légitime en soi, je soutiendrai qu'une autre conception dite *architecturale* permet une approche plus rigoureuse de la rationalité, apte à identifier des espèces naturelles. Elle consiste à identifier les mécanismes responsable de ce que je nommerai la *compétence rationnelle*. Je soutiens, par la suite, que l'écologie du comportement supporte une description du comportement animal comme étant économiquement rationnel (partie 4), puis qu'on peut aujourd'hui identifier, dans leur grandes ligne, les mécanismes neuronaux responsables de la compétence rationnelle. Je suggère, conséquemment, que la

catégorie des agents *économiquement* rationnels incluent d'autres espèces que la nôtre si on veut proposer une conception naturaliste de la rationalité (partie 5).

## 2 La conception interprétative de la rationalité animale

Dans son *Traité de la nature humaine*, Hume attribuait sans hésitation une raison aux animaux :

*Nous sommes conscients que nous-mêmes, en adaptant des moyens aux fins, nous sommes guidés par la raison et l'intention (...)  
Quand donc nous voyons d'autres créatures (...), accomplir de semblables actions et les diriger vers des fins semblables, tous nos principes de raisonnement et de probabilité nous portent avec une force invincible à croire à l'existence d'une cause semblable (Traité de la nature humaine, I, III, XVI)*

La méthode par laquelle Hume en arrivait à une conclusion sur la nature de la rationalité animale est de nature essentiellement interprétative. Donald Davidson et Daniel Dennett ont, chacun dans une perspective différente, employé une méthode semblable pour déterminer l'extension de la catégorie AGENT RATIONNEL<sup>2</sup>. Elle consiste en trois étapes :

- (a) Définir RATIONALITÉ, ou le concept d' AGENT RATIONNEL et spécifier les différents traits, caractères ou compétences T que possèdent les agents rationnels.
- (b) Déterminer si une créature X est rationnelle en déterminant si X est rationnellement interprétable selon les concepts définis en (a)

(c) Si X s'avère rationnel au sens de (b), attribuer les traits T à X.

Déterminer, dans le monde vivant, qui sont les agents rationnels demande qu'on isole certains traits, ou caractères<sup>3</sup> (T), qu'une espèce possède et qui font que cette espèce est constituée d'agents rationnels. T peut par exemple être le langage ou la psychologie intuitive (*folk-psychology*), ainsi que le soutient Davidson.

Un animal rationnel, affirme Davidson, est une créature capable d'entretenir des attitudes propositionnelles (Davidson, 1991: 63). Lorsqu'on décrit le comportement d'une personne en terme de croyances et désirs, on montre « la rationalité de l'action à la lumière du contenu de la croyance et de l'objet du désir (Davidson, 1993: 234) ». Avoir une attitude propositionnelle, c'est aussi en avoir une panoplie : si un agent possède le concept CHÊNE, et que ce concept spécifie qu'un chêne est un arbre, alors un agent qui croit qu'il est près d'un chêne doit aussi pouvoir se faire créditer la croyance qu'il est près d'un arbre. Parmi les concepts qu'un agent rationnellement interprétable doit posséder figure le concept de CROYANCE. En effet, croire que P implique croire que l'on croit P. On croit que l'on croit que P lorsqu'il y a possibilité de surprise, laquelle suppose un contraste entre ce qu'on tenait pour vrai et ce qui est vrai, ce qui signifie qu'on distingue le vrai du faux, et l'objectif du subjectif. L'interprétabilité et les capacités cognitives et langagières qu'elle présupposent figurent dans une théorie implicite, « l'idéal constitutif de rationalité » (1993: 222), condition de possibilité des attributions de croyances et de désirs.

Daniel Dennett a proposé une analyse semblable de rationalité, qui prolonge la réflexion de Davidson, dans un cadre naturaliste. La notion de rationalité, selon lui, s'articule dans le cadre de stratégies prédictives; nous avons intuitivement des attitudes, points de vue ou des « postures » (*stances*) à l'égard de certains types de systèmes avec lesquels nous pouvons interagir. Que ce

soit une pomme qui tombe d'un arbre (posture physique), un logiciel d'échecs (posture du design) ou un vendeur d'aspirateurs, (posture intentionnelle) dès que nous entrons en contact avec ces systèmes nous avons une certaine tendance à générer des prédictions quant à leur évolution. Dans le cas des systèmes intentionnels, nous présupposons la rationalité de l'agent, à savoir que le système en question est doté de croyances et de désirs.

Dans son principe général, la méthode interprétative consiste donc à faire de la question de la rationalité animale une tentative d'interprétation du comportement animal, laquelle utilise les concepts dérivés de notre psychologie intuitive (*folk-psychology*)<sup>4</sup>. Cette façon de faire est relativement courante dans les discussions sur la rationalité animale : si être rationnel, c'est être rationnellement interprétable, y a-t-il un sens selon lequel des animaux sont rationnellement interprétables ? On peut relaxer certaines contraintes (la possession du langage par exemple) et essayer de donner un sens aux actions animales comme on en donne aux actions humaines, afin de voir s'ils sont capables d'avoir des croyances et des désirs, des états conscients, des capacités conceptuelles, inférentielles ou communicationnelles équivalentes ou semblables aux nôtres. Comprendre l'esprit animal devient alors un projet qui ne succèdera que si on peut se « mettre dans la peau » de l'animal en question pour deviner ce qu'il pense. Dennett donne comme exemple de l'application de la posture intentionnelle aux animaux celui du pluvier et de son comportement de « tromperie »; lorsque celui-ci détecte un prédateur, il se comporte comme si son aile était brisée, et « feint » d'être blessé. La tromperie est-elle une réelle tromperie intentionnelle, qui révèle la rationalité de l'agent, ou un simple réflexe ? Pour ce faire (Dennett, 1987, p. 258) propose d'interpréter le comportement du pluvier en lui attribuant un soliloque :

*[Un prédateur] pourrait être distrait par son désir de m'attraper  
et de me manger, mais seulement s'il pensait qu'il y a une chance*

*raisonnable qu'il me mange; il contracterait juste cette croyance si je lui donnais des raisons de croire que je ne peux plus voler; je pourrais faire cela en feignant d'avoir une aile brisée.*

Si on peut utiliser avec succès la posture intentionnelle, si on montre que l'animal est sensible à un ensemble complexe de conditions environnementales ainsi que le présuppose le soliloque, on peut dire qu'il entretient et attribue des croyances et des désirs. Si des tests montrent que son contrôle cognitif réagit à des patrons environnementaux plus simples, il faut alors réévaluer sa rationalité à la baisse. Cette procédure est la même, selon Dennett, que nous utilisons entre nous : on suppose les agents informés et capables d'inférences logiques, on prédit leur comportement, et on réévalue cette idéal de rationalité à la baisse s'il y a lieu.

On peut alors se demande si un autre animal, par exemple le chimpanzé, a lui aussi une « théorie de l'esprit » (Premack & Woodruff, 1978). Supposons que le chimpanzé ait réussi tous les tests nécessaires et que les primatologues soient conduits à lui attribuer une théorie de l'esprit<sup>5</sup> : il serait alors justifié de dire que le trait *avoir une théorie de l'esprit* est réalisé chez le chimpanzé et l'humain, et donc la présence de ce trait, qu'on tient pour un critère de rationalité, justifie le fait qu'on considère le chimpanzé comme un agent rationnel. C'est donc par une *analogie* qu'on justifierait l'extension de la classe des agents rationnels aux chimpanzés, un raisonnement récurrent depuis Hume.

Plus récemment, Susan Hurley (2003) a défendu une stratégie interprétive pour étudier la possibilité d'une rationalité animale, selon laquelle l'agentivité rationnelle dépend d'une caractérisation de l'animal au niveau « personnel » (distinct du niveau *subpersonnel*) : si on peut lui attribuer des capacités qui font office de raisons d'agir (des perceptions et des intentions motrices), que celles-ci satisfont aux réquisits holistes (à savoir qu'elles peuvent être combinées

et recombinaées) et normatifs (les raisons d'agir peuvent faire que l'agent commet des erreurs), alors on peut considérer que l'animal est rationnel.

Cette façon de procéder est toutefois problématique non pas relativement aux conclusions qu'on en tire, mais au type de connaissance qu'elle peut produire. En effet, quand Davidson affirme qu'« être un animal rationnel, c'est précisément avoir des attitudes propositionnelles » (Davidson, 1991: 63), et que ce critère constitue une condition nécessaire et suffisante pour être un animal rationnel, nous n'en savons pas plus sur les animaux, rationnels ou non, mais disposons plutôt d'un « mode d'emploi du langage intentionnel » (Proust : 1997 :54). Cette méthode est moins susceptible de nous apporter de nouvelles connaissances que de consolider des connaissances ou croyances déjà établies. L'interprétivisme souscrit en effet à une règle de catégorisation et d'individuation des traits selon leur *fonction*, une façon de faire qui, bien qu'elle soit populaire chez les philosophes<sup>6</sup>, ne trouve que peu d'écho en biologie.

Selon l'interprétivisme, si les animaux peuvent être interprétés, c'est que, comme nous, ils possèdent un trait ou une collection de traits T qui en font des agents rationnels. La catégorie AGENTS RATIONNELS est alors réduite à la catégorie AGENTS DOTÉS DE CAPACITÉS QUI LES REND INTERPRÉTABLES, la capacité variant selon la théorie de la rationalité. Cela équivaut à regrouper des traits selon une identité de fonction : le trait T<sub>1</sub> chez l'humain et T<sub>2</sub> chez le chimpanzé (ou tout autre animal) sont identiques parce qu'ils ont la même fonction, à savoir entretenir des croyances sur les croyances d'autrui. Le trait en question ne peut qu'être individué par sa fonction, car ce n'est pas un trait anatomique, physiologique ou comportemental auquel on fait référence lorsqu'on identifie les traits qui font qu'un animal est rationnel. La fonction en question peut être conçue comme la fonction adaptative ou son rôle causal (les deux principales familles de théories des fonctions biologiques). En anatomie comparative et en cladistique,<sup>7</sup> un

regroupement de traits par des critères d'identité de fonction est une *analogie* ou *homoplasie*<sup>8</sup> : les traits A et B des espèces X et Y sont homoplasiques ssi A et B ont la même fonction F.

Le problème que j'aimerais soulever ici ne concerne pas le raisonnement analogique en soi, mais le raisonnement analogique *en biologie*. En spécifiant un critère qui individue un trait, on détermine par le fait même une catégorie de traits ou d'organismes dotés de ce trait. Lorsqu'on individue le trait « avoir des ailes » par la fonction *se propulser de façon autonome dans un milieu aérien*, on crée une catégorie parmi laquelle on compte entre autres : des ailes d'insectes, de chauves-souris, d'oiseaux (si on élimine pingouins et autruches), et de ptérodactyles (si on inclut les espèces éteintes). Les classes d'entités créées par des règles de catégorisation analogiques, soutiendrai-je dans la section suivante, sont des espèces nominales, alors que l'ontologie de la science requiert des espèces *naturelles*.

### **3 Espèce naturelles et méthode architecturale**

Le problème de l'intéprétivisme touche à la nature des catégories ontologiques que cette méthode identifie : plutôt que des espèces naturelles<sup>9</sup> ce sont des espèces nominales qui sont identifiées. Or ce sont les premières, mais non les secondes, qui constituent à la fois une connaissance authentique et qui offrent la possibilité d'un avancement de la recherche.

Une espèce  *nominale* est une catégorie d'entités déterminée par des définitions, des règles, des conditions nécessaires et suffisantes ou plus généralement par des essences nominales. Une catégorie comme JADE est nominale : tous les échantillons de jade sont faits d'une pierre précieuse verdâtre mais cette catégorie n'est pas une espèce naturelle. Ce qu'on appelle « jade » est en fait une catégorie hétérogène composée de deux types de minéraux, la jadéite  $[\text{Na}(\text{Al}, \text{Fe})\text{Si}_2\text{O}_6]$  et la néphrite  $[\text{Ca}_2(\text{Mg}, \text{Fe})_5\text{Si}_8\text{O}_{22}(\text{OH})]$ . Bien que la catégorie puisse être décrite par

des conditions nécessaires et suffisantes, elle constitue un ensemble hétéroclite. En général, outre les espèces mathématiques, les espèces nominales ne sont pas scientifiquement pertinentes.

Les catégories scientifiquement pertinentes sont en général des catégories d'entités dont l'homogénéité ou la ressemblance ne provient pas d'une définition, mais d'une nature commune considérée comme étant objective; elles sont « formé[e]s par la nature, non par de simples définitions » (Whewell [1860] 1971, 290). Ces catégories justifient des inductions en vertu de la possession d'un ensemble de propriétés unies par des relations empiriques (causalement ou par corrélation) plutôt que logiques.

Les espèces naturelles comportent des propriétés de surfaces et une structure sous-jacente, habituellement identifiée par des découvertes scientifiques. Les propriétés de surface des membres de la classe sont expliquées par une structure sous-jacente qu'une théorie aura mis au jour. L'eau, par exemple, possède certaines propriétés de surface par lesquelles nous l'identifions : sa masse volumique, son point d'ébullition, sa liquidité à température normale, etc. Ces propriétés sont expliquées par le fait que tout échantillon d'eau possède la même structure moléculaire, à savoir  $H_2O$ . La structure sous-jacente partagée par chaque membre de la catégorie EAU fait en sorte qu'un échantillon de liquide doté de cette structure sera de l'eau. Si en revanche, un liquide ne possède pas cette structure, il ne peut être considéré comme de l'eau, même s'il possède les mêmes propriétés superficielles.

L'identification d'une espèce naturelle circonscrit des corrélations de propriétés « projectibles » (Goodman, 1984[1954]), c'est-à-dire qu'elles permettent de faire des prédictions: ce qui sera découvert avec un échantillon d'eau E1 pourra être généralisé à l'ensemble des membres de la classe, et on est justifié de croire que le prochain échantillon d'eau E2 se comportera comme E1. La ressemblance superficielle d'échantillons de jade, en revanche, ne

s'explique pas par une structure sous-jacente commune et donc on ne peut justifier des inductions à propos des échantillons de jade.

De la sorte, toute enquête sérieuse—et *a fortiori* les programmes de recherches naturalistes en philosophie—aura avantage à identifier des espèces naturelles. Le problème des espèces nominales est qu'elles risquent d'être identifiées par des conditions qui ne sont, en fait, ni nécessaires ni suffisantes pour individuer une espèce, comme Kripke et Putnam l'ont souligné<sup>10</sup>. On peut très bien faire référence à X tout en se trompant à propos de X, il demeure qu'on fait toujours référence à la même substance. Lorsqu'on a découvert que les baleines sont des mammifères et non des poissons, l'énoncé « les baleines sont des poissons » n'était plus vrai, mais « baleine » et BALEINE réfèrent toujours aux baleines. De même, si sur une autre planète se trouve un liquide doté des mêmes propriétés de surface que l'eau mais dont la structure moléculaire est différente, alors ce liquide n'est pas de l'eau, mais quelque chose d'autre.

La connaissance avance donc par la mise au jour des espèces naturelles; si en plus elles peuvent être décrites par des conditions nécessaires et suffisantes, tant mieux; la rectitude définitionnelle n'est cependant pas la première vertu des concepts scientifiques, car on lui préfère souvent le potentiel inductif et explicatif. GÈNE est une espèce naturelle bien qu'on soit loin d'avoir une définition en termes de conditions nécessaires et suffisantes (Falk, 2000) : cette catégorie est toutefois pertinente, parce qu'elle guide des recherches en biologie.

On peut distinguer parmi les espèces naturelle des espèces qui, suivant l'usage, sont dites *historiques*<sup>11</sup>. Les espèces historiques qui concernent particulièrement la problématique de la naturalisation de la rationalité sont celles de la biologie, à savoir les lignées<sup>12</sup> animales et végétales. Ce sont des catégories homogènes qui supportent des inductions en vertu de la possession d'un ensemble de propriétés scientifiquement pertinentes. Si on sait qu'un cactus *Ferocactus pilosus* peut supporter au plus un froid de  $-4^{\circ}\text{C}$ , alors, *ceteris paribus*, on sait aussi

qu'un autre cactus de la même lignée tolérera le froid jusqu'à cette température. Les lignées sont des espèces naturelles historiques plutôt qu'éternelles (comme l'eau ou les électrons). Un animal ou une plante appartient à une espèce non pas selon son adéquation à un type abstrait ou à un ensemble de propriétés essentielles, mais selon sa généalogie et son appartenance à une population issue de la même généalogie<sup>13</sup>.

Pour qualifier de manière générale les espèces naturelles historiques, la philosophie des sciences contemporaine a souvent recours à la définition de Boyd (1991, 1999) : une espèce naturelle historique est un agrégat homéostatique de propriétés (*homeostatic property cluster*, ci-après). Elles comportent deux caractéristiques : 1) ce sont des espèces naturelles (des agrégats de propriétés co-occurentes qui supportent l'induction en spécifiant des structures sous-jacentes) et 2) ces espèces sont unies par des structures sous-jacentes d'un type particulier, à savoir des mécanismes qui font en sorte que ces propriétés tendent à survenir chez les membres de la classe (1999, 70). L'unité et l'homogénéité d'une lignée ne sont pas explicités par des conditions nécessaires et suffisantes, mais par des *mécanismes* (la reproduction, le développement, la sélection naturelle, les mutations, la niche écologique, etc.) qui expliquent la persistance et l'identité qualitative<sup>14</sup> de l'espèce. Prenons par exemple le concept intuitif, pré-darwinien, de POISSON\* : celui-ci regroupe les différentes lignées qui sont encore aujourd'hui appelées « poissons »<sup>15</sup> en plus des baleines et des dauphins, classées aujourd'hui comme des mammifères. Pour différencier les poissons des mammifères et réviser l'ancien concept POISSON\*, on doit reconstruire une histoire évolutionniste qui montre la spéciation des différentes lignées, leurs adaptations analogues mais séparées ainsi que les différences développementales (les uns ont des branchies, les autres des poumons, etc.). En bref, la sélection naturelle, les interactions écologiques et le développement sont des mécanismes qui, chez les cétacés et les poissons, maintiennent deux agrégats de propriétés qui supportent deux classes distinctes d'inductions.

Outre les lignées, l'ontologie de la biologie comporte un autre type d'espèce naturelle historique : les traits homologue ou *homologies*. La biologie évolutionniste et la cladistique, depuis les travaux fondateurs de Hennig (Hennig, 1966), classent les espèces par ascendance commune, plutôt que par une hiérarchie linnéenne des espèces, groupes, genres, etc. L'analyse cladistique suppose qu'on puisse analyser un organisme comme un ensemble de traits, et que les liens généalogiques entre ces traits permettent de reconstruire la phylogénie des espèces qui portent ces traits. Lorsqu'un trait est individué, son individuation est en fait une hypothèse à l'effet qu'il soit une homologie. Dans son acceptation classique, l'homologie est l'identité de traits dus à une ascendance commune :

- Homologie =<sub>df</sub> les traits A et B des espèces X et Y sont homologues ssi X et Y sont des descendants de l'espèce  $\alpha$ , le dernier ancêtre commun de X et Y doté de T, la forme ancestrale de A et B

Ce concept « taxique » (fondé sur les taxonomies) d'homologie tend maintenant à être complété (ou pour certains, remplacé) par un concept développemental d'homologie<sup>16</sup> :

- Homologie<sub>dev</sub> =<sub>df</sub> les traits A et B des espèces X et Y sont homologues ssi A et B sont des instances du même phénomène développemental au même niveau d'analyse.

Par *phénomène développemental*, il faut entendre la maturation ou encore la trajectoire développementale d'un trait déterminée par les mêmes contraintes développementales. Le code génétique, la morphologie, l'environnement et d'autres facteurs constituent des contraintes développementales qui contribuent ensemble à la construction du trait.

De la sorte, l'individuation des traits en biologie comparative ne procède pas par la mise au jour de caractéristique fonctionnelles, mais plutôt en identifiant l'ascendance commune ou les mécanismes de développement communs (l'homologie), une pratique courante de la biologie. Ainsi, « aile » est une catégorie fonctionnelle, mais « membre pentadactyle<sup>17</sup> » ne l'est pas. La catégorie fonctionnelle des ailes, identifiée comme organes dont la fonction est le vol, est à la biologie ce que le jade est à la géologie: une catégorie hétérogène qui ne supporte pas d'inductions sur la nature du trait, car elle regroupe des organes non homologues et exclus des organes homologues (les ailes des autruches ne servent pas à voler mais sont pourtant des ailes). La similarité de fonctions sera peu informative sur la catégorie qu'elle individue.<sup>18</sup> Inversement, des traits homologues ne garantissent pas une identité de fonctions, particulièrement pour des traits qui ont divergé depuis longtemps. Le même membre pentadactyle peut avoir différentes fonctions : la préhension, la marche, la course, le vol, la nage, etc.

Aussi identifier des traits, ou même généralement des « catégories biologiques »<sup>19</sup> *uniquement* par leurs fonctions, de façon homoplasique, ne nous donne donc que des essences nominales. En revanche, des traits ou catégories définies par l'ascendance commune—comme les lignées et homologies—même s'ils ne partagent pas la même fonction, sont assez homogènes pour supporter les inductions : on peut s'attendre à trouver par exemple la même ossature chez le pingouin et le flamand rose. Dans certains cas, on peut aussi, grâce aux développements de la génétique moléculaire et la biologie développementale, isoler des contraintes développementales ou des structures moléculaires homologues. Le meilleur exemple est sans doute les gènes *hox*, responsables de la segmentation du corps des animaux selon un axe antérieur/postérieur.<sup>20</sup> Ces gènes sont des homologues partagés par une grande partie des animaux; les interactions entre ces gènes et d'autres gènes, elles, ne le sont pas, ce qui explique la grande diversité des plans corporels malgré l'homologie de ces gènes. La relation entre le génotype et le phénotype étant

passablement complexe, il est plutôt rare voire impossible d'identifier un seul gène ou un seul ensemble de gènes à un trait en particulier : les gènes peuvent être homologues entre eux, sans que les structures anatomiques le soient.

Il est alors justifié de dire que l'homogénéité des homologues en font, avec les lignées, les espèces naturelles de la biologie<sup>21</sup>: chaque classe d'équivalence est assez homogène pour être considérée comme une espèce naturelle, et lorsqu'on possède des connaissances sur un membre de la classe, on peut généraliser à l'ensemble de la classe. Ce sont aussi des agrégats homéostatiques de propriétés : les homologues se maintiennent à travers les générations par la sélection naturelle, la reproduction et le développement.

La raison principale pour préférer l'analyse homologique (outre l'adhérence à l'ontologie de la biologie comparative) est la « profondeur » des similarités entre les membres de la classe. Plus on analyse les ressemblances de surface entre des membres de classes uniquement unies par une relation homoplasique, moins on trouve de principes sous-jacents. À l'inverse, plus on analyse les ressemblances de surfaces entre des homologues, plus on découvre des principes unificateurs qui concourent à faire avancer le savoir. De plus, le concept d'homologie est relativement indépendant de la conception théorique de l'évolution et de l'adaptation : sa formulation est sensiblement la même depuis Owen<sup>22</sup>, et fut acceptée en biologie avant *L'Origine des espèces*. Darwin a vu dans la sélection naturelle une façon d'expliquer l'homologie. Pour reprendre une boutade de Griffiths, Richard Dawkins et le Pape n'ont sans doute pas la même opinion sur la fonction propre de la sexualité, mais ils seront du même avis quant à la conception des parties génitales.

De la sorte, si l'on veut spécifier l'étendue réelle de la catégorie AGENT RATIONNEL dans le monde vivant, il faut recourir à une méthode qui ne crée pas de classes nominales, mais qui identifie plutôt des espèces naturelles historiques, soit des lignées et des homologues. Plutôt

qu'une méthode interprétive, il faut recourir à une méthode *architecturale*. Pour les tenants de cette méthode<sup>23</sup>, le point de départ d'une analyse de la rationalité est le vivant et son évolution plutôt que l'humain. On recherche les conditions d'émergences (pressions sélectives, contraintes développementales et morphologiques) des comportements rationnels et les mécanismes qui en sont responsables.

Comme le type de rationalité étudié ici est celui de l'action, les propriétés de surfaces de cette catégorie, si elle existe, concerneront le comportement et la prise de décision, ce qu'on peut appeler les *performances* des agents. Performer une action rationnelle n'est toutefois pas un gage de rationalité, tout comme la production d'une phrase par une machine ou un animal n'est pas une assurance que l'animal ou la machine soit un être linguistique. Il faut montrer que la performance est sous-tendue par une *compétence*. La rationalité des agents tient à leur capacité à générer des actions qui peuvent être, mais qui ne le sont pas nécessairement, rationnelles. L'agentivité rationnelle est de l'ordre de la disposition ou de la propension, ce qu'on pourrait appeler la « compétence rationnelle »<sup>24</sup>. Une compétence est une disposition à produire des actions rationnelles, en dehors des diverses erreurs de performances. Comme la compétence grammaticale (Chomsky, 1965), la compétence rationnelle peut faillir lors de différentes performances, mais qu'une personne fasse une faute de grammaire n'implique pas qu'elle n'a pas de compétence grammaticale; de même, poser une action irrationnelle n'entraîne pas que l'agent ne soit pas rationnel. À l'inverse, un agent sans compétence rationnelle ne peut pas poser d'action rationnelle ou irrationnelle. La compétence rationnelle a un rôle particulièrement important pour l'extension de la catégorie des agents rationnels : elle constitue la structure sous-jacente qui unit cette catégorie, car c'est elle qui fait en sorte que les propriétés de la catégorie (les performances) surviennent. Si les agents rationnels constituent une espèce naturelle, c'est parce qu'ils ont cette compétence., tout comme les individus linguistiques possèdent tous une compétence linguistique

et constituent ainsi une espèce naturelle. Comme le dit Proust, ce qu'on cherche à comprendre, en suivant une méthode architecturale, ou plus généralement une méthode naturaliste et réaliste,

*ce n'est pas comment la projection de la rationalité humaine nous aide à reconstruire ce que cela fait [what it is like] d'être un agent non-humain. Il s'agit plutôt [de comprendre] jusqu'où dans la phylogénie le concept de rationalité peut être adéquatement appliqué (Proust, 2006 op.cit.).*

S'il s'avère que les agents rationnels sont une espèce naturelle, on doit pouvoir différencier les propriétés de surface (performance) de la structure sous-jacente (compétence). Et si cette espèce naturelle est adéquatement inscrite dans l'ontologie de la biologie, alors son étendue déterminera un clade monophylétique, dont les performances rationnelles s'expliqueront par l'activité d'une compétence rationnelle, laquelle sera réalisée par des structures neuronales homologues. Les sections suivantes suggèrent qu'on peut adéquatement appliquer le concept de rationalité à l'ensemble du clade vertébré, que l'écologie du comportement décrit les performances rationnelles des vertébrés et que les neurosciences cognitives ont identifié des structures neuronales homologues responsables de la compétence rationnelle.

#### **4 La performance rationnelle en écologie du comportement**

On peut admettre que des animaux prennent des décisions, mais ces décisions, pour être rationnelles, doivent littéralement maximiser une fonction d'utilité : autrement dit, la description de l'animal comme maximisateur ne doit pas être qu'un artefact de l'interprétation du comportement, mais une propriété réelle du comportement et des mécanismes qui en sont responsables.

La maximisation de l'utilité est une des hypothèses de base de l'écologie comportementale<sup>25</sup> (ou évolutionniste), la branche de la biologie qui étudie le comportement adaptatif des animaux. Suite aux pressions évolutionnistes, on peut s'attendre à ce que la sélection naturelle crée des individus qui tendent à être performants dans la survie et la reproduction : dans la recherche de nourriture, de partenaires, dans l'évitement de prédateurs, la sélection d'habitat, etc., en bref, de meilleurs décideurs. Les fondateurs de l'écologie du comportement, MacArthur et Pianka (1966), proposaient d'étudier la recherche de ressources nutritives, ou fourrageage (*foraging*) par les animaux en se demandant :

*dans quelles zones une espèce se nourrirait et quels items  
formeraient sa diète si l'espèce agissait de façon entièrement économique  
(the most economic fashion) (MacArthur & Pianka, 1966, p. 603)*

Ainsi, les animaux sont modélisés comme étant des fourrageurs optimaux (Stephens & Krebs, 1986). À l'aide de la théorie de l'optimisation et des connaissances sur l'écologie d'une espèce, on essaie d'identifier les objectifs et contraintes de l'animal dans sa quête de nourriture. On en déduit un modèle algorithmique, soit un ensemble de règles de décisions qui maximiseraient l'efficacité de l'animal dans sa tâche. Ces règles permettent alors de prédire le comportement, lequel peut valider ou invalider la règle postulée.

Un modèle de fourrageage incorpore donc des agents, leurs choix possibles, la devise à maximiser—la plupart du temps le gain calorique net—et un ensemble de contraintes intrinsèques ou extrinsèques à l'agent. Les décisions étudiées sont généralement le choix d'une zone de fourrageage (*patch*), le choix d'une proie ou le temps optimal alloué à l'exploration d'une zone. L'animal à la recherche de nutriments doit maximiser le nombre de calories ingérées, compte tenu de celles qu'il dépense à rechercher et à capturer sa proie; à ce problème s'ajoutent

aussi la fréquence de rencontre de la proie, le temps consacré à la recherche et les calories que peuvent lui apporter sa proie, etc. On peut alors mettre en équation le taux de rendement énergétique pour les différents types de proie et, à partir de ces équations et à l'aide de techniques comme la programmation dynamique, on dérive des algorithmes qu'un fourrageur optimal devrait implémenter pour optimiser la fonction objectif (lesquelles servent ensuite à la prédiction du comportement).

Les études des performances des fourrageurs en milieu naturel démontrent les vertus prédictives et explicatives de l'hypothèse maximisatrice. Elle prédit par exemple l'arbitrage entre la défense du territoire et le fourragement chez des étourneaux (Kacelnik *et al.*, 1981) et l'arbitrage entre le chant et le fourragement chez les mésanges charbonnières (Thomas, 1999), explique les différences comportementales entre le fourragement chez des oiseaux vivant dans le désert (risque de prédation plus élevé et difficultés d'accès aux ressources) et les oiseaux urbains (Shochat *et al.*, 2004), la modulation, chez les rats, du temps alloué à l'ingestion de nourriture en fonction de la luminosité et de leur visibilité (ils mangeront plus vite dans un environnement lumineux où ils peuvent être facilement visibles) (Whishaw *et al.*, 1992) ou encore pourquoi des oiseaux huîtres préfèrent des huîtres à coquilles brunes plutôt que noires, alors que les deux sont identiques en énergie et en épaisseur : les brunes contiennent moins d'humidité, et donc en les sélectionnant les huîtres maximisent leur gain d'énergie à long terme (l'humidité fait en sorte que les huîtres occupent plus de place dans l'estomac de l'oiseau) (Nagarajan *et al.*, 2002).

Dans leur méta-analyse, Sih et Christensen (Sih & Christensen, 2001) ont revu 134 études (en laboratoire et en contexte naturel, expérimentales ou observationnelles) de fourragement, et en concluent que, bien que le succès prédictif ne soit pas parfait, la théorie est efficace selon le type de proie que recherche le fourrageur. Lorsque la proie est en mouvement, la prédictivité chute à 37%; en revanche, lorsque la proie (la *proie* peut être une plante, des graines, du miel,

etc.) est immobile, près de 74% des études réussissent à prédire le comportement. On peut cependant expliquer cette différence importante par l'ordre de rationalité impliqué : lorsque la proie est mobile, l'individu est en situation de décision *stratégique* plutôt qu'individuelle, et donc la théorie des jeux<sup>26</sup> est plus appropriée que la théorie du fourrage optimal, qui est en fait une théorie de la décision séquentielle.

Le succès d'un prédateur à la recherche de proies mobiles ne dépend donc pas que de ses décisions, mais aussi de celles de sa proie. De même, lorsque le fourrageur est à la recherche d'une proie, il peut lui aussi être la proie d'un autre fourrageur, ou en compétition avec ses conspécifiques. Un modèle de Vickery *et al.* (Vickery *et al.*, 1991) prédit ainsi que la co-occurrence de trois stratégies de fourrage en contexte social, PRODUCTEUR (qui amasse des nutriments) CHAPARDEUR (qui vole le butin du premier) et OPPORTUNISTE (producteur et charpateur en alternance) survient rarement (seulement dans le cas peu probable où la perte que l'opportuniste encourrait en cherchant est exactement équivalente à son gain s'il charpait ses conspécifiques) bien que des distributions de paires de stratégies sont possibles : elles constituent des stratégies évolutivement stables (SES). La proportion de zones partagées par les charpateurs, la taille du groupe et le degré de compatibilité entre l'activité charpateuse et productive (par exemple s'il est facile pour l'animal de performer les deux activités) déterminent la répartition des stratégies dans une population, ce qui a été par la suite confirmé entre autres chez des oiseaux (*Conchura punctulata*, ou capucins damiers) (Mottley & Giraldeau, 2000). Tel que prédit par le modèle, les capucins réduisent l'utilisation de la stratégie PRODUCTEUR lorsque le coût du fourrage individuel augmente.

La rationalité stratégique trouve aussi son application en écologie du comportement dans la théorie des marchés biologiques<sup>27</sup>. Évidemment, les marchés animaux n'ont pas de monnaies symboliques et conventionnelles, mais de nombreuses interactions entre animaux instituent des

structures d'échanges. Dès qu'une population d'échangeurs est en mesure de fournir des commodités en vue d'un profit mutuel, la compétition pour l'obtention de la commodité crée une surenchère. Les animaux qui maximisent leur utilité recherchent et sélectionnent des partenaires selon le principe de l'offre et la demande, qui s'applique à des domaines aussi variés que le mutualisme interspécifique<sup>28</sup>, la sélection sexuelle ou la coopération intraspécifique. Pour les babouins *Papio cynocephalus ursinus*, pouvoir s'amuser avec des nourrissons constitue une commodité : les adultes cherchent à y avoir accès en échangeant la compagnie des jeunes en échange d'épouillage. Plus les enfants sont rares, plus les femelles doivent épouiller pour y avoir accès (Henzi & Barrett, 2002) et inversement. De même, la femelle de la souris *Apodemus sylvaticus* échange du toilettage contre de l'information sur sa fertilité : plus une femelle souris est disponible pour la reproduction, plus elle exige de toilettage en échange (Stopka & Macdonald, 1999). Finalement, dans le domaine de la coopération intraspécifique, les poissons *Hipposcarus harid* et les poissons-nettoyeurs *Labroides dimidiatus* instituent un marché du nettoyage : les « clients » utilisent les services des nettoyeurs pour se faire enlever des parasites qui les importunent, alors que les nettoyeurs, à l'occasion, trichent et se nourrissent des tissus sains des clients. Comme les nettoyeurs offre un service qui ne peut être trouvé ailleurs, ces derniers bénéficient d'un certain avantage économique. Un client n'a pas le choix d'être exploité ou non, alors que le nettoyeur choisit de coopérer ou non. Les clients, des prédateurs qui pourraient manger les nettoyeurs, s'abstiennent de le faire dans la grande majorité des cas, étant donné l'avantage réciproque. Bshary et Schaffer (Bshary & Schaffer, 2002) ont aussi observé que les nettoyeurs soignent plus les clients de passage que les réguliers et sont en compétition entre eux pour les clients occasionnels, plus faciles à exploiter.

L'écologie du comportement abonde d'études qui, plutôt que de considérer les animaux non humains uniquement comme des unités monadiques d'une population, en font des agents qui

prennent des décisions, interagissent et négocient, en prédisant et expliquant leur comportement avec les mêmes outils théoriques que la microéconomie.

On pourrait bien sûr reformuler chacune de ces expériences, et mettre les mots *décision* ou *échange* entre guillemets pour signifier que l'emploi du terme est uniquement instrumental, que ce ne sont que des façons de parler et qu'au fond il n'y pas vraiment de décisions et d'échanges, mais plutôt des comportement instinctifs préservés par la sélection naturelle. Si tel est le cas, il faudra alors mettre des guillemets lorsqu'on parlera de l'être humain : l'écologie du comportement humain, appliquant la même logique économique et biologique, en arrive à des prédictions aussi robustes que l'écologie du comportement animal non-humain<sup>29</sup>. On modélise les agents comme des fourrageurs optimaux sous l'effet d'une multitude de contraintes, on identifie les ressources disponibles dans l'environnement d'une collectivité, et on génère un modèle qui prédit l'allocation optimale de ressources dans la collecte de nourriture. Les modèles sont bien sûr plus complexes en ce qu'ils intègrent des paramètres sociaux comme les coutumes, la technologie ou les structures économiques locales, mais dans l'ensemble, l'écologie du comportement humain s'est révélée plus prédictive que la microéconomie classique. En incluant l'ancienneté du lieu d'établissement, la distance à parcourir et la technologie utilisée par 3 tribus d'Amazonie, un modèle de fourragement optimal (Hames & Vickers, 1982) prédit la différence dans le style de fourragement. D'autres modèles incorporent des relations sociales comme le partage de nourriture (Kaplan *et al.*, 1984) ou encore la division du travail entre hommes et femmes (Jochim, 1988), les cultures agricoles jusqu'à la recherche d'information sur Internet (Pirolli & Card, 1999), où la denrée recherchée est de l'information.

## 5 La compétence rationnelle animale : mécanismes et neurocladistique

L'écologie du comportement décrit les performances rationnelles des animaux, et montre qu'une description des animaux comme des agents économiquement rationnels est prédictivement adéquate. S'il y a une réelle unité chez les animaux rationnels, on doit pouvoir identifier la structure sous-jacente, à savoir la compétence rationnelle. Les biologistes du comportement animal ont cependant, règle générale, une interprétation instrumentaliste des modèles d'optimisation : ils ne les tiennent pas pour des énoncés sur des mécanismes internes de contrôle du comportement et ne se prononce pas quant à la capacité réelle (ou non) des animaux à effectuer des décisions. Je propose cependant, dans la présente section, de faire un pas de plus vers une théorie réaliste, naturaliste et architecturale de la rationalité, en proposant un mécanisme neuronal de prise de décision.

Les neurosciences du comportement reconnaissent que pour la sélection de l'action et la maximisation de l'utilité, une foule de zones et de processus sont sollicités. On peut cependant identifier certains composants actifs dans la décision maximisatrice d'utilité. De nombreux travaux montrent que parmi ceux-ci, les systèmes dopaminergiques sont particulièrement efficaces pour guider le comportement et pour l'apprentissage à base de rétroaction évaluative<sup>30</sup>.

L'évaluation de l'utilité est effectuée, dans le cerveau vertébré<sup>31</sup>, en grande partie par les neurones dopaminergiques des régions mésencéphaliques (en particulier l'aire tegmentale ventrale), la région du cerveau reliant le tronc cérébral à l'encéphale. La dopamine agit comme un neurotransmetteur, mais aussi et surtout comme un neuromodulateur : elle ne fait pas qu'activer des neurones-cible, mais vient aussi moduler la réponse de ces neurones.

Les systèmes dopaminergiques dirigent le maintien de buts en mémoire de travail, l'intégration, l'information sur les prédicteurs d'utilité détectée, la modulation de l'apprentissage,

la détection des erreurs de prédictions et le contrôle moteur. On observe ainsi une activation de ces neurones lors de stimuli prédictifs de récompenses (Schultz & Dickinson, 2000) : des stimuli liés aux récompenses de façon primaire (jus, nourriture), des stimuli auditifs et visuels qui indiquent la possibilité de récompenses, des stimuli semblables aux récompenses, ou encore des stimuli qui élicitent des réactions d'orientation et des comportements d'approches, comme des stimuli qui recrutent l'attention, (nouveaux ou particulièrement intenses), potentiellement liés à des récompenses (la saillance motivationnelle). Montague *et al* comparent son rôle à celui d'un portier : les neurones dopaminergiques filtrent l'information qui doit être signalée en mémoire de travail, et ne laissent passer que celle qui représente un intérêt pour le système. Les représentations en mémoire de travail peuvent alors être mises à jour, modifiées ou être utilisées pour le contrôle du comportement (P.R. Montague *et al.*, 2004).

Globalement, les neurones dopaminergiques démontrent une sensibilité aux qualités évaluatives de l'environnement. Ils « étiquettent » les stimuli environnementaux d'une valeur appétitive, prédisent les récompenses et signalent des événements motivants et importants (alertes). En plus d'un rôle évaluatif, les recherches de Berridge et de ses collègues montrent que les systèmes dopaminergiques sont en grande partie responsables de la motivation<sup>32</sup>. L'activité de ces neurones est nécessaire pour que des indices reliés à des récompenses acquièrent une signification et une saillance motivante. Des souris génétiquement modifiées, qui possèdent des récepteurs dopaminergiques mais n'en produisent pas, sont hypoactives et hypophagiques: la détection de nourriture ne motive pas l'ingestion. Lorsqu'on administre un produit qui, une fois métabolisé, est converti en dopamine, les souris deviennent alors hyperactives et hyperphagiques et développent des stéréotypies<sup>33</sup>. On sait aussi que ces systèmes de motivation participent au contrôle de l'animal en étant dissociés de l'affect hédonique. Des rats hypodopaminergiques ont une réaction orofaciale au sucrose différente de la quinine, démontrant un affect hédonique pour

le premier mais non par le second (Berridge & Robinson, 1998). Ils n'auront toutefois aucune motivation à ingérer du sucrose ou tout autre nourriture auquel l'animal est sensible normalement. Le même phénomène a été constaté chez des humains, par exemple des fumeurs à qui on administre un antagoniste de la dopamine éprouvent le même plaisir à fumer sans avoir de désir (Berridge & Robinson, 2003). En situation de fourragement, ces animaux hypo- ou hyperdopaminergiques auraient peine à recueillir les ressources énergétiques nécessaires : les uns par manque de motivation, les autres par manque de concentration. Lorsqu'un antagoniste de la dopamine est injecté dans le nucleus accumbens, des rats se sont montrés incapables d'explorer un environnement à la recherche de nourriture (Floresco *et al.*, 1996). Si l'environnement est déjà connu, en revanche, ils sont capables de réutiliser de l'information déjà acquise à son sujet et de retrouver des sources de nutriments; de la sorte, les systèmes dopaminergiques n'interviennent dans le fourragement que lorsqu'il y a ambiguïté quant à localisation de la nourriture, bref lorsqu'une décision s'impose.

On a aussi observé que les neurones dopaminergiques ne répondent pas directement à la présence de récompenses, mais uniquement lorsqu'une différence est constatée entre un état attendu et l'état occurrent. Les enregistrements d'activités de neurones individuels (Schultz *et al.*, 1997) montrent que les neurones dopaminergiques envoient un signal d'erreur—seulement lorsque la valeur de la récompense attendue, et la valeur de la récompense occurrente, est non nulle—analogue à celui des algorithmes DT (différence temporelle), un type d'apprentissage non supervisé qui opère à partir des erreurs de prédiction. Autrement dit, si les neurones sont « surpris » (une récompense survient alors qu'elle n'était pas attendue), ou si ils sont « déçus » (la récompense attendue ne s'est pas présentée), les neurones renvoient un signal d'erreur. La réponse dopaminergique fonctionne comme un prédicteur d'utilité<sup>34</sup>.

La modélisation computationnelle du rôle des systèmes dopaminergiques montre que ceux-ci sont bien décrits par des algorithmes d'apprentissage DT. Ces algorithmes, utilisés dans des réseaux de neurones artificiels ou en robotique autonome, apprennent et contrôlent un système en effectuant de façon constante des prédictions quant à l'utilité attendue, et réajustent cette prédiction lorsque que le signal ne correspond pas à la prédiction. À force de corriger ses erreurs de prédiction, le système apprend des règles (*policies*) de contrôle pour l'atteinte de ses buts. Toute sa capacité d'apprentissage et de contrôle repose donc sur la possibilité qu'il soit dans l'erreur et qu'il reconnaisse son erreur, qu'il soit surpris.

Ces modèles d'apprentissage, en plus d'être biologiquement plausibles, sont tractables, obligés d'acquérir de l'information pour prendre des décisions, et nécessairement situés dans des environnements incertains avec lesquels ils sont en contact continu. On a par exemple soumis des sujets humains à une simple tâche de décision sous incertitude radicale: ils devaient choisir entre deux options A et B (deux icônes affichés sur un écran d'ordinateur) sans savoir la probabilité de paiement de chaque option. À chaque coup, l'ordinateur les informait de leur gain total. Le même problème a été présenté à un réseau de neurones artificiels munis d'un algorithme d'apprentissage DT, et les humains comme le réseau ont produit les mêmes résultats (Egelman *et al.*, 1998). La performance de singes dans des tâches similaires peut aussi être prédite par un modèle DT (Glimcher *et al.*, 2005).

Il ressort de toutes ces études que les systèmes dopaminergiques ne sont pas des systèmes perceptifs ou moteurs, mais plutôt des *contrôleurs* du comportement, de la cognition, de l'apprentissage et de l'attention qui manipulent les paramètres de structures neuronales impliqués dans ces fonctions (Nieoullon, 2002). Une architecture de contrôle dotée de ces capacités peut découvrir d'elle-même ce qui, dans son environnement, est apte à satisfaire ses besoins. Cette plasticité dans le comportement est plus efficace, en contexte incertain, qu'un système pré-cablé

qui spécifierait déjà les ressources à recueillir. La génétique peut toujours pré-spécifier les mécanismes de renforcement, mais la trajectoire d'atteinte des *satisfecits* est laissée au développement et à l'apprentissage. Le système est alors apte, en intégrant l'information sur l'utilité, la probabilité et l'incertitude, à guider son comportement d'optimisation dans la recherche de ressources.

Une caractéristique importante des systèmes dopaminergiques et de leur algorithmes DT, pour la question de la rationalité animale, est leur rôle dans la prise de décision dans l'ensemble du clade vertébré. En effet, des rats au humains, en passant par les primates et les oiseaux<sup>35</sup>, les processus de valuation et de prise de décision font systématiquement appel aux mécanismes dopaminergique :

*La réception des objets culturels tels que les œuvres d'art ne peuvent pas être considérés comme des récompenses primaires [primary reward], et pourtant ces objets suscitent l'activité des mêmes structures neuronales activées par la réception des récompenses primaires (P. R. Montague et al., 2006, p. 420),*

Ainsi, on a pu montrer par exemple que les préférences dûes aux effets de marque (« *branding* ») étaient aussi le résultat des systèmes dopaminergique. Dans une expérience, les sujets devaient évaluer une boisson gazeuse (Pepsi ou Coke). Sans la marque, les sujets américains ont tendance à préférer légèrement Pepsi. Avec la marque, cependant, les sujets ont tendance à affirmer une préférence pour le Coke, une préférence reflétée par une activation marquée des systèmes dopaminergiques (McClure *et al.*, 2004). On sait aussi que l'action de drogues psychomotrices telles que la cocaïne recrutent les neurones dopaminergiques (Kauer, 2004); l'incapacité à inhiber un comportement de consommation de drogue découle de la

surabondance de transmissions dopaminergiques (van Gaalen *et al.*), créée par ces substances, qui biaisent le système motivationnel. La dépendance à ces drogues peut être forte au point où le sujet éprouve une envie de consommer même s'il n'en retire aucun plaisir.

De la sorte, que ce soit des stimuli primaires ou secondaires, essentiels à la survie ou non, chargés d'une signification culturelle ou non, la machinerie d'évaluation et de sélection de l'action reste la même. Étant donné que le plan du cerveau vertébré est relativement uniforme et que les cerveaux actuels sont des homologues, la neurocladistique<sup>36</sup> est donc en mesure de justifier l'induction suivante : les systèmes dopaminergiques sont des mécanismes de prises de décision répandus dans l'ensemble du clade vertébré. Leurs opérations sont donc des composantes primordiales de la compétence rationnelle. Des études neurocladistiques ont par exemple permis de révéler la phylogénie des différents types de récepteurs dopaminergiques. Bien que ces récepteurs ne soient pas uniquement présents chez les vertébrés, on assiste toutefois dans ce clade, à une duplication de ces récepteurs, qui serait due à la demande accrue de traitement de l'information et de neuromodulation qu'exigent des processus comme la motivation, l'émotion, la planification et la mémoire<sup>37</sup>.

## **6 Conclusion**

De la sorte, plutôt que de se demander dans quelle mesure on peut adopter la posture intentionnelle face à des êtres d'une autre espèce ou dans quelle mesure notre discours psychologique intuitif s'applique aux animaux, il est aujourd'hui possible d'aborder la question de la rationalité animale sous un autre angle : dans quelle mesure peut-on identifier une espèce naturelle constituant l'extension de la catégorie AGENT RATIONNEL ? Pour ce faire, il faut identifier à la fois les propriétés de surface et les structures sous-jacentes qui expliquent lesdites propriétés, ce que j'ai appelé les performances et compétences rationnelles. On peut donc montrer,

grâce aux recherches en écologie du comportement, que les propriétés de surface des agents rationnels, à savoir la prise de décisions rationnelles, surviennent au moins dans tout ce clade. Ces performances rationnelles, elles, peuvent être expliquées par une compétence rationnelle, à savoir la possession de mécanismes de prises de décisions. La décomposition mécaniste de cette compétence fait intervenir des entités homologues—les systèmes dopaminergiques—dans tout le clade vertébré, ce qui justifie la généralisation de la compétence à tout le clade.

On pourrait bien sûr tenter de voir dans quelle mesure le concept s'étend à d'autres clades, car il ne serait pas étonnant de constater que ces propriétés soient partagées par d'autres clades, une espèce naturelle n'étant pas déterminée par des conditions nécessaires et suffisantes, mais plutôt par des agrégats de propriétés co-occurentes. Cependant, en maintenant l'extension au clade vertébré, on dispose alors d'une capacité inductive scientifiquement intéressante, qui permet de supporter des inférences d'une espèce à l'autre en vertu de la relative homogénéité du cerveau vertébré.

## 7 Références

- Allen, C., Bekoff, M., & Lauder, G. V. (1998). *Nature's Purposes : Analyses of Function and Design in Biology*. Cambridge, Mass.: The MIT Press.
- Bayer, H. M., & Glimcher, P. W. (2005). Midbrain Dopamine Neurons Encode a Quantitative Reward Prediction Error Signal. *Neuron*, 47(1), 129.
- Bennett, J. F. (1971). *Rationality. An Essay Towards an Analysis*. London,: Routledge and Kegan Paul.
- Bermúdez, J. L. (2003). *Thinking without Words*. Oxford ; New York: Oxford University Press.
- Berridge, K. C. (2004). Motivation Concepts in Behavioral Neuroscience. *Physiol Behav*, 81(2), 179-209.

- Berridge, K. C., & Robinson, T. E. (1998). What Is the Role of Dopamine in Reward: Hedonic Impact, Reward Learning, or Incentive Saliency? *Brain Res Brain Res Rev*, 28(3), 309-369.
- Berridge, K. C., & Robinson, T. E. (2003). Parsing Reward. *Trends Neurosci*, 26(9), 507-513.
- Boyd, R. (1999a). Homeostasis, Species, and Higher Taxa. In R. Wilson (Ed.), *Species: New Interdisciplinary Essays* (pp. 141-185). Cambridge: MIT Press.
- Boyd, R. (1999b). Kinds, Complexity and Multiple Realization. *Philosophical Studies*, 95(1 - 2), 67.
- Brigandt, I. (2002). Homology and the Origin of Correspondence. *Biology and Philosophy*, 17(3), 389.
- Bshary, R., & Schaffer, D. (2002). Choosy Reef Fish Select Cleaner Fish That Provide High-Quality Service. *Animal Behaviour*, 63(3), 557.
- Callier, S., Snapyan, M., Le Crom, S., Prou, D., Vincent, J. D., & Vernier, P. (2003). Evolution and Cell Biology of Dopamine Receptors in Vertebrates. *Biol Cell*, 95(7), 489-502.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Collier, J. (1996). On the Necessity of Natural Kinds. In P. Riggs (Ed.), *Natural Kinds, Laws of Nature, and Scientific Methodology* (pp. 1-10). Dordrecht: Kluwer.
- Davidson, D. (1993). *Actions Et Événements*. Paris: Presses universitaires de France.
- Dayan, P., & Balleine, B. W. (2002). Reward, Motivation, and Reinforcement Learning. *Neuron*, 36(2), 285-298.
- Dennett, D. C. (1987). *The Intentional Stance*. Cambridge, Mass.: MIT Press.
- Dennett, D. C. (1995). *Darwin's Dangerous Idea : Evolution and the Meanings of Life*. New York: Simon & Schuster.

- Dugatkin, L. A., & Reeve, H. K. (1998). *Game Theory & Animal Behavior*. New York ; Oxford: Oxford University Press.
- Egelman, D. M., Person, C., & Montague, P. R. (1998). A Computational Role for Dopamine Delivery in Human Decision-Making. *J Cogn Neurosci*, 10(5), 623-630.
- Falk, R. (2000). The Gene: A Concept in Tension. In P. Beurton, R. Falk & H.-J. Rheinberger (Eds.), *The Concept of the Gene in Development and Evolution*. Cambridge: Cambridge University Press.
- Floresco, S. B., Seamans, J. K., & Phillips, A. G. (1996). A Selective Role for Dopamine in the Nucleus Accumbens of the Rat in Random Foraging but Not Delayed Spatial Win-Shift-Based Foraging. *Behav Brain Res*, 80(1-2), 161-168.
- Ghiselin, M. T. (1974). A Radical Solution to the Species Problem. *Systematic Zoology*, 23, 536-544.
- Gilbert, S. F. (2006). *Developmental Biology* (8th ed.). Sunderland, Mass.: Sinauer Associates, Inc. Publishers.
- Giraldeau, L.-A., & Caraco, T. (2000). *Social Foraging Theory*. Princeton, N.J.: Princeton University Press.
- Glimcher, P. W., Dorris, M. C., & Bayer, H. M. (2005). Physiological Utility Theory and the Neuroeconomics of Choice. *Games Econ Behav*, 52(2), 213-256.
- Godfrey-Smith, P. (1996). *Complexity and the Function of Mind in Nature*. Cambridge ; New York: Cambridge University Press.
- Godfrey-Smith, P. (2003). Folk Psychology under Stress: Comments on Susan Hurley's 'Animal Action in the Space of Reasons'. *Mind & Language*, 18(3), 266-272.
- Goodman, N. (1984[1954]). *Faits, Fictions Et Prédications*. Paris: Minuit.

- Griffiths, P. E. (1994). Cladistic Classification & Functional Explanation. *Philosophy of Science* 61(2), 206-227.
- Griffiths, P. E. (1997). *What Emotions Really Are : The Problem of Psychological Categories*. Chicago, Ill.: University of Chicago Press.
- Griffiths, P. E. (1999). Squaring the Circle: Natural Kinds with Historical Essences. In R. Wilson (Ed.), *Species: New Interdisciplinary Studies*. Cambridge, MA: MIT Press.
- Griffiths, P. E. (2006, sous presse). Function, Homology and Character Individuation. *Philosophy of Science*.
- Hall, B. K. (Ed.). (1994). *Homology: The Hierarchical Basis of Comparative Biology*. San Diego: Academic Press.
- Hames, R. B., & Vickers, W. T. (1982). Optimal Diet Breadth Theory as a Model to Explain Variability in Amazonian Hunting. *American Ethnologist*, 9(2, Economic and Ecological Processes in Society and Culture), 358-378.
- Hennig, W. (1966). *Phylogenetic Systematics*. Urbana, IL: University of Illinois Press.
- Henzi, S. P., & Barrett, L. (2002). Infants as a Commodity in a Baboon Market. *Animal Behaviour*, 63(5), 915.
- Hull, D. L. (1978). A Matter of Individuality. *Philosophy of Science*, 45, 335-360.
- Hurley, S. (2003). Animal Action in the Space of Reasons. *Mind and Language*, 18(3), 231-257.
- Hurley, S., & Nudds, M. (2006). *Rational Animals ?* Oxford: Oxford University Press.
- Ingo, B. (2003). Homology in Comparative, Molecular, and Evolutionary Developmental Biology: The Radiation of a Concept. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, 299B(1), 9-17.
- Jochim, M. A. (1988). Optimal Foraging and the Division of Labor. *American Anthropologist*, 90(1), 130-136.

- Kacelnik, A., Houston, A. I., & Krebs, J. R. (1981). Optimal Foraging and Territorial Defence in the Great Tit (*Parus Major*). *Behavioral Ecology and Sociobiology*, 8(1), 35.
- Kagel, J. H., Battalio, R. C., & Green, L. (1995). *Economic Choice Theory : An Experimental Analysis of Animal Behavior*. Cambridge [England] ; New York: Cambridge University Press.
- Kaplan, H., Hill, K., Hawkes, K., & Hurtado, A. (1984). Food Sharing among Ache Hunter-Gatherers of Eastern Paraguay. *Current Anthropology*, 25(1), 113-115.
- Kapsimali, M., Le Crom, S., & Vernier, P. (2003). A Natural History of Vertebrate Dopamine Receptors. In A. Sidhu, M. Laruelle & P. Vernier (Eds.), *Dopamine Receptors and Transporters : Function, Imaging, and Clinical Implication* (2nd ed., pp. xviii, 741 , [748] of plates). New York: M. Dekker.
- Kauer, J. A. (2004). Learning Mechanisms in Addiction: Synaptic Plasticity in the Ventral Tegmental Area as a Result of Exposure to Drugs of Abuse. *Annual Review of Physiology*, 66(1), 447-475.
- Krebs, J. R., & Davies, N. B. (1997). *Behavioural Ecology : An Evolutionary Approach* (4th ed.). Oxford, England ; Malden, MA: Blackwell Science.
- Kripke, S. (1972). *Naming and Necessity*. Cambridge, MA: Harvard University Press.
- Laubichler, M. D. (2000). Homology in Development and the Development of the Homology Concept. *Integr. Comp. Biol.* %R 10.1093/icb/40.5.777, 40(5), 777-788.
- Le Crom, S. p., Kapsimali, M., Bar, äöme, P.-O., & Vernier, P. (2003). Dopamine Receptors for Every Species: Gene Duplications and Functional Diversification in Craniates. *Journal of Structural and Functional Genomics*, 3(1 - 4), 161.
- Lecointre, G., & Le Guyader, H. (2001). *Classification Phylogénétique Du Vivant* (2. ed.). Paris: Belin.

- Locke, J. ([1689]2001). *Essai Sur L'entendement Humain* (J.-M. Vienne, Trans.). Paris: J. Vrin.
- MacArthur, R. H., & Pianka, E. R. (1966). On Optimal Use of a Patchy Environment. *American Naturalist*(100), 603-609.
- Matthen, M. (2000). What Is a Hand? What Is a Mind? *Revue Internationale de Philosophie*, 214, 653-673.
- McClure, S. M., Li, J., Tomlin, D., Cypert, K. S., Montague, L. M., & Montague, P. R. (2004). Neural Correlates of Behavioral Preference for Culturally Familiar Drinks. *Neuron*, 44(2), 379-387.
- McFarland, D. (1999). *Animal Behaviour : Psychobiology, Ethology, and Evolution* (3rd ed.). Harlow, Essex, England & New York: Wiley.
- Mill, J. S. (1848). *A System of Logic, Ratiocinative and Inductive; Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation*. New-York,: Harper & Brothers.
- Millikan, R. G. (1984). *Language, Thought, and Other Biological Categories : New Foundations for Realism*. Cambridge, Mass.: MIT Press.
- Millikan, R. G. (2005). *Language: A Biological Model*. Oxford: Oxford University Press.
- Montague, P. R., & Berns, G. S. (2002). Neural Economics and the Biological Substrates of Valuation. *Neuron*, 36(2), 265-284.
- Montague, P. R., Hyman, S. E., & Cohen, J. D. (2004). Computational Roles for Dopamine in Behavioural Control. *Nature*, 431(7010), 760.
- Montague, P. R., King-Casas, B., & Cohen, J. D. (2006). Imaging Valuation Models in Human Choice. *Annu Rev Neurosci*, 29, 417-448.
- Mottley, K., & Giraldeau, L. A. (2000). Experimental Evidence That Group Foragers Can Converge on Predicted Producer-Scrounger Equilibria. *Anim Behav*, 60(3), 341-350.

- Nagarajan, R., Goss-Custard, J. D., & Lea, S. E. (2002). Oystercatchers Use Colour Preference to Achieve Longer-Term Optimality. *Proc Biol Sci*, 269(1490), 523-528.
- Neander, K. (2002). Types of Traits: Function, Structure and Homology in the Classification of Traits. In A. Ariew, R. Cummins & M. Perlman (Eds.), *Functions: New Essays in the Philosophy of Biology and Psychology*. Oxford: Oxford University Press.
- Nieoullon, A. (2002). Dopamine and the Regulation of Cognition and Attention. *Progress in Neurobiology*, 67(1), 53.
- Nieuwenhuys, R., Donkelaar, H. J. t., & Nicholson, C. (1998). *The Central Nervous System of Vertebrates*. Berlin ; New York: Springer.
- Noë, R., & Hammerstein, P. (1995). Biological Markets. *10*(8), 336.
- Noë, R., Hooff, J. A. R. A. M. v., & Hammerstein, P. (2001). *Economics in Nature : Social Dilemmas, Mate Choice, and Biological Markets*. Cambridge, UK: New York, NY Cambridge University Press.
- Owen, R. (1843). *Lectures on the Comparative Anatomy and Physiology of the Invertebrate Animals, Delivered at the Royal College of Surgeons, in 1843*. London: Longman, Brown, Green, and Longmans.
- Pecina, S., Cagniard, B., Berridge, K. C., Aldridge, J. W., & Zhuang, X. (2003). Hyperdopaminergic Mutant Mice Have Higher "Wanting" But Not "Liking" For Sweet Rewards. *J Neurosci*, 23(28), 9395-9402.
- Penney, D., & McGee, G. (2005). Chemical Trust: Oxytocin Oxymoron? *Am J Bioeth*, 5(3), 1-2.
- Pirolli, P., & Card, S. (1999). Information Foraging. *Psychological Review*, 106(4), 643.
- Premack, D., & Woodruff, G. (1978). Does the Chimpanzee Have a Theory of Mind? *Behavioural and Brain Sciences*, 1, 515-526.
- Proust, J. (2003). *Les Animaux Pensent-Ils ?* Paris: Bayard.

- Proust, J. (2006). Metacognition and Animal Rationality. In S. Hurley & M. Nudds (Eds.), *Rational Animals*. Oxford: Oxford University Press.
- Putnam, H. (1975). The Meaning of Meaning In K. Gunderson (Ed.), *Language, Mind and Knowledge*. Minneapolis: University of Minnesota Press.
- Quine, W. V. O. (1969). *Ontological Relativity, and Other Essays*. New York,: Columbia University Press.
- Rieppel, O. (2005a). Monophyly, Paraphyly, and Natural Kinds. *Biology and Philosophy*, 20(2 - 3), 465.
- Rieppel, O. (2005b). Modules, Kinds, and Homology *J Exp Zool: Mol Dev Evol*, 304B, 18–27.
- Robinson, S., Sandstrom, S. M., Denenberg, V. H., & Palmiter, R. D. (2005). Distinguishing Whether Dopamine Regulates Liking, Wanting, and/or Learning About Rewards. *Behavioral Neuroscience*, 119(1), 5.
- Samuels, R., & Stich, S. (2004). Rationality and Psychology. In A. Mele & P. Rawling (Eds.), *The Oxford Handbook of Rationality* (pp. 279-300). Oxford: Oxford Reference Library.
- Schultz, W. (1999). The Reward Signal of Midbrain Dopamine Neurons. *News Physiol Sci*, 14(6), 249-255.
- Schultz, W. (2001). Reward Signaling by Dopamine Neurons. *Neuroscientist*, 7(4), 293-302.
- Schultz, W., Dayan, P., & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, 275(5306), 1593-1599.
- Schultz, W., & Dickinson, A. (2000). Neuronal Coding of Prediction Errors. *Annu Rev Neurosci*, 23, 473-500.
- Schwartz, M. W., & Hoeksema, J. D. (1998). Specialization and Resource Trade: Biological Markets as a Model of Mutualisms. *Ecology*, 79(3), 1029-1038.

- Shochat, E., Lerman, S. B., Katti, M., & Lewis, D. B. (2004). Linking Optimal Foraging Behavior to Bird Community Structure in an Urban-Desert Landscape: Field Experiments with Artificial Food Patches. *Am Nat*, *164*(2), 232-243.
- Sih, A., & Christensen, B. (2001). Optimal Diet Theory: When Does It Work, and When and Why Does It Fail? *Animal Behaviour*, *61*(2), 379.
- Smith, E. A. (1991). *Inujjamiut Foraging Strategies : Evolutionary Ecology of an Arctic Hunting Economy*. New York: A. de Gruyter.
- Stein, E. (1996). *Without Good Reason : The Rationality Debate in Philosophy and Cognitive Science*. Oxford: Oxford University Press.
- Stephens, D. W., & Krebs, J. R. (1986). *Foraging Theory*. Princeton, N.J.: Princeton University Press.
- Sterelny, K. (2003a). Charting Control Space: Comments on Susan Hurley's 'Animal Action in the Space of Reasons'. *Mind & Language*, *18*(3), 257-266.
- Sterelny, K. (2003b). *Thought in a Hostile World : The Evolution of Human Cognition*. Malden, MA ; Oxford: Blackwell Pub.
- Stopka, P., & Macdonald, D. W. (1999). The Market Effect in the Wood Mouse, *Apodemus Sylvaticus*: Selling Information on Reproductive Status. *Ethology*, *105*, 969-982.
- Thomas, R. J. (1999). Two Tests of a Stochastic Dynamic Programming Model of Daily Singing Routines in Birds. *Anim Behav*, *57*(2), 277-284.
- van Gaalen, M. M., van Koten, R., Schoffelmeer, A. N. M., & Vanderschuren, L. J. M. J. Critical Involvement of Dopaminergic Neurotransmission in Impulsive Decision Making. *Biological Psychiatry, In Press, Corrected Proof*.
- Vickery, W. L., Giraldeau, L.-A., Templeton, J. J., Kramer, D. L., & Chapman, C. A. (1991). Producers, Scroungers, and Group Foraging. *American Naturalist*, *137*(6), 847-863.

- Wagner, G. P. (1989). The Biological Homology Concept. *Annual Review of Ecology and Systematics*, 20(1), 51-69.
- Wagner, G. P. (1996). Homologues, Natural Kinds and the Evolution of Modularity. *American Zoologist*, 36, 36-43.
- Wagner, G. P. (Ed.). (2001). *The Character Concept in Evolutionary Biology*. San Diego: Academic Press.
- Whewell, W. ([1860] 1971). *On the Philosophy of Discovery*. New York: Burt Franklin.
- Whishaw, I. Q., Dringenberg, H. C., & Comery, T. A. (1992). Rats (*Rattus Norvegicus*) Modulate Eating Speed and Vigilance to Optimize Food Consumption: Effects of Cover, Circadian Rhythm, Food Deprivation, and Individual Differences. *J Comp Psychol*, 106(4), 411-419.
- Williams, G. V., & Goldman-Rakic, P. S. (1995). Modulation of Memory Fields by Dopamine D1 Receptors in Prefrontal Cortex. *Nature*, 376(6541), 572.
- Wilson, R. A. (1999). *Species : New Interdisciplinary Essays*. Cambridge, Mass.: MIT Press.
- Winterhalder, B., & Smith, E. A. (2000). Analyzing Adaptive Strategies: Human Behavioral Ecology at Twenty-Five. *Evolutionary Anthropology: Issues, News, and Reviews*, 9(2), 51-72.

## Notes

---

<sup>1</sup> (Glimcher *et al.*, 2005; Kagel *et al.*, 1995)

---

<sup>2</sup> Conformément à l'usage en sciences cognitives, nous utilisons les majuscules pour désigner un concept, et non le référent du concept ou son expression linguistique.

<sup>3</sup> Un trait ou caractère est une partie d'un animal ou un processus ayant lieu en l'animal : un organe, un os, un réflexe, un système de régulation, un mécanisme immunitaire, etc., sont des traits identifiables (Wagner, 2001). Les individus biologiques sont des collections de traits.

<sup>44</sup> Voir par exemple (Bennett, 1971; Bermúdez, 2003; Dennett, 1995; Hurley, 2003) de même que les essais dans (Hurley & Nudds, 2006).

<sup>5</sup> Dans les faits, les primatologues reconnaissent des compétences métacognitives aux primates, mais pas de « théorie de l'esprit » (Proust, 2003).

<sup>6</sup> Par exemple (Millikan, 1984; Neander, 2002) et les essais dans (Allen *et al.*, 1998).

<sup>7</sup> La cladistique est la classification en graphe acyclique du vivant en fonction de l'histoire phylogénétique (Lecointre & Le Guyader, 2001); un clade, ou taxon monophylétique, est un groupe d'organismes qui inclut le dernier ancêtre commun de tous les membres du groupe ainsi que les descendants de cet ancêtre commun.

<sup>8</sup> Cf. (Hall, 1994). L'analogie est une forme d'homoplasie, soit des identités de fonctions, mais le concept d'analogie ajoute à cette identité une propriété adaptative, alors que l'homoplasie se limite à l'identité.

<sup>9</sup> La notion d'espèces et d'essences naturelles ou réelles, contrastées avec les espèces et essences nominales, a une longue histoire en philosophie, continuellement associée à la question de

---

l'induction. Voir (Locke, [1689]2001; Mill, 1848; Whewell, [1860] 1971) pour les références classiques, (Kripke, 1972; Putnam, 1975; Quine, 1969) pour la réintroduction du concept dans la philosophie contemporaine et (Boyd, 1999b; Collier, 1996; Griffiths, 1997) pour des discussions plus récentes.

<sup>10</sup> (Kripke, 1972; Putnam, 1975)

<sup>11</sup> (Boyd, 1999a; Griffiths, 1999; Millikan, 2005).

<sup>12</sup> Comme *natural kind* a été traduit par *espèces naturelles*, il peut y avoir une confusion entre les espèces comme « kind » et les espèces comme « species ». Nous utiliserons le terme *lignée* ou *espèce biologique* pour parler des « species » (les souris, les criquets, etc.) et *espèce naturelle biologique* pour parler des « kinds » biologiques.

<sup>13</sup> (Ghiselin, 1974; Griffiths, 1999; Hull, 1978) voir les essais dans (Wilson, 1999) pour une défense et une critique. Notons aussi, pour éviter toute confusion, que lorsque des biologistes et philosophes de la biologie affirment que des lignées ne sont pas des espèces naturelles, ils impliquent par là qu'elles ne sont pas des espèces naturelles *éternelles*; la plupart reconnaissent que ce sont des espèces historiques.

<sup>14</sup> On distingue usuellement l'identité *numérique* (A et B sont identiques en tout point, et donc A=B) de l'identité *qualitative* (deux jumeaux sont qualitativement identiques parce qu'ils possèdent plusieurs qualités en commun, mais sont numériquement distincts).

---

<sup>15</sup> Précisons tout de même que « poisson » n'est pas un terme qui cadre exactement avec les taxonomies zoologiques (il s'agit d'un clade paraphylétique) mais nous ignorerons ces détails pour l'instant.

<sup>16</sup> (Hall, 1994; Ingo, 2003; Laubichler, 2000; Wagner, 1989, 2001). Plusieurs s'entendent aussi pour dire que le concept d'homologie peut varier selon les disciplines.

<sup>17</sup> 5 doigts/orteils, une « main », un membre inférieur doté de deux os et un membre supérieur doté d'un seul os.

<sup>18</sup> Voir (Griffiths, 1997, 2006, sous presse; Matthen, 2000) pour les problèmes des définitions téléofonctionnelles en biologie.

<sup>19</sup> (Millikan 1984; Neander 2002).

<sup>20</sup> Voir (Gilbert, 2006; Hall, 1994) pour plus de détails.

<sup>21</sup> (Brigandt, 2002; Griffiths, 1994, 2006, sous presse; Rieppel, 2005a, 2005b; Wagner, 1996, 2001) .

<sup>22</sup> L'inventeur du concept biologique d'homologie : (« le même organe chez différents animaux quelque soit [*under every variety*] la variété de forme et de fonction » (Owen, 1843, p. 379)).

<sup>23</sup> (Godfrey-Smith, 1996, 2003; Proust, 2006 ; Sterelny, 2003a, 2003b).

---

<sup>24</sup> Notons, à la différence de (Stein, 1996) et (Samuels & Stich, 2004), que la compétence rationnelle qui nous intéresse ici n'est pas la compétence à réfléchir, mais à agir de façon rationnelle.

<sup>25</sup> (Krebs & Davies, 1997; McFarland, 1999; Stephens & Krebs, 1986).

<sup>26</sup> Voir (Dugatkin & Reeve, 1998; Giraldeau & Caraco, 2000).

<sup>27</sup> (Noë & Hammerstein, 1995; Noë *et al.*, 2001).

<sup>28</sup> Voir (Schwartz & Hoeksema, 1998).

<sup>29</sup> (Smith, 1991; Winterhalder & Smith, 2000).

<sup>30</sup> Cf. (Nieoullon, 2002; Schultz, 2001; Schultz & Dickinson, 2000; Williams & Goldman-Rakic, 1995).

<sup>31</sup> (Callier *et al.*, 2003; Le Crom *et al.*, 2003).

<sup>32</sup> cf. (Berridge, 2004; Pecina *et al.*, 2003).

<sup>33</sup> (Pecina *et al.*, 2003; Robinson *et al.*, 2005).

<sup>34</sup> (Bayer & Glimcher, 2005; P.R. Montague *et al.*, 2004; Schultz, 1999; Schultz *et al.*, 1997).

<sup>35</sup> (Dayan & Balleine, 2002; P. Read Montague & Berns, 2002; P. R. Montague *et al.*, 2006),

---

<sup>36</sup> La branche de la cladistique spécialisée dans les homologie neuronales (Nieuwenhuys *et al.*, 1998)

<sup>37</sup> (Kapsimali *et al.*, 2003, p. 33; Le Crom *et al.*, 2003, p. 33)