

# La métacognition chez l'animal non humain

KEBA COLOMA CAMARA, *Université Laval*

## *Introduction*

L'homme est depuis très longtemps considéré comme un être supérieur à l'animal non humain. Cette opinion est largement défendue par Descartes<sup>1</sup> lorsqu'il considère l'animal comme une simple machine dépourvue de raison qui n'agit que par instinct; contrairement à l'homme qui possède une âme pensante. L'anthropomorphisme défend également la même opinion : il consiste à attribuer certaines capacités humaines aux animaux non humains sans pour autant accepter que les animaux possèdent réellement ces capacités<sup>2</sup>. À la question les animaux pensent-ils?, certains philosophes tels que Donald Davidson<sup>3</sup> répondent par la négative, pour la simple raison que les animaux ne possèdent pas de croyance. Toutes ces thèses considèrent l'homme comme une exception dans la nature et l'animal non humain, privé de raison, lui reste inférieur. Mais, avec les recherches et découvertes contemporaines sur la cognition animale et l'organisation sociale des animaux non humains, il semble que nous assistons à la fin du mythe de l'exception humaine professée par Jean-Marie Schaeffer<sup>4</sup> et le début de la redéfinition de l'homme proposée par Joëlle Proust<sup>5</sup>.

La rationalité qui représentait jusque-là la frontière entre l'homme et l'animal peut être conçue, selon certains chercheurs, comme un produit de l'évolution<sup>6</sup> et il existe des formes primitives de rationalité chez les animaux non humains<sup>7</sup>. De nombreuses expériences accréditent cette thèse en s'appuyant sur les principes de l'évolution biologique. La raison n'est plus identifiée à l'âme, elle est plutôt considérée comme un système de commande cognitif qui permet à l'animal et à l'homme de réaliser et de contrôler leurs cognitions. Si les chercheurs, dans le domaine de la cognition animale, acceptent cette définition minimale de la rationalité; ils

ne sont pas, pour la plupart, convaincus que l'animal non humain possède la capacité métacognitive : c'est-à-dire la capacité à suivre et à évaluer les cognitions. Pour certains, cette aptitude relève du niveau supérieur de la rationalité, exclusif à l'homme, et l'animal qui n'a atteint que le niveau primitif est incapable de manifester des aptitudes métacognitives. Est-ce un mythe philosophique ou un raisonnement scientifique ? La métacognition<sup>8</sup> représente, pour le moment, une frontière cognitive qui sépare l'homme de l'animal non humain.

Les animaux non humains ont-ils des aptitudes métacognitives ? La réponse à cette question varie selon les théories et les philosophes. Deux conceptions s'opposent : la première, la conception auto-attributive, considère que pour qu'un organisme soit capable d'aptitude métacognitive, il doit posséder au préalable une théorie de l'esprit et le mindreading<sup>9</sup> (la capacité d'attribution d'états mentaux à autrui). Selon cette conception, si l'animal non humain ne dispose pas de ces éléments alors il ne possède tout simplement pas la capacité métacognitive. Cette hypothèse est rejetée par la seconde conception : la conception auto-évaluative. Cette conception, en s'appuyant sur des résultats en neurosciences cognitives, estime que les animaux non humains ont la capacité métacognitive ; car la métacognition est un système de commande qui ne nécessite pas l'acquisition du mindreading ni d'une théorie de l'esprit.

Dans les sections qui suivent, nous proposerons, d'abord, une définition du concept de métacognition. Ensuite, nous examinerons les arguments de la conception auto-attributive de la métacognition et, notamment, l'approche de Peter Carruthers qui considère que les preuves métacognitives découvertes chez les animaux non humains, défendues par la conception auto-évaluative, ne sont pas suffisantes pour attribuer la capacité métacognitive aux animaux non humains. Dans la troisième section, nous discuterons des arguments de la conception auto-évaluative qui défend l'existence des capacités métacognitives chez les animaux et les problèmes auxquels ils sont confrontés, avant d'examiner les différentes formes de métacognition. Nous terminerons par les perspectives de Joëlle Proust qui mettent

en valeur les enjeux de l'attribution de la métacognition chez l'animal non humain.

### *1. Qu'est-ce que la métacognition ?*

La définition la plus simple à retenir de la métacognition est la suivante: c'est une cognition de la cognition. Lorsqu'on parle de cognition, on fait référence aux processus cérébraux et mentaux tels que le traitement de l'information, la mémoire, la connaissance, le raisonnement, l'apprentissage. La métacognition est alors la cognition qui permet d'évaluer, de contrôler et de réguler le fonctionnement des autres cognitions.

Le concept de métacognition est utilisé tout aussi bien en psychologie cognitive, en philosophie qu'en sciences de l'éducation. Même si son usage est relatif par rapport aux préoccupations de chaque domaine, il y a tout au moins deux aspects qui entrent en jeu lorsqu'on utilise ce concept<sup>10</sup>: il s'agit du savoir métacognitif et de l'expérience métacognitive. Le premier représente les métaconnaissances c'est-à-dire les savoirs et les croyances stockées dans la mémoire à long terme qui affectent les processus cognitifs. Ces métaconnaissances permettent aux individus, par exemple, de savoir comment :

- agir dans le cadre des relations humaines,
- trouver la disponibilité d'une information,
- opter pour les meilleures stratégies face aux situations, etc.

Le second aspect, les expériences métacognitives, renvoie à l'évaluation « des expériences cognitives et affectives conscientes<sup>11</sup> ». Le contrôle, le suivi et l'évaluation de ces expériences cognitives et affectives permettent le contrôle des émotions et d'anticiper ou de planifier l'action du sujet face aux situations qu'il rencontre. Ainsi l'individu est-il généralement en mesure de contrôler son niveau de confiance par rapport aux situations qu'il rencontre.

Proust<sup>12</sup> distingue deux fonctions de l'activité métacognitive. La première, écrit-elle, permet de décider de la faisabilité d'une activité mentale. Par exemple, si nous proposons à un enfant de

l'école primaire de résoudre une équation mathématique à deux inconnus, il n'a certainement pas encore appris à résoudre ce type d'équation ; il doit être en mesure de savoir, après une rapide évaluation métacognitive, s'il est compétent ou non à résoudre ce type d'équation avant même de tenter quoi que ce soit. Si on propose la même équation, par ailleurs, à un étudiant en mathématique, qui a probablement rencontré ce type de problème mathématique dans son cursus scolaire, et qu'il accepte de le résoudre, c'est parce qu'il sait, après évaluation métacognitive, qu'il en est capable. La décision métacognitive permet, en effet, d'évaluer les aptitudes cognitives de l'individu. L'individu sait ce qu'il est capable de réaliser ou non, après avoir évalué la situation dans laquelle il se trouve par rapport à ces aptitudes cognitives.

La seconde fonction de la métacognition consiste à évaluer si l'activité cognitive ou mentale a été réussie ou non. Reprenons notre exemple : l'enfant va certainement refuser de résoudre une équation qu'il ne connaît pas ; par contre, l'étudiant en mathématique qui s'engage à résoudre le problème doit être en mesure de savoir si le résultat qu'il a trouvé est conforme au raisonnement mathématique pour lequel il a opté, faute de quoi, il peut douter de son résultat.

La métacognition est une activité cognitive qui permet, d'une part, d'évaluer, de contrôler ou de faire le suivi des activités cognitives prospectives, c'est-à-dire prendre une décision sur la faisabilité ou non d'une action cognitive (première fonction) ; d'autre part, elle permet de faire un suivi rétrospectif de l'activité cognitive, par exemple l'évaluation de la conformité du résultat par l'étudiant en mathématique (seconde fonction). La métacognition ne se limite pas, cependant, à ces deux cas de figure ; elle se manifeste également dans l'évaluation de la cognition dans les tâches quotidiennes telles que l'apprentissage, le jugement ; etc.

Il existe, en outre, un autre élément essentiel dans l'activité métacognitive : c'est le rôle des sentiments. Certains philosophent<sup>13</sup>, en effet, ont défendu la thèse selon laquelle les émotions et la raison représentent deux niveaux distincts, les émotions sont un obstacle à la raison<sup>14</sup>. Pour parvenir à un raisonnement ou un jugement rationnel, l'individu, croient-ils, doit pouvoir contrôler ses émotions

en s'en détachant, car c'est seule la raison qui permet d'atteindre l'objectivité dans la réflexion. Les sentiments représentent quelquefois des obstacles au raisonnement rationnel et objectif; mais est-ce une raison de les négliger dans les activités rationnelles? D'autres philosophes<sup>15</sup> ont montré, par contre, l'importance des sentiments dans le raisonnement. Les sentiments jouent, en effet, un rôle épistémique dans l'évaluation d'une activité cognitive. Proust<sup>16</sup> distingue différentes fonctions des sentiments dans l'activité métacognitive: par exemple, dans une tâche métacognitive prospective ou rétrospective, la réponse métacognitive peut être un sentiment d'incertitude ou de confiance, de familiarité, de savoir, d'avoir un mot sur le bout de la langue, d'échec, de cohérence, d'incohérence, etc.

La raison et les sentiments entretiennent alors une relation complexe et complémentaire. L'être rationnel est tout aussi bien l'être sentimental. La raison n'a pas tout à fait le pouvoir d'exclure les sentiments; ces dernières sont tout aussi bien en mesure de l'orienter pour les choix meilleurs dans une tâche cognitive. Le rôle des sentiments métacognitifs dans les évaluations cognitives doit nous orienter à penser que la capacité métacognitive n'est pas l'apanage exclusif de l'homme, l'animal non humain doué de sentiment et, au moins, une forme primitive de raison peut aussi en être capable. Car l'évaluation métacognitive est, d'ailleurs, motivée à la fois par les émotions et par la raison.

Le problème que nous soulevons à propos du concept de métacognition est le suivant: à l'instar du sujet humain qui fait usage des aptitudes métacognitives pour évaluer et contrôler sa cognition, l'animal non humain a-t-il des aptitudes métacognitives? Il existe une hypothèse à cette question: si l'animal non humain a des aptitudes cognitives, il est tout à fait possible que ses activités cognitives soient accompagnées d'activités métacognitives, et ces dernières permettent de contrôler et d'évaluer les premières. Mais, cette hypothèse, aussi cohérente qu'elle soit, fait face à des problèmes conceptuels tels que le *mindreading* et la théorie de l'esprit.

Deux approches vont s'opposer dans l'étude de la métacognition chez l'animal non humain. En effet, les défenseurs du *mindreading*<sup>17</sup>

estiment qu'il ne peut exister une activité métacognitive chez un individu qui ne possède pas une théorie de l'esprit<sup>18</sup> : c'est la conception auto-attributive ou la Self-Ascriptive View (SAV). Le mindreading et la métacognition, défendent-ils, jouent des rôles analogues : le mindreading à propos d'autrui et la métacognition à propos de soi. Alors la théorie de l'esprit est une condition nécessaire et suffisante pour la métacognition. Contrairement aux partisans de la SAV, ceux de la conception auto-évaluative ou la Self-evaluative View (SEV) estiment que la métacognition et le mindreading jouent deux fonctions<sup>19</sup> distinctes qui peuvent néanmoins se compléter. Alors les animaux non humains qui ne disposent pas d'une théorie de l'esprit sont tout à fait capables de posséder des aptitudes métacognitives.

L'enjeu du problème de la métacognition est d'abord théorique : la définition de la métacognition est différente dans la SAV et la SEV. Pour la première conception, défendue par les spécialistes de la théorie de l'esprit, la métacognition renvoie à la métareprésentation de ses propres états mentaux<sup>20</sup>. Cela signifie que la capacité métacognitive est déterminée par la théorie de l'esprit c'est-à-dire les connaissances dont on dispose sur son propre esprit. Or, ces connaissances sont formées par des représentations portant sur d'autres représentations acquises culturellement ou de façon innée. La métacognition a, selon cette conception, une structure métareprésentationnelle puisque les connaissances métacognitives portent nécessairement sur des contenus représentationnelles. La métacognition et la métareprésentation font donc appel à des capacités similaires. Cependant, pour la seconde conception, acceptée principalement en sciences cognitives, les connaissances métacognitives ne sont pas nécessairement métareprésentationnelles. Selon cette conception, la métacognition est la capacité dont dispose un sujet pour évaluer sa performance cognitive<sup>21</sup>. La capacité métareprésentationnelle ou la théorie de l'esprit n'est donc pas une condition nécessaire à l'activité métacognitive. Les animaux non humains, en effet, qui ne disposent pas de théorie de l'esprit sont tout à fait en mesure de posséder la capacité métacognitive. C'est la perspective de la SEV que nous espérons défendre et approfondir dans les sections qui suivent. Mais, pour le moment, examinons l'orientation de la SAV.

## *2. La SAV: Carruthers et la HOT.*

### *2.1 L'approche auto-attributive (SAV): la métacognition une capacité strictement humaine*

Même si les versions sont différentes, il existe néanmoins un point sur lequel les théoriciens de la SAV sont unanimes: pour qu'un individu puisse posséder une capacité métacognitive, il doit nécessairement posséder une théorie de l'esprit et avoir la capacité de lire l'esprit des autres. La métacognition passe par la métareprésentation et cette dernière nécessite l'acquisition du mindreading. Ils considèrent, en général, la métacognition comme une activité du second ordre.

Certains théoriciens de la SAV s'inspirent, en effet, de ce schéma de Nelson et Narens<sup>22</sup>: le fonctionnement de tout système, précisent-ils, a besoin nécessairement de deux niveaux, un premier niveau orienté dans l'exécution des tâches et un second niveau qui permet le suivi et le contrôle des tâches exécutées au premier niveau. Tout ce qui est réalisé, en effet, au premier niveau est vérifié et contrôlé au second niveau, cela permet d'assurer le bon fonctionnement du système, c'est-à-dire la correspondance entre la commande, l'exécution et le résultat attendu. La conception auto-attributive, s'appuie sur ce modèle pour expliquer que la métacognition est une activité cognitive du second ordre qui doit nécessairement représenter la tâche cognitive du premier ordre. Cela signifie que la métacognition passe par le mindreading qui implique la métareprésentation. En effet, la représentation métacognitive s'appuie nécessairement, selon les théoriciens de la SAV, sur d'autres représentations mentales c'est-à-dire d'autres états mentaux du sujet qui s'auto-attribue, finalement, le résultat métacognitif. Si je crois, par exemple, que P est Y c'est que j'ai déjà des croyances à propos de P ce qui me permet de le qualifier de Y.

Selon John H. Flavell<sup>23</sup>, un des premiers théoriciens de la SAV, il y a une relation entre la métacognition et l'attribution d'états mentaux à autrui. Les enfants âgés de moins de 4 ans, selon cet auteur, ne possèdent pas le mindreading car ils sont incapables de métareprésenter leurs croyances et celles des autres. Aussi sont-ils incapables de distinguer l'apparence et la réalité des objets et

phénomènes qui les entourent. Mais, dès qu'ils comprennent, après 4 ans, qu'ils possèdent un esprit capable de se représenter soi-même, les autres et le monde, alors ils parviennent à discriminer les croyances et à distinguer ce qui relève de la réalité et de l'apparence. Flavell considère en fait, à l'instar des praticiens des sciences de l'éducation, que la métacognition fait partie de la cognition sociale. Les connaissances métacognitives sont inséparables de la connaissance de soi et des autres. Il est donc nécessaire de posséder une théorie de l'esprit qui permet d'évaluer les connaissances métacognitives, car sans elle, il n'y aura pas de capacité métacognitive. Un autre théoricien de la SAV Joseph Perner<sup>24</sup> abonde dans le même sens : pour lui, l'enfant n'acquiert une théorie de l'esprit et une capacité métacognitive qu'au moment où il est capable de se faire une représentation du rapport entre une représentation mentale X et l'état d'un objet Z. L'activité métacognitive a, selon la SAV, la même structure cognitive que celle du mindreading<sup>25</sup> (connaître ses propres états mentaux et inférer ceux des autres).

Si tel est le cas, il est impossible d'attribuer aux animaux la capacité métacognitive étant donné que l'animal non humain ne possède pas une théorie de l'esprit. La métacognition est, selon la SAV, une capacité strictement humaine. L'animal peut, toutefois, avoir une pensée du premier ordre pour agir dans son environnement, mais il ne possède pas la pensée du second ordre, c'est-à-dire l'aptitude métacognitive, puisqu'il ne possède pas de mindreading.

## *2.2 Carruthers, la théorie de la pensée d'ordre supérieur (HOT) et la métacognition*

Peter Carruthers<sup>26</sup>, un partisan de la SAV, accepte l'argument selon lequel l'activité métacognitive a la même structure cognitive que le mindreading ; autrement dit, la métacognition et le mindreading n'utilisent qu'une seule voie d'accès : l'agent s'attribue des représentations mentales qu'il infère d'autres représentations. Mais il propose, avec sa théorie de la pensée d'ordre supérieur<sup>27</sup> (Higher-order Thought ou HOT), une explication du premier ordre de la plupart des activités métacognitives. Selon cet auteur, la métacognition et le mindreading se font de façon conceptuelle,



c'est-à-dire inférentielle ou par métareprésentation. En effet, si l'accès à la mémoire ou à la perception se fait de façon directe et non inférentielle, la métacognition, cependant, passe nécessairement par inférence ou par métareprésentation. L'idée est que les connaissances métacognitives passent nécessairement par le mindreading et que ce dernier requiert des capacités métareprésentationnelles. Or, sans le mindreading, il ne peut y avoir de métacognition. Les connaissances métacognitives représentent les capacités représentationnelles dont on dispose : après inférence le sujet s'attribue ces connaissances.

Carruthers défend la thèse suivante : la métacognition est une représentation contenue dans une autre représentation. Il est possible à l'homme de recourir à l'explication métacognitive, puisqu'il possède une pensée d'ordre supérieur, c'est-à-dire une théorie de l'esprit ; mais l'animal non humain qui ne possède pas ce type de pensée est, du coup, incapable d'exercer une activité métacognitive. Tout ce que l'animal est capable de faire relève de la cognition et il n'existe pas du tout d'activité métacognitive. Pourtant les études récentes ont montré que les singes et les dauphins ont des aptitudes métacognitives<sup>28</sup>. En outre, selon certaines expériences, qui sont pour le moment discutables, les singes possèdent la capacité du mindreading<sup>29</sup>. Carruthers écrit :

I shall subject the main studies referred to above to sustained critique, arguing that there is no need to postulate *meta-cognitive*<sup>30</sup> processing in order to explain the data. Rather, I shall show how those data admit of explanation in first-order terms, appealing only to states and processes that are world-directed rather than self-directed. I shall argue, in consequence, that we should, at present, refuse to attribute *meta-cognitive*<sup>31</sup> processes to animals<sup>32</sup>.

Carruthers parvient, en effet, à proposer une explication du premier ordre à tous les exemples<sup>33</sup> que certains auteurs<sup>34</sup> qualifient de preuves métacognitives chez l'animal non humain. Mais l'explication qu'il propose est finalement très sophistiquée puisqu'elle traite de différents types de cas. Étant donné qu'il n'accorde pas l'aptitude métacognitive à l'animal non humain, il applique la même rigueur

d'explication du premier ordre aux croyances de l'homme dans le but de réduire la frontière cognitive entre l'homme et l'animal. En d'autres termes, il est possible d'expliquer à la fois le comportement de l'animal comme celui de l'homme en faisant référence au mécanisme du premier ordre c'est-à-dire comment l'homme et l'animal parviennent à avoir des désirs et des croyances, à les évaluer et à réaliser des actions cognitives.

Même si la démarche de Carruthers paraît quelquefois convaincante, il n'en demeure pas moins que la thèse qu'il défend dans sa théorie de la HOT fait face à un enjeu déterminant quant au statut de la cognition animale. Deux éléments nous paraissent suffisants pour s'écarter de la perspective des théoriciens de la SAV et notamment de celle de Carruthers. Premièrement, les études récentes en sciences cognitives et en éthologie démontrent de plus en plus que l'action cognitive des animaux non humains est tout aussi complexe que celle des humains ; et il existe assez de preuves convaincantes que les animaux non humains possèdent une capacité métacognitive.

Le second élément est épistémologique, il est structuré à deux niveaux. D'abord, il nous semble impossible qu'un système puisse causer des activités cognitives sans pour autant être capable de réaliser des activités métacognitives qui sont censées évaluer les activités cognitives. Alors il est tout à fait cohérent de présumer que les animaux non humains possèdent la capacité métacognitive ; sinon comment parviendront-ils à discriminer les phénomènes, à choisir leur proie et à vivre dans leur milieu ? Or toutes ces situations nécessitent, à notre avis, l'activité métacognitive, même si les partisans de la SAV, principalement Carruthers, préfèrent analyser ces situations sous l'angle de la cognition plutôt que la métacognition. Ensuite, l'outil conceptuel utilisé par les théoriciens de la SAV et principalement Carruthers (HOT) impose des critères censés être simples et suffisants pour évaluer la cognition animale, mais nous constatons qu'ils sont, paradoxalement, complexes et leur application s'avère difficile dans l'étude de l'esprit des animaux non humains. Ces auteurs commettent une erreur de principe lorsqu'ils estiment que si l'animal ne possède pas, par exemple, le mindreading

alors il n'a pas de capacité métacognitive. Comme nous le verrons dans la section suivante, le mindreading n'est pas nécessaire pour la métacognition.

### *3. La SEV et l'approche de Proust*

#### *3.1 L'approche auto-évaluative (SEV)*

Contrairement à la thèse défendue par les partisans de la SAV, les auteurs de la SEV tels que Proust et Koriat<sup>35</sup> considèrent que la fonction de la métacognition est différente de celle du mindreading. Cela signifie qu'il est possible pour un organisme de posséder la capacité métacognitive sans pour autant acquérir une théorie de l'esprit et le mindreading. En ce sens, les animaux non humains peuvent avoir la capacité métacognitive. Mais comment parviendront-ils à évaluer et à contrôler leurs états mentaux sans pour autant lire le contenu des croyances? L'enjeu de cette question représente l'essentiel du défi lancé par la SAV. Il existe quelques arguments pour répondre à cette question, examinons-les.

D'abord, les progrès réalisés dans la dernière décennie en psychologie comparative ont été favorables à la conception auto-évaluative. Les études sur la cognition animale faites par Hampton<sup>36</sup> et par Smith<sup>37</sup> ont permis de démontrer que certains animaux tels que le singe macaque et le dauphin manifestent des capacités métacognitives remarquables alors qu'ils ne possèdent ni la théorie de l'esprit ni le mindreading. Dans les différents tests auxquels ces animaux ont été soumis, lorsqu'ils sont en face de deux tâches A et B, ils choisissent la tâche la plus facile qu'ils sont capables d'effectuer; lorsqu'ils sont incapables de réaliser les tâches, ils refusent tout simplement de poursuivre les tests. Le macaque et le dauphin ont réussi les tests de discrimination visuelle et du rappel de mémoire, établis comme des critères démontrant des aptitudes à évaluer leur incertitude, donc qu'ils possèdent des capacités métacognitives. Pourtant ils échouent à des expériences mentalisatrices quand il s'agit de prédire le comportement d'un agent à partir de ses croyances et intentions. Ces animaux non humains parviennent à faire le contrôle et le suivi de leur cognition sans pour autant disposer d'une théorie de l'esprit ou du mindreading.

De ces expériences, on peut retenir la remarque suivante : la théorie de l'esprit n'est pas nécessaire pour la capacité métacognitive ; un organisme peut alors évaluer ses aptitudes cognitives sans posséder le mindreading. Mais, selon Carruthers, ce qui est considéré ici comme preuve métacognitive est tout simplement une cognition du premier ordre. Car la métacognition nécessite la métareprésentation et un organisme qui ne possède pas la pensée d'ordre supérieur (HOT) serait incapable de faire des inférences métacognitives.

Ensuite, des découvertes récentes faites en phylogénie et en neurophysiologie ont permis de distinguer les mécanismes de fonctionnement de la métacognition et du mindreading<sup>38</sup>. En neurophysiologie<sup>39</sup>, il a été démontré chez l'homme que les aires cérébrales qui entrent en jeu dans la métacognition et le mindreading sont distinctes : pour la métacognition, ce sont les aires sensorielles, le cortex préfrontal dorsal, le cortex préfrontal ventro-médian et le cortex cingulaire antérieur, etc. ; pour le mindreading, c'est le cortex préfrontal antéro-médian, la zone temporo-pariétale, etc. En phylogénie il y a une distinction entre le mécanisme de fonctionnement de la métacognition procédurale et la métacognition analytique<sup>40</sup> : «the first type relies on a variety of mechanisms to detect and control error ; the second is a distinct adaptation, which enables agents to understand error as false belief»<sup>41</sup>.

Enfin, l'argument décisif sur lequel la conception auto-évaluative s'appuie est conceptuel. En effet, les concepts de métacognition, de métareprésentation ou de mindreading peuvent être considérés comme des sous-ensembles distincts et complémentaires, mais ils ne sont pas nécessairement dépendants<sup>42</sup>. Il y a une hypothèse logique : le mindreading implique la métareprésentation et, souvent, la métacognition ; mais, une activité métacognitive peut être réalisée sans pour autant impliquer la métareprésentation. Les schizophrènes, par exemple, sont atteints de différentes sortes de perturbations psychologiques ; ils attribuent leurs actions à des agents extérieurs. Mais, c'est ce qui est remarquable, ces patients sont capables de rapporter et d'évaluer leurs états mentaux qu'ils attribuent aux autres. Ils délirent, mais ils rapportent parfaitement le contenu de leurs délires : ils sont d'ailleurs hyper réflexifs et rapportent même

les détails de leurs délires. En ce sens, la capacité métacognitive qui consiste à évaluer une action cognitive est distincte de la capacité métareprésentationnelle et du mindreading qui permettent de conceptualiser et d'inférer le contenu des représentations cognitives.

La métacognition est, généralement, directe et évaluative alors que le mindreading et la métareprésentation se font de façon indirecte et par inférence. Lorsqu'on me demande si je connais la nationalité du philosophe Kierkegaard<sup>43</sup>, avant de répondre à cette question, le mécanisme métacognitif me permet de savoir si je dispose des informations à propos de ce nom : voici une évaluation métacognitive. Le résultat de l'évaluation déterminera, par ailleurs, si je dois me souvenir au cas où je sais que je possède des informations sur cette personne ; à défaut, je répondrai que je n'ai aucune connaissance de cette personne.

Pour dépasser l'opposition entre ces deux principales conceptions, à savoir la SAV et la SEV, nous proposons d'examiner les différentes formes de métacognition dans lesquelles chacune de ces approches peut trouver un écho de ses arguments.

### *3.2 Les formes de métacognition*

La métacognition définie comme une cognition sur la cognition est sans doute difficile à vérifier chez les animaux non humains. Robert Hampton<sup>44</sup> a proposé quatre critères pour opérationnaliser le concept de métacognition :

- a) il faut nécessairement un comportement primaire identifiable ;
- b) ensuite, une incertitude du résultat : l'agent doit faire face à une situation d'incertitude à partir de laquelle il doit chercher une solution métacognitive à partir de ses capacités cognitives ;
- c) puis, il faut un comportement secondaire qui permet de réguler le comportement primaire : suivi et évaluation des capacités pour l'exécution ou non du comportement primaire ;
- d) enfin, le résultat du comportement secondaire est nécessairement évalué.

En appliquant ces critères dans la cognition animale, Hampton a réussi à distinguer deux formes de métacognition : *la métacognition publique et la métacognition privée*. La première fait référence aux mécanismes métacognitifs publiquement évaluables. Par exemple, un enseignant pose une question à un élève sur le nombre de planètes, la réponse de l'élève sera non seulement basée sur ses connaissances acquises, mais le reste de la classe peut évaluer la conformité de la réponse de leur camarade par rapport aux connaissances acquises durant la leçon sur les planètes. Il est également possible d'évaluer les aptitudes métacognitives de l'animal non humain dans des tests publiquement observables. Dans la métacognition publique, les mécanismes métacognitifs se traduisent dans le comportement observable (hésitation dans le test, refus, acceptation, etc.) La seconde forme permet de surveiller et de contrôler ses processus cognitifs non observables publiquement. En fait, dans la métacognition privée il s'agit d'une évaluation personnelle sur une situation cognitive particulière dont on ne peut observer les mécanismes. L'introspection est un moyen qui permet d'expérimenter et de vérifier les mécanismes de la métacognition privée.

Proust<sup>45</sup> propose également une distinction importante entre deux types de métacognition. Il s'agit de la *métacognition théorique et la métacognition procédurale*. Les mécanismes de la métacognition théorique dépendent de la métareprésentation, par exemple, une représentation de la croyance et de son contenu. La métacognition théorique est ce qu'on appelle également métacognition analytique. Ce type de métacognition nécessite une théorie de l'esprit avec le mindreading c'est-à-dire la capacité à lire le contenu des croyances. La métacognition théorique ou analytique se fait alors par inférence par le moyen de la métareprésentation.

La métacognition procédurale s'articule principalement sur l'évaluation des aptitudes cognitives en fonction des objectifs cognitifs visés, elle ne nécessite pas une métareprésentation parce qu'elle accède directement à l'information sans faire des inférences. Proust note deux fonctions importantes de cette forme de métacognition : une fonction instrumentale et une fonction exécutive. La fonction instrumentale est orientée vers l'obtention du

résultat ou la récompense visée par l'agent. Par exemple, l'animal dans son environnement est capable de discriminer le type de proie qui l'intéresse le mieux et choisir le moyen (la technique de chasse) qu'il doit utiliser pour atteindre son objectif. La fonction exécutive quant à elle consiste à organiser et à allouer les ressources cognitives indispensables au succès de l'événement métacognitif. Dans l'organisation et l'utilisation des ressources, il y a nécessairement l'évaluation et la surveillance de la cognition par rapport à la norme ou l'objectif métacognitif visé. On peut dire que la fonction instrumentale de la métacognition se rend possible par le concours de la fonction exécutive.

On peut noter, à la lumière de ces formes de métacognition, que la conception auto-attributive ne prend en charge que la métacognition analytique ou théorique. Car ces formes de métacognition nécessitent l'acquisition d'une théorie de l'esprit, le mindreading et la métareprésentation. Le mécanisme de ces formes de métacognition est indirect parce qu'il se réalise par inférence. La métacognition procédurale, essentiellement épistémique, est prise en charge par l'orientation de la SEV. Le mécanisme de cette forme de métacognition est direct et permet d'évaluer les aptitudes en fonction des situations précises. En somme, l'usage du concept de métacognition doit être déterminé par les formes de métacognition. On peut donc affirmer que l'animal non humain a des capacités métacognitives en se référant à la métacognition procédurale.

Mais la métacognition procédurale essentiellement épistémique est une forme métacognitive pratique et limitée. Cela représente un problème dans l'approche de la SEV et donne souvent raison aux partisans de la SAV. En effet, là où l'homme est capable de s'attribuer des connaissances métacognitives qui passent par la métareprésentation; l'animal non humain, par la métacognition procédurale, n'agit que de façon directe. En ce sens, il serait difficile à l'animal de faire un retour sur ses propres états mentaux, savoir que c'est lui qui est l'auteur de ses pensées. Le problème pour la SEV est finalement d'aller au-delà de cette forme restrictive de la métacognition procédurale et de voir si les animaux non humains possèdent une théorie de l'esprit qui représente l'élément central

pour l'acquisition de toutes les formes de métacognition. Des études faites sur les singes sont prometteuses, mais elles sont confrontées, pour le moment, à des problèmes méthodologiques que les partisans de la SEV doivent surmonter.

Le dernier élément à explorer dans ce travail est le suivant : à quoi sert l'attribution de la capacité métacognitive à l'animal non humain ? En analysant la perspective des travaux de Joëlle Proust, nous trouvons des réponses à cette question.

### *3.3 Cognition, métacognition et statut de l'animal non humain*

Le débat sur la rationalité de l'animal non humain a opposé différentes thèses à propos de l'attribution ou non de la raison aux animaux non humains. Contrairement à certains philosophes comme Descartes<sup>46</sup> qui estiment que la raison est une propriété strictement humaine ; d'autres comme Davidson<sup>47</sup> qui pensent que le langage est un critère de la représentation mentale ; d'autres comme Bermudez et Proust ont essayé de démontrer les aptitudes rationnelles chez les animaux non humains. Proust<sup>48</sup> définit la raison comme un système de commande capable de s'adapter aux situations auxquelles l'agent est confronté et ce système de commande est déterminé par les aptitudes et connaissances dont dispose l'agent.

Il existe, en effet, une relation entre la rationalité et la métacognition. Mais le sens que l'on attribue au concept de rationalité détermine son rapport avec la métacognition. En général, la rationalité est définie comme la capacité dont dispose un agent à respecter les normes du raisonnement logique<sup>49</sup>. Chez les partisans de la SAV, la rationalité correspond à l'acquisition de la théorie de l'esprit et du mindreading. Être rationnel dans la perspective de la SAV, exige posséder un esprit, être capable de se représenter ses croyances et celles des autres. Cette forme de rationalité implique le mindreading, la métareprésentation et, du coup, ne prend en charge que la forme de la métacognition analytique ou théorique. La rationalité définie chez Proust comme une flexibilité ou une adaptabilité d'un système de commande est différente de la perspective de la SAV ; elle ne prend en compte que la forme de métacognition procédurale. Il est donc tout à fait possible qu'un animal soit rationnel et possède la capacité métacognitive,



dans la perspective de Proust, sans nécessairement posséder une théorie de l'esprit et le mindreading ; et sans qu'il soit nécessairement conscient. La rationalité, entendue comme adaptabilité et flexibilité du système de commande représente la capacité cognitive de l'agent. Cette capacité cognitive entretient alors un rapport étroit avec la métacognition procédurale qui permet, par exemple, d'évaluer la relation entre les ressources cognitives mobilisées par le système de commande d'un animal, et le but et les stratégies d'action à entreprendre.

Dans ce cadre, la métacognition procédurale permet aux animaux non humains d'évaluer leur cognition et le comportement des autres agents dans leur environnement de façon mentaliste. Autant que le sujet humain évalue sa pensée avant d'agir, autant l'animal non humain se sert de la métacognition procédurale avant d'engager une action en vue d'obtenir un résultat. Être rationnel ne nécessite pas chez un agent de posséder une théorie de l'esprit ou de représenter ses croyances par le langage ; c'est en fait la capacité cognitive de l'animal non humain. La métacognition procédurale permet à ce dernier d'avoir un comportement rationnel, en fonction des situations, en évaluant et en optimisant les ressources cognitives indispensables pour obtenir un résultat. Les capacités métacognitives sont en fait des preuves indubitables et élémentaires de la rationalité d'un animal non humain.

La propriété fondamentale de la métacognition est, écrit Proust, par définition, celle d'un système de contrôle opérant sur le système cognitif. La métacognition vise à assurer l'adéquation cognitive des moyens dont dispose le système et de ses buts cognitifs : il s'agit d'obtenir de l'information au niveau nécessaire et suffisant, ou de maintenir l'intensité de l'engagement ou du contrôle critique au niveau requis<sup>50</sup>.

La métacognition conduit également à une forme de connaissance et de contrôle de soi. Elle est, d'une part, une forme de contrôle de soi en ce sens qu'elle permet à l'agent de faire le suivi et l'évaluation de sa cognition ; d'autre part, elle permet à l'agent d'avoir certaines formes supérieures de conscience : la conscience de soi. Les preuves

métacognitives retrouvées chez le singe macaque par Hampton indiquent que la métacognition est un moyen de conscience de soi. Si l'animal non humain parvient à contrôler et à évaluer ses cognitions, il peut également être conscient qu'il est l'auteur de ses activités cognitives. La connaissance de ses aptitudes cognitives est aussi un moyen de connaissance de soi. «C'est en effet, écrit Proust, grâce à cette structure de contrôle que l'on peut se reconnaître soi-même comme tel, par opposition au cas où l'on se perçoit «objectivement», c'est-à-dire sans se reconnaître comme étant soi-même<sup>51</sup>» La métacognition est alors une condition nécessaire à la connaissance de soi. Cette condition n'est, cependant, pas suffisante, car il n'est pas évident que l'animal non humain puisse se reconnaître comme étant toujours l'auteur de ses propres cognitions. Cela ne peut être possible qu'avec la forme de métacognition analytique qui passe par la métareprésentation. C'est d'ailleurs, nous le rappelons, la difficulté principale à laquelle la SEV est confrontée. La métacognition procédurale est certes utile dans l'action pratique de la vie des animaux non humains, mais elle est limitée puisqu'elle ne permet pas de prouver rigoureusement d'autres formes de pensées telles que la conscience phénoménale et la représentation des croyances. Seule la métacognition analytique implique ces formes de pensée. Par contre, cela ne signifie pas que les animaux ne possèdent pas ces formes supérieures de pensée. Gennaro<sup>52</sup>, un partisan de la SAV, accepte non seulement que les animaux non humains aient la capacité métacognitive, mais qu'ils possèdent la conscience phénoménale sans pour autant acquérir la HOT. Le problème de la métacognition chez l'animal non humain est à la fois un défi, pour prouver que les animaux non humains possèdent certaines capacités métacognitives et une perspective, dans les recherches sur la relation que cette capacité mentale peut entretenir ou non avec les autres formes de pensées complexes.

### *Conclusion*

Le problème de la métacognition chez l'animal non humain est, de prime abord, un problème théorique. Le problème est tout à fait lié à la définition de la métacognition selon les différentes conceptions.

La métacognition est définie, en effet, selon les préoccupations conceptuelles de chaque conception: dans la SAV elle sert à s'auto-attribuer des cognitions tandis que dans la SEV elle sert à auto-évaluer ses aptitudes cognitives. Toutefois, en opérationnalisant ce concept dans la perspective de Hampton, on se rend compte qu'il n'existe pas une seule forme de métacognition, mais plusieurs formes qui se retrouvent dans les usages des différentes théories. Cela permet, néanmoins, de valider l'option de l'approche auto-évaluative sans invalider entièrement les considérations conceptuelles de l'approche auto-attributive. La métacognition procédurale permet de reconnaître les capacités métacognitives à l'animal non humain.

Les découvertes en éthologie et en sciences cognitives accréditent davantage les arguments de la SEV. Le problème est de savoir si la théorie de l'esprit et le mindreading, chers aux partisans de la SAV, sont présents chez les animaux non humains. Les arguments ne manquent pas pour espérer cette découverte et l'organisation sociale complexe et dynamique des animaux ne cesse de dissiper les doutes et les préjugés sur le statut de l'esprit des animaux non humains. Les recherches interdisciplinaires allant de la sociologie, à la linguistique, la philosophie et les sciences cognitives peuvent encore nous surprendre. Mais, pour le moment, certains philosophes résistent avec leurs arguments et les théories s'opposent.

- 
1. René Descartes, *Méditations métaphysiques*, Paris, Flammarion, 2010 (1641).
  2. Sandra D. Mitchell, «Anthropomorphism and Cross-Species Modeling Thinking with Animals» dans Lorraine Daston and Gregg Mitman (eds.) *New Perspectives on Anthropomorphism*, New York, Columbia University Press, 2005, V.
  3. D. Davidson, «Rational Animals» dans *Dialectica*, 1982, vol. 36, n° 4.
  4. J-M. Schaeffer, *La fin de l'exception humaine*, Paris, Gallimard, 2007.
  5. Joëlle Proust, «Redéfinir l'humain», *Le Débat* 3, 2014, n ° 180, pp. 56-69.
  6. *Id.* «Rationality and metacognition in non-human animals» dans Susan Hurley & Matthew Nudds (eds.) *Rational Animals?* Oxford, Oxford University Press, 2006, pp. 247-274.

7. J-L. Bermúdez, «Rationality and Psychological Explanation without Language» dans José Luis Bermúdez & Alan Millar *Reason and Nature : Essays in the Theory of Rationality*, Oxford, Oxford University Press, 2002.
8. Joëlle Proust, «Metacognition» dans *Philosophy Compass*, 2010, 5, 11, pp. 989-998.
9. Le lecteur doit garder à l'esprit la signification donnée à ce concept. En plus, je l'utiliserai couramment dans le texte sans l'usage des guillemets à des fins de commodité de lecture.
10. Olivier Houdé, *Vocabulaire des sciences cognitives*, Paris, Quadrige, 2003, pp. 289-294.
11. *Ibid.* p. 290.
12. J. Proust, «Metacognition and mindreading: one or two functions?» dans M. Beran, J. Brandl, J. Perner & J. Proust (Eds.) *The Foundations of Metacognition*, Oxford, Oxford University Press, 2012, pp. 234-251.
13. Les philosophes tels que Platon, Descartes et Kant considèrent l'émotion et les passions de façon générale comme un obstacle à la raison.
14. E. Kant écrit à ce sujet: «L'émotion est le sentiment d'un plaisir ou d'un déplaisir actuel qui ne laisse pas le sujet parvenir à la réflexion. Dans l'émotion, l'esprit surpris par l'impression perd l'empire sur lui-même» cf. E. Kant, *Anthropologie du point de vue pragmatique*, Paris, Flammarion, 1997 (1863).
15. R. de Sousa , *The rationality of emotion*, MIT, 1990 et Jonh Elster, *Alchemies of the Mind: Rationality and the Emotions*, Cambridge University Press, 1999.
16. J. Proust, «Metacognition and mindreading: one or two functions?» dans M. Beran, J. Brandl, J. Perner & J. Proust (Eds.), *The Foundations of Metacognition*, Oxford, Oxford University Press, 2012, pp. 246-247.
17. Nous soulignons.
18. P. Carruthers, «Meta-cognition in animals : A skeptical look» dans *Mind and Language*, 2008, 23, pp. 58-89.
19. J. Proust, *op.cit.* pp. 234-251.
20. Peter Carruthers, «How we know our own minds: The relationship between mindreading and metacognition» dans *Behavioral and Brain Sciences*, 2009, 32, pp. 121-138.
21. A. Koriat et al., «The intricate relationships between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behavior» dans *Journal of Experimental Psychology*, General, 2006, 135, 1, pp. 36-69.

22. T. O. Nelson and L. Narens, «Metamemory : a theoretical framework and new findings» dans T. O. Nelson (ed.) *Metacognition, Core Readings*, 1992, pp. 117-130.
23. J. H. Flavell, «Metacognition and cognitive monitoring: A new area of cognitive developmental inquiry» dans *American Psychologist*, 1979, 34, pp. 906-911.
24. J. Perner, T. Ruffman, «Episodic memory and automatic consciousness: Developmental evidence and a theory of childhood amnesia. Special Issue : Early memory» dans *Journal of Experimental Child Psychology*, 1995, 59, 3, pp. 516-548.
25. Onishi et Baillargeon ont essayé de prouver l'existence d'un mindreading biologique chez le bébé de 15 mois, et probablement chez les singes macaques puisque le singe macaque adulte a la même performance cognitive qu'un enfant de 3 ans. Ils estiment qu'il existe des formes précoces de mindreading, disons le proto-mindreading. La capacité à lire le contenu des croyances n'est pas strictement, selon ces auteurs, une acquisition sociale ; elle est d'abord innée. cf. K. H. Onishi, & R. Baillargeon, «Do 15-month-old infants understand false beliefs ?» in *Science*, 2005, 308, pp. 255-258.
26. Peter Carruthers, «Meta-cognition in animals: A skeptical look» dans *Mind and Language*, 2008, 23, pp. 58-89.
27. La pensée d'ordre supérieur est une pensée qui se situe au-delà de la simple cognition. Elle est par exemple rattachée aux croyances, à la conscience de soi et à la conscience phénoménale. Ces formes de pensées sont nécessairement liées à l'acquisition d'une théorie de l'esprit et du mindreading. La pensée d'ordre supérieur fait référence à la théorie de l'esprit chez Carruthers.
28. Pour le singe rhésus, cf. R. R. Hampton, «Rhesus monkeys know when they remember.» dans *Proceedings of the National Academy of Sciences U.S.A.*, 2001, 98, pp. 5359-5362 ; pour le dauphin cf. J. D. Smith et al. «The uncertain response in the bottle nosed dolphin» dans *Tursiops truncatus. Journal of Experimental Psychology: General*, 1995, 124, pp. 391-408. Ces deux auteurs démontrent par des expériences empiriques que les animaux non humains ont la capacité métacognitive.
29. Joëlle Proust, «Can non-human primates read minds?» dans *Philosophical Topics*, 2000, 27, 1, pp. 203-232.
30. Nous soulignons.
31. *Idem*.
32. Peter Carruthers, «Meta-cognition in animals : A skeptical look.» dans *Mind and Language*, 2008, 23, pp. 58-59.

33. Nous reviendrons sur certains de ces exemples dans les sections suivantes.
34. Il vise principalement l'article de Joëlle Proust, «Does metacognition necessarily involve metarepresentation?» dans *Behavior and Brain Sciences*, 2003, 26, 3, p. 352.
35. Parmi ces auteurs, on peut citer Koriat et al. (2006) Proust (2012).
36. R. R. Hampton, «Multiple demonstrations of metacognition in nonhumans: Converging Evidence or multiple mechanisms?» dans *Comparative Cognition and Behavior Reviews*, 2009, 4, pp. 17-28.
37. J.D Smith et al., «The comparative psychology of uncertainty monitoring and metacognition.» dans *Behavioral and Brain Sciences*, 2003, 26, pp. 317-373.
38. Joëlle Proust, «Metacognition and mindreading : one or two functions ?» dans M. Beran, J. Brandl, J. Perner & J. Proust (Eds.) *The Foundations of Metacognition*, Oxford, Oxford University Press, 2012, p. 248.
39. R. Kiani, R. & M. N. Shadlen, «Representation of Confidence associated with a decision by neurons in the parietal cortex» dans *Science*, 2009, 324, 5928, pp. 759-764.
40. Il existe différentes formes de métacognition, nous reviendrons sur ces deux formes de métacognition dans la section suivante. Cela évitera de confondre l'usage du terme «métacognition» chez Proust et chez Carruthers, puisqu'ils ne traitent pas de la même forme de métacognition.
41. Joëlle Proust, «Metacognition and mindreading : one or two functions ?» dans M. Beran, J. Brandl, J. Perner & J. Proust (Eds.) *The Foundations of Metacognition*, Oxford, Oxford University Press, 2012, p. 247.
42. *Id.* «La métacognition comme contrôle de soi» dans *Psychiatrie, Sciences Humaines, Neurosciences*, 2008, 6, 1, pp. 31-37.
43. Søren Kierkegaard (1813-1855) est un philosophe danois.
44. R. R. Hampton, «Multiple demonstrations of metacognition in nonhumans: Converging evidence or multiple mechanisms?» dans *Comparative Cognition and Behavior Reviews*, 2009, 4, pp. 17-28.
45. Joëlle Proust, «Metacognition and mindreading : one or two functions ?» dans M. Beran, J. Brandl, J. Perner & J. Proust (Eds.), *The Foundations of Metacognition*, Oxford, Oxford University Press, 2012, pp. 242-244
46. René Descartes, *op.cit.*
47. Donald Davidson, *op.cit.*
48. Joëlle Proust (2006) «Rationality and metacognition in non-human animals» dans Susan Hurley & Matthew Nudds (eds.) *Rational Animals ?*, Oxford, Oxford University Press, 2006, pp. 247-274.
49. O. Houdé *Op.cit.*, pp. 372-374.

50. J.Proust, «Les conditions de la connaissance de soi» dans *Philosophiques*, 2008, 27, 1, pp. 161-186.
51. *Ibid.*
52. R.Gennaro, «Animals, consciousness and I-thoughts Metacognition and mindreading : one or two functions?» dans Lurz R. (ed.) *Philosophy of Animal Minds*, Cambridge University Press, 2009, pp. 184-200.