DOIT-ON SE PASSER DE LA NOTION DE CONCEPT ?

*Edouard Machery
Department of History and Philosophy of Science
University of Pittsburgh*

Les concepts sont les représentations mentales qui figurent dans les processus cognitifs supérieurs (catégorisation, raisonnement inductif et déductif, création d'analogies, compréhension linguistique...).¹ La plupart des chercheurs en sciences cognitives supposent qu'ils forment une classe homogène de représentations mentales. Bien sûr, ils admettent qu'il existe de nombreuses différences entre les concepts ainsi qu'entre types de concepts : il existe des différences entre deux concepts comme CHIEN et CHAT ainsi qu'entre les concepts biologiques naïfs et les concepts des artefacts. Mais ils supposent que par-delà ces différences, les représentations mentales qui figurent dans les processus cognitifs supérieurs ont en commun de nombreuses propriétés. Le but d'une théorie psychologique des concepts est de décrire ces propriétés.

Le psychologue Gregory Murphy résume cette idée au début de son livre sur la psychologie des concepts (2002, 2-3) :

« The psychology of concepts cannot by itself provide a full explanation of the concepts of all the different domains that psychologists are interested in. This book will not explore the psychology of concepts of persons, musical forms, numbers, physical motions, and political systems. *The details of each of these must be discovered by the specific disciplines that study them; (...) Nonetheless, the general processes of concept*

¹ Dans ce qui suit, j'utilise strictement « concept » en ce sens. Les noms des concepts sont écrits en majuscules.

learning and representation may well be found in each of these domains. »²

Comme la plupart des psychologues, Murphy soutient ici que *les concepts forment une espèce naturelle*. Car, comme nous allons le voir, une espèce naturelle est une classe d'entités à propos desquelles de nombreuses généralisations inductives peuvent être formulées. J'appelle la thèse selon laquelle les concepts forment une espèce naturelle *l'hypothèse de l'espèce naturelle* (l'hypothèse EN).

L'hypothèse EN est souvent acceptée par les philosophes de la psychologie (par exemple, Prinz 2002), les psychologues et, plus généralement, les chercheurs en sciences cognitives.³ C'est en fait une hypothèse de base de la psychologie des concepts. Car le terme « concept » est un terme théorique. De manière générale, introduire et utiliser un terme théorique dans une science empirique revient à supposer que l'on peut formuler des généralisations scientifiquement pertinentes à propos de la classe dénotée – en d'autres mots, que cette classe forme une espèce naturelle. Inversement, si l'on ne peut formuler des généralisations inductives à propos d'une classe dénotée par un terme scientifique, autrement dit si la classe dénotée ne forme pas une espèce naturelle, la pertinence scientifique de la notion doit être rejetée (voir, par exemple, Griffiths 1996 sur la notion d'émotion). Les chercheurs en sciences cognitives croient que la notion de concept dénote une classe de représentations mentales qui est scientifiquement importante. Ils cherchent à formuler des généralisations à propos de cette classe.⁴ Ils admettent donc l'hypothèse EN.

Bien que l'hypothèse EN soit largement acceptée, je me propose de la réfuter dans cet article. Je soutiens en effet la thèse suivante :

² Je souligne.

³ Un nombre croissant de psychologues rejettent cette thèse (par exemple, Smith et al., 1998).

⁴ Voir les recensions de la psychologie des concepts dans Smith et Medin, 1981, Lambert et Shanks, 1997, Laurence et Margolis, 1999 et Murphy, 2002.

De nombreuses données empiriques montrent que les concepts ne forment pas une espèce naturelle.

Je soutiens que contrairement à ce que l'on a supposé en sciences cognitives depuis trente ans et dans la psychologie philosophique depuis plusieurs siècles, les représentations mentales qui figurent dans nos processus cognitifs supérieurs ne forment pas une classe homogène à propos de laquelle de nombreuses généralisations inductives pourraient être formulées.

Dans la première section, j'explique la notion d'espèce naturelle. Dans la deuxième section, j'applique cette analyse à l'hypothèse EN et je présente un argument contre cette hypothèse, l'argument de la nécessité explicative (ANE). En bref, je soutiens qu'il faut admettre l'existence de plusieurs types de concepts qui ont peu en commun afin d'expliquer le cours des processus cognitifs supérieurs. Dans la troisième section, je présente les principales entités théoriques admises par les théories psychologiques des concepts. Finalement, je présente des données empiriques, tirées de la recherche psychologique sur la composition des concepts, qui montrent qu'il est légitime d'admettre ces différents types d'entités théoriques. Je conclus que l'hypothèse EN est fautive.

1. QU'EST-CE QU'UNE ESPÈCE NATURELLE ?

Dans la première section, j'analyse la notion d'espèce naturelle : que veut-on dire quand l'on qualifie une classe d'espèce naturelle ? Je cherche une notion d'espèce naturelle qui soit pertinente pour caractériser l'hypothèse EN. En particulier, cette notion ne doit pas être trop restrictive. Autrement, le rejet de l'hypothèse EN serait trivial.

1.1. ESPÈCE NATURELLE ET INDUCTION

La notion d'espèce naturelle est intrinsèquement liée au problème de *l'induction* (Quine, 1969, Boyd, 1990, 1991, Hacking, 1991). Car la notion d'espèce naturelle réfère aux classes à propos

desquelles des généralisations empiriques peuvent être faites. L'idée est la suivante : certaines classes, *les espèces naturelles*, sont telles que leurs membres ont de nombreuses propriétés en commun, tandis que les membres d'autres classes, par exemple la classe des objets qui pèsent plus de trente kilos, n'ont que très peu de propriétés en commun. Les membres des espèces naturelles sont donc supposés posséder de nombreuses propriétés en commun en sus des propriétés qui sont utilisées pour les identifier (Mill, 1905).

Puisque les espèces naturelles sont les classes qui donnent lieu à des généralisations inductives, elles sont les fondements des généralisations scientifiques. Ce sont les classes pertinentes en science (ou, au moins, dans de nombreuses sciences). La psychologie ne fait pas exception. Les psychologues cherchent à isoler les classes d'entités qui donnent lieu à des généralisations scientifiquement pertinentes. Ils introduisent des notions nouvelles pour dénoter de telles classes (par exemple, la notion d'émotion de base) et ils rejettent les notions qui n'identifient pas des espèces naturelles (par exemple, la notion d'émotion selon Griffiths).

1.2 DEUX FAUSSES PISTES : ESSENCES ET LOIS

Cette idée a été développée de plusieurs manières. On a ainsi traditionnellement associé les espèces naturelles à la possession d'une essence. On a aussi souvent associé les espèces naturelles aux lois.

Malgré leur intérêt, ces notions d'espèce naturelle sont inadéquates pour expliciter le sens de l'hypothèse EN. Car elles ne sont pas pertinentes pour caractériser des classes d'entités psychologiques. Car il est largement admis que les classes d'entités psychologiques sont fonctionnelles. Par conséquent, elles n'ont pas d'essence. Puisque les concepts sont des entités psychologiques, la notion d'espèce naturelle associée aux essences ne peut s'appliquer aux concepts. Il est aussi largement admis que, comme les autres sciences spéciales, la psychologie ne cherche pas à formuler des lois, mais des généralisations *ceteris paribus* (Fodor, 1974). Puisque la notion de concept est psychologique, elle n'est pas utilisée pour formuler des lois au sens strict. Donc, la notion

d'espèce naturelle associée aux lois ne s'applique pas aux concepts. Si une classe devait posséder une essence ou donner lieu à des lois pour être une espèce naturelle, trivialement, « concept » ne dénoterait pas une espèce naturelle. La négation de l'hypothèse EN serait sans intérêt.

1.3. LA NOTION CAUSALE D'ESPÈCE NATURELLE

La notion d'espèce naturelle n'est toutefois pas essentiellement liée aux notions d'essence et de loi. Les espèces naturelles peuvent en effet être caractérisées comme suit (Boyd, 1990, 1991, Griffiths, 1996). Une classe d'entités est une espèce naturelle si et seulement si c'est la classe maximale d'entités qui ont en commun ou qui tendent à avoir en commun un ensemble de propriétés à cause d'un ou plusieurs mécanismes causaux.

Cette définition résume trois idées principales. Une espèce naturelle est une classe à propos de laquelle des généralisations peuvent être faites : ses membres tendent à avoir des propriétés en commun. Ces généralisations ne sont pas accidentelles : il y a au moins un mécanisme causal qui explique pourquoi les membres de la classe concernée tendent à avoir ces propriétés. Enfin, cette classe n'est pas un sous-ensemble d'une classe à propos de laquelle on pourrait faire les mêmes généralisations. Une classe qui possède ces trois propriétés est dite être une espèce naturelle.

Dans le présent contexte, trois points sont importants. Tout d'abord, les propriétés qui caractérisent une espèce naturelle n'ont pas à être possédées par *tous* les membres de cette classe. Il est seulement requis que les membres d'une espèce naturelle *tendent* à avoir ces propriétés. En conséquence, les espèces naturelles n'ont pas à donner lieu à des lois au sens strict. Il est seulement requis qu'elles figurent dans des généralisations *ceteris paribus*.

Ensuite, les essences sont seulement l'un des mécanismes causaux par lesquels les membres d'une espèce naturelle ont en commun ou tendent à avoir en commun de nombreuses propriétés. Boyd a par exemple suggéré que l'homéostasie est un autre mécanisme (par exemple, Boyd, 1990, pour d'autres mécanismes, voir Griffiths, 1996).

Enfin, la notion causale d'espèce naturelle a une large extension : les matières, par exemple l'or, les particules physiques, par exemple les atomes, les espèces, par exemple, *macaca fuscata*, les artefacts... sont des espèces naturelles. Cependant, cette notion n'est pas vide. Tout d'abord, les classes nominales, par exemple la classe des objets physiques qui pèsent plus de trente kilos, ne sont pas des espèces naturelles selon cette définition, puisque leurs membres n'ont pas en commun de nombreuses propriétés en sus des propriétés utilisées pour les identifier. En outre, de nombreux sous-ensembles d'espèces naturelles, par exemple les chiens blancs, ne sont pas des espèces naturelles. En effet, les chiens blancs ne possèdent pas des propriétés spécifiques par rapport à l'espèce naturelle que forment les chiens – si ce n'est, bien sûr, la propriété *être blanc* qui nous sert à les identifier : la classe des chiens blancs n'est donc pas une classe *maximale* des entités qui ont ces propriétés.

Contrairement aux autres notions d'espèce naturelle, la notion causale est pertinente pour spécifier le sens de l'hypothèse de l'espèce naturelle. En effet, si l'on adopte cette notion d'espèce naturelle, les classes d'entités psychologiques peuvent être des espèces naturelles : les classes fonctionnelles ne sont pas exclues et les espèces naturelles n'ont pas à donner lieu à des lois au sens strict. *Dans ce qui suit, j'utilise donc la notion d'espèce naturelle exclusivement en ce sens.*

2. L'ARGUMENT DE LA NÉCESSITÉ EXPLICATIVE (ANE)

Dans la seconde section, j'applique l'analyse de la notion d'espèce naturelle menée dans la première section à l'hypothèse EN. Sur cette base, je présente l'argument contre l'hypothèse EN.

2.1. QUE SIGNIFIE L'HYPOTHÈSE DE L'ESPÈCE NATURELLE ?

Soutenir que les concepts forment une espèce naturelle, c'est affirmer que la classe des concepts possède les propriétés qui caractérisent les classes pertinentes en science. Quelles sont ces propriétés ? La notion causale d'espèce naturelle implique que la classe des concepts doit posséder les trois propriétés suivantes :

1. Il y a un large ensemble de propriétés S que les représentations mentales qui figurent dans nos processus cognitifs supérieurs tendent à posséder.
2. Cela résulte d'un ou plusieurs mécanismes causaux.
3. Cette classe de représentations n'est pas strictement incluse dans une classe à propos de laquelle les mêmes généralisations peuvent être formulées.

2.2. L'HYPOTHÈSE DE L'HÉTÉROGÉNÉITÉ

Les propriétés (1)-(3) sont nécessaires pour qu'une classe soit une espèce naturelle. Nier que les concepts forment une espèce naturelle, c'est donc affirmer que l'une des trois possibilités suivantes est réalisée :

- a. Les concepts n'ont pas en commun un grand nombre de propriétés.
- b. Ils ont bien des propriétés en commun, mais cela est accidentel.
- c. Ils ont bien des propriétés en commun, mais la classe des concepts est incluse dans une classe plus vaste de représentations qui possèdent les mêmes propriétés ; autrement dit, aucun ensemble de propriétés ne distingue la classe des concepts.

Dans le présent article, je défends la première option : *je soutiens que les concepts n'ont pas en commun un grand nombre de propriétés. Pourquoi est-ce le cas ? Parce que la classe des concepts se divise en types qui ont très peu en commun. Il existe plusieurs types de concepts, et ces types de concepts ont peu en*

commun. J'appelle cette thèse *l'hypothèse de l'hétérogénéité* (HH) :

(HH) La classe des concepts se divise en types de représentations qui ont peu en commun.

Il y a plusieurs lectures possibles de cette hypothèse. HH doit être comprise comme suit : *la plupart des catégories sont dénotées par de nombreux concepts qui appartiennent à des types différents.* De nombreuses catégories sont dénotées par plusieurs concepts. Et ces concepts appartiennent à des types de représentations qui ont peu en commun les uns avec les autres. Par exemple, il y a plusieurs concepts de chiens qui coréfèrent à la classe des chiens. Ces concepts de chien appartiennent à des types de représentations qui ont peu en commun.

2.3. L'ARGUMENT DE LA NÉCESSITÉ EXPLICATIVE

J'ai expliqué en détail la position que je défends dans cet article. Je présente maintenant la stratégie que je vais mettre en œuvre pour la défendre. L'idée principale est la suivante : il est nécessaire de postuler l'existence de plusieurs types de représentations qui ont peu en commun les uns avec les autres *si l'on veut expliquer nos compétences cognitives supérieures.* Lesquels ? Au moins ceux qui ont été postulés par les principales théories psychologiques des concepts (voir ci-dessous). En conséquence, l'image traditionnelle d'une classe homogène de représentations gouvernant les processus conceptuels est erronée. Si « concept » est le nom de cette classe, ces concepts ne forment pas une espèce naturelle. Je présente cette idée comme suit :

L'argument de la nécessité explicative (ANE)

1. Les concepts sont les représentations mentales dont les propriétés expliquent le déroulement des processus cognitifs supérieurs.

2. Les entités théoriques postulées par les principales théories des concepts ont peu de propriétés en commun.
3. Il est légitime de postuler *toutes* ces entités théoriques afin d'expliquer le déroulement des processus cognitifs supérieurs.
4. Donc, peu de propriétés sont communes aux représentations qui expliquent le déroulement des processus cognitifs supérieurs (de 2 et 3).
5. Donc, les concepts ne forment pas une espèce naturelle (de 1 et 4).

Dans cet article, j'admets simplement les deux premières prémisses (mais voir Machery, ms). Presque tous les psychologues et de nombreux philosophes acceptent *la prémisses 1* (par exemple, Laurence et Margolis, 1999, Prinz, 2002). Soulignons cependant qu'elle est controversée. En effet, elle a été constamment rejetée par certains philosophes, au premier rang desquels Jerry Fodor (1994, 1998). Selon Fodor, les concepts n'expliquent pas les propriétés de nos processus cognitifs supérieurs. Ils sont les composants de nos attitudes propositionnelles (croyances, désirs...) et ce sont ces attitudes qui expliquent le déroulement de nos processus cognitifs supérieurs. Les lecteurs qui acceptent une telle position peuvent lire ce qui suit de manière conditionnelle : s'il peut être montré que les concepts expliquent le cours des processus cognitifs, alors les concepts ne forment pas une espèce naturelle.

J'admets aussi *la prémisses 2*. Selon la prémisses 2, les théories psychologiques des concepts (voir ci-dessous) postulent des entités théoriques qui ont en fait fort peu en commun.

Dans cet article, je me concentre donc sur la prémisses centrale de l'argument de la nécessité explicative, à savoir la prémisses 3. Je présente certaines données empiriques qui montrent qu'il est justifié de postuler l'existence des différents types de représentations mentales proposés par les théories psychologiques contemporaines des concepts (voir Machery ms, pour d'autres données empiriques).

3. TROIS ENTITÉS THÉORIQUES

Depuis le rejet de ce que l'on appelle d'habitude le paradigme classique des concepts (PC) dans les années 70 (Posner et Keele, 1968, Rosch, 1978, Rosch et Mervis, 1975), trois paradigmes dominent la psychologie contemporaine des concepts, le paradigme postulant des prototypes (PP), le paradigme postulant des exemplaires (PE) et le paradigme postulant des théories (PT). On les traite habituellement comme des théories rivales à propos de la classe prétendument homogène des concepts (voir, par exemple, Laurence et Margolis, 1999, Murphy, 2002). Je soutiens que c'est une erreur : ces paradigmes portent en fait sur trois types différents de concepts dont il faut postuler l'existence pour expliquer les propriétés des compétences cognitives supérieures. Dans cette section, je décris brièvement ces trois entités théoriques. Dans la quatrième section, je montre qu'il est légitime de postuler leur existence.

3.1. LES PROTOTYPES

Le PP repose sur l'idée que les concepts sont des prototypes⁵ :

Un concept d'une classe d'objets est un *prototype*, c'est-à-dire *un ensemble de connaissances statistiques* à propos des propriétés possédées par les membres de cette classe, en particulier des propriétés possédées par *la plupart* des membres de cette classe (Rosch et Mervis, 1975).

Selon le PP, un concept est un ensemble de croyances statistiques : le possesseur du concept croit que les membres de la classe dénotée ont la probabilité x d'avoir la propriété P_1 , la probabilité y d'avoir la propriété P_2 etc. Ces propriétés ne sont pas jugées nécessaires. En outre, dans la plupart des modèles des prototypes (mais pas dans tous), ces croyances portent sur les

⁵ Le terme « prototype » est utilisé de manière ambiguë pour désigner le meilleur exemplaire d'une catégorie ou la représentation d'une catégorie. J'utilise strictement le terme pour désigner la représentation mentale d'une catégorie.

propriétés qui sont jugées être possédées par *la plupart* des membres de la classe dénotée (les propriétés *typiques*). Par exemple, posséder un concept de chien, c'est avoir une représentation d'une classe d'objets comme possédant en général, mais pas nécessairement, telle forme, comme étant en général composés de telles parties, comme ayant en général une queue...

3.2. LES EXEMPLAIRES

Peu après l'introduction du PP, certains psychologues ont suggéré une autre approche des concepts, le paradigme postulant des exemplaires (Brooks, 1978, Medin et Shaffer, 1978). Le PE repose sur l'idée suivante : nos performances cognitives supérieures font appel à des représentations d'instances spécifiques, qui sont utilisées comme des représentants de la catégorie entière.

Un concept d'une classe d'objets est un ensemble d'exemplaires, c'est-à-dire de représentations de membres particuliers de cette classe.

Par exemple, avoir un concept de chien, c'est avoir un ensemble de représentations de chiens particuliers, Pluto, Médor, Fido etc. et c'est utiliser ces représentations pour raisonner à propos des chiens en général. La classe des chiens est caractérisée comme la classe des entités qui sont semblables à Fido, Pluto et Médor.

3.3. LES CONCEPTS THÉORIQUES

Le PT a été développé indépendamment au milieu des années 80 par Gregory Murphy et Douglas Medin (1985) d'un côté et par Susan Carey (1985) de l'autre. L'idée de base est la suivante : le concept d'une classe d'objets exprime un ensemble de propositions théoriques à propos de cette classe. Une proposition est théorique dans la mesure où elle peut servir de prémisse dans des *explications* des propriétés possédées par le membre de cette classe (Murphy et Medin, 1985, Keil, 1989, Keil et Wilson, 1999).

A quoi ressemblent ces propositions théoriques ? Des généralisations nomologiques, causales, génériques et/ou fonctionnelles semblent pouvoir figurer dans des explications. En conséquence, un concept théorique est supposé encoder des généralisations nomologiques, causales, génériques, et/ou fonctionnelles qui expliquent les propriétés des membres de son extension. Par exemple, posséder un concept de chien, c'est posséder une espèce de théorie à propos des chiens, c'est-à-dire avoir un ensemble de connaissances nomologiques, causales, génériques et/ou fonctionnelles à propos des chiens, par exemple la connaissance que les chiens ont un pelage pour se protéger du froid etc.

Dans la section suivante, je montre qu'il faut admettre ces trois types de concepts.

4. LA COMPOSITION DES CONCEPTS

Malgré les nombreux désaccords qui existent à propos des concepts, philosophes et psychologues s'accordent pour soutenir que les concepts se composent. Depuis l'article fondateur de Osherson et Smith (1981), les psychologues cognitifs ont beaucoup travaillé sur ce sujet. Dans cette section, je montre que les travaux empiriques sur cette capacité cognitive viennent à l'appui de la prémisse 3 de l'ANE.

4.1. QU'EST-CE QUE LA COMPOSITION DES CONCEPTS ?

La composition des concepts est la capacité de former des représentations de classes pour lesquelles nous n'avons pas de représentations stockées en MLT. Ces représentations sont formées sur la base des représentations stockées en MLT. En d'autres termes, il y a deux types de concepts, les concepts complexes et les concepts simples. Par définition, un concept est simple si et seulement si c'est une représentation permanente en MLT.⁶ Il y a

⁶ Qu'en est-il alors de concepts comme FORFAIT HEBDOMADAIRE qui sont stockés en MLT ? Je soutiens que ce sont des concepts simples. En effet, nous n'avons pas à composer des concepts pour les activer en mémoire de travail.

un stock fini de concepts simples. Par exemple, « chien » exprime un concept simple (ou plus probablement plusieurs), puisque j'ai une (ou plusieurs) représentation(s) permanente(s) de chien en MLT. Un concept complexe n'est pas stocké de manière permanente en MLT, mais doit être produit de manière instantanée pour différentes finalités cognitives (par exemple, la compréhension linguistique, le raisonnement contrefactuel...). Il existe un mécanisme qui opère sur les concepts primitifs afin de produire les concepts complexes. Par exemple, si CHIEN et FOU sont deux concepts simples, il existe un mécanisme qui prend CHIEN et FOU pour inputs et produit le concept complexe CHIEN FOU.

Etant donné cette manière de caractériser la composition des concepts, l'objet d'une théorie de la composition des concepts est de déterminer *comment un concept complexe en vient à encoder les connaissances qu'il encode*. Par exemple, elle doit être en mesure d'expliquer comment le concept CHIEN FOU encode l'ensemble des connaissances à propos des chiens fous qu'il encode en fait.

4.2. DES PROTOTYPES SONT DES INPUTS DU PROCESSUS DE COMPOSITION

Dans certaines expériences, on demande aux sujets de faire la liste des propriétés qui sont typiques des deux catégories originelles et la liste des propriétés représentées par le concept complexe (Hampton, 1987, 1996, Kunda et al., 1990, Johnson et Keil, 2000). On pourrait par exemple demander aux sujets de faire la liste des étudiants d'Oxford, des maçons et des étudiants d'Oxford qui sont des maçons. Le psychologue James Hampton a montré qu'il y a une corrélation entre les propriétés qui sont jugées être typiques des membres des catégories originelles, par exemple des étudiants d'Oxford et des maçons, et les propriétés qui sont jugées typiques des membres de la catégorie complexe nouvelle, par exemple des étudiants d'Oxford qui sont des maçons (Hampton, 1987, 1996).

Pour expliquer cette observation, on a proposé le mécanisme suivant. Nous avons en MLT des représentations qui encodent des connaissances sur les propriétés qui sont jugées être typiques de leurs catégories, à savoir des *prototypes*. Quand nous produisons un

concept complexe, nous extrayons ces prototypes de notre MLT et nous transmettons certaines des propriétés qu'ils représentent au concept complexe. Produire un concept complexe consiste (entre autres, comme nous allons le voir) à déterminer quelles propriétés sont typiques de la catégorie dénotée par le concept complexe, dans notre exemple des étudiants d'Oxford qui sont des maçons, sur la base des prototypes des catégories originelles, dans notre exemple des étudiants d'Oxford et des maçons. Ce processus est connu sous le nom de *transmission des propriétés* (TP).

La transmission des propriétés au concept complexe est déterminée par le principe que les propriétés typiques des membres des catégories dénotées par les concepts simples sont aussi typiques des membres de la catégorie dénotée par le concept complexe. Cela soulève bien sûr la question suivante : *parmi les propriétés typiques, lesquelles sont transmises ?* Comme nous allons le voir ci-dessous, la question est complexe. Néanmoins, les résultats de Hampton montrent que le processus de composition suit, entre autres, *la règle de la typicalité élevée* : seules les propriétés qui sont *les plus typiques* des membres des catégories composées sont transmises au concept complexe (Hampton, 1987).

Soulignons que je ne soutiens pas que les concepts complexes sont des prototypes. Bien que j'affirme qu'ils encodent des connaissances sur les propriétés typiques des catégories dénotées, comme le font aussi les prototypes, ils ont des propriétés spécifiques. En particulier, ils encodent des informations modales et causales – comme nous allons le voir ci-dessous. En conséquence, ce ne sont pas des prototypes.

4.3. DES THÉORIES SONT DES INPUTS DU PROCESSUS DE COMPOSITION

La description du processus de composition proposée ci-dessus est incomplète. Si la TP était seulement gouvernée par la règle de la typicalité élevée, les instances typiques des concepts complexes seraient des membres typiques des catégories combinées. Or cela n'est pas toujours le cas. Par exemple, un étudiant d'Oxford qui est

un maçon est sans doute une instance atypique de la catégorie des étudiants d'Oxford et de la catégorie des maçons.

Cela suggère donc que la TP n'est pas seulement gouvernée par la typicalité des propriétés. D'autres considérations interviennent pour déterminer quelles propriétés sont représentées par les concepts complexes. Nous utilisons aussi *nos connaissances modales à propos des catégories originelles* pour sélectionner les propriétés *nécessaires*.

Hampton a montré que quand nous croyons qu'une propriété P est nécessairement possédée par les membres d'une catégorie C dénotée par un ou plusieurs concepts, nous transmettons toujours P aux concepts complexes produits à partir des concepts de C (Hampton, 1987). Par exemple, si nous croyons que les grand-mères sont nécessairement des mères de parents, la propriété *être une mère de parent* est toujours transmise aux concepts complexes produits à partir des concepts de grand-mères, par exemple au concept de grand-mère espion.

Or, les prototypes n'encodent pas des connaissances sur les propriétés nécessaires ou impossibles des membres des classes dénotées, mais seulement sur leur fréquence. En d'autres termes, les prototypes n'encodent pas de connaissances modales, mais seulement des connaissances statistiques. En revanche, les concepts théoriques sont précisément supposés encoder ce type de connaissances. Cela suggère que quand un prototype des A et un prototype des B sont composés, des concepts théoriques des A et des B sont aussi impliqués pour produire le concept complexe : ils sélectionnent les propriétés qui sont nécessaires.⁷

⁷ Un argument analogue montre que nous utilisons des connaissances causales pour produire un concept complexe (Medin et Shoben, 1988, Murphy, 1988, 1990, Johnson et Keil, 2000). Comme les prototypes ne sont pas supposés encoder ce type de connaissances, cela montre aussi qu'il faut postuler des concepts théoriques.

4.4. DES EXEMPLAIRES SONT DES INPUTS DU PROCESSUS DE COMPOSITION

Quand nous produisons un concept complexe à partir de deux prototypes, certaines connaissances qui sont encodées par le concept complexe ne sont pas héritées des prototypes composés (Hampton, 1987, 1997, Kunda et al., 1990, Johnson et Keil, 2000). Ce phénomène est appelé l'émergence des propriétés. Par exemple, la propriété *généreux* est jugée être typique des étudiants d'Oxford qui sont des maçons, bien que ni les étudiants d'Oxford ni les maçons ne soient jugés être généreux.

Les propriétés émergentes proviennent souvent de représentations d'individus singuliers. Hampton a appelé ce phénomène le feedback extensionnel (Hampton, 1988). Par exemple, il est probable que l'expression « étudiant de l'ENS qui est un O.S. » n'exprime pas un concept simple. Un concept complexe doit donc être produit à partir (entre autres) des prototypes des étudiants de l'ENS et des O.S.. Cependant, si je devais faire la liste des propriétés des étudiants de l'ENS qui sont des O.S., je mentionnerais des propriétés que je ne mentionnerais ni pour les étudiants de l'ENS ni pour les O.S.. L'explication est sans doute la suivante : je connais un étudiant de l'ENS qui est un O.S., à savoir Robert Linhart (voir *L'Etabli*). J'ai donc en MLT une représentation singulière de Linhart. Quand je construis le concept complexe, j'extrais de ma MLT cette représentation singulière et j'encode dans le concept complexe les propriétés représentées par la représentation singulière de Linhart, par exemple l'engagement à l'extrême gauche.

Le fait que nous extrayons certaines représentations singulières de notre MLT pour construire un concept complexe suggère que des représentations singulières sont disponibles pour être utilisées dans nos processus conceptuels supérieurs, autrement dit que nous possédons des exemplaires. Dans notre exemple, la représentation singulière de Linhart sert d'exemplaire de la catégorie des étudiants de l'ENS qui sont des O.S. : les autres membres de la catégorie sont supposés ressembler à Linhart.

4.5. RÉSUMÉ

J'ai montré ci-dessous que le processus de composition des concepts repose sur différents types de représentations mentales, à savoir des prototypes, des exemplaires et des théories. Cela vient bien sûr à l'appui de la prémisse 3 de l'ANE. Les données empiriques montrent en effet que la composition des concepts fait appel à trois types différents de connaissances : des connaissances sur les propriétés typiques des catégories, des connaissances modales et causales et des connaissances sur les membres particuliers de ces catégories. Les prototypes, les exemplaires et les théories sont précisément supposés encoder ces connaissances. Pour expliquer la composition des concepts, il faut donc postuler ces trois types de concepts. C'est ce que soutient la prémisse 3.

CONCLUSION

Il est sans doute utile de reformuler l'argument développé dans cet article. Les trois théories psychologiques des concepts postulent l'existence de trois types de représentations mentales, les exemplaires, les théories et les prototypes. J'ai admis que ces entités théoriques avaient peu en commun. La recherche sur une compétence psychologique importante, à savoir la composition des concepts, montre qu'il faut postuler l'existence de ces trois types de représentations mentales (pour des données empiriques concernant d'autres compétences cognitives supérieures, voir Machery ms). Par conséquent, les représentations mentales qui figurent dans les processus cognitifs supérieurs ont peu en commun. Même si l'on admet la notion la plus vaste d'espèce naturelle, à savoir la notion causale, les concepts ne forment donc pas une espèce naturelle.

Les conséquences de cette position sont importantes. Tout d'abord, les psychologues des concepts et des fonctions cognitives supérieures opposent depuis plus de trente ans leur théorie des concepts préférée aux théories défendues par leurs adversaires. *L'argument développé dans cet article montre que c'est une erreur.* Les concepts ne sont ni des prototypes, ni des exemplaires ni des théories. Certains concepts sont des prototypes, d'autres sont des

exemplaires, d'autres encore des théories. Les différentes théories des concepts ne sont donc pas contradictoires. Elles portent en fait sur différentes classes de représentations mentales. *Il est temps que les psychologues mettent fin à des controverses qui n'ont guère lieu d'être.*

Deuxièmement, la notion de concept est inutile pour formuler des généralisations empiriques à propos des représentations mentales. C'est une manière inadéquate de classer nos états mentaux si nous cherchons à étudier scientifiquement la cognition humaine. Au lieu de chercher à formuler des généralisations sur la classe des représentations mentales qui figurent dans les processus cognitifs supérieurs, en d'autres termes sur les concepts, il vaut mieux chercher à étudier les propriétés d'autres classes de représentations mentales, à savoir les prototypes, les exemplaires et les théories (et, éventuellement, d'autres classes encore). Les chercheurs en sciences cognitives et les philosophes de la psychologie devraient enfin se résoudre à renoncer à la notion de concept. *La psychologie scientifique n'a guère besoin de cette vieille lune.*

RÉFÉRENCES BIBLIOGRAPHIQUES

- BOYD, R.
1990 « What Realism implies and what it does not », dans *Dialectica*, vol. 43, pp. 5-29.
- 1991 « Realism, Anti-foundationalism and the Enthusiasm for Natural Kinds », dans *Philosophical Studies*, vol. 61, 1-2, pp. 127-148.
- BROOKS, L.
1978 « Nonanalytic concept formation and memory for instances », dans E. Rosch et B. B. Lloyd (dir), 1978.
- CAREY, S.
1985 *Conceptual Change in Childhood*, MIT Press, Cambridge.
- FODOR, J.A.
1974 « Special sciences », dans *Synthese*, vol. 28, pp. 97-115.
- 1994 « Concepts: A potboiler ». dans *Cognition*, vol. 50, 1, 95-113.
- 1998 *Concepts: Where Cognitive Science Went Wrong*, Oxford University Press, New York.
- GRIFFITHS, P.E.
1996 *What Emotions Really Are*, Chicago University Press, Chicago.
- HACKING, I.
1991 « A tradition of natural kinds », dans *Philosophical Studies*, vol. 61, pp. 109-126.

HAMPTON, J.A.

1987 « Inheritance of attributes in natural concept conjunctions », dans *Memory and Cognition*, vol. 15, pp. 55-71.

1988 « Overextension of conjunctive concepts: Evidence for a unitary model of concept typicality and class inclusion », dans *Journal of Experimental Psychology/ Learning, Memory, and Cognition*, vol. 14, pp. 12-32.

1996 « Conjunctions of visually-based categories: overextension and compensation », dans *Journal of Experimental Psychology: Learning, Memory and Cognition*, vol. 22, pp. 378-396.

1997 « Conceptual combination », dans K. Lamberts and D. Shanks (dir.), 1997.

JOHNSON, C., et KEIL, F.C.

2000 « Explanatory understanding and conceptual combination », dans F. C. Keil and R. A. Wilson (dir.), 2000.

KEIL, F.C.

1989 *Concepts, Kinds, and Cognitive Development*, MIT Press, Cambridge.

KEIL, F.C., et WILSON, R.A. (dir.)

2000 *Explanation and Cognition*, MIT Press, Cambridge.

KUNDA, Z., MILLER, D.T., et CLAIRE, T.

1990 « Combining social concepts: The role of causal reasoning », *Cognitive Science*, vol. 14, pp. 551-577.

- LAMBERTS, K., et SHANKS, D., (dir.)
 1997 *Knowledge, Concepts and Categories*, MIT Press, Cambridge.
- LAURENCE, S., et MARGOLIS, E.
 1999 « Concepts and cognitive science », dans S. Laurence et E. Margolis (dir.), *Concepts, Core Readings*, MIT Press, Cambridge.
- MACHERY, E. ms
 2004 *Les Concepts ne Forment pas une Espèce Naturelle : la Notion de Concept en Sciences Cognitives*, Thèse de doctorat, Université de Paris-Sorbonne.
- MEDIN, D.L., et SHAFFER, M.M.
 1978 « Context theory of classification learning », dans *Psychological Review*, vol. 85, pp. 207-238.
- MEDIN, D.L., et SHOBEN, E.J.
 1988 « Context and structure in conceptual combination », dans *Cognitive Psychology*, vol. 20, pp. 158-190.
- MILL, J.S.
 1905 *A System of Logic*, Routledge, Londres.
- MURPHY, G.L.
 1988 « Comprehending complex concepts », dans *Cognitive Science*, vol. 12, pp. 529-562.
 1990 « Noun-phrase interpretation and conceptual combination », dans *Journal of Memory and Language*, vol. 29, pp. 259-288.
 2002 *The Big Book of Concepts*, MIT Press, Cambridge.

- MURPHY, G.L., et MEDIN, D.L.
1985 « The role of theories in conceptual coherence », dans *Psychological Review*, vol. 92, pp. 289-316.
- OSHERSON, D.N., et SMITH, E.E.
1981 « On the Adequacy of Prototype Theory as a Theory of Concepts », dans *Cognition*, vol. 9, pp. 35-58.
- POSNER, M.I., et KEELE, S.W.
1968 « On the genesis of abstract ideas », dans *Journal of Experimental Psychology*, vol. 77, pp. 353-363.
- PRINZ, J.J.
2001 *Furnishing the mind: Concepts and their perceptual basis*, MIT Press, Cambridge.
- QUINE, W.V.O.
1969 « Natural Kinds », dans W.V.O. Quine, *Ontological Relativity and Other Essays*, Columbia University Press, New York.
- ROSCH, E.
1978 « Principles of categorization », dans E. Rosch et B. B. Lloyd (dir.), 1978.
- ROSCH, E., et LLOYD, B.B. (dir.)
1978 *Cognition and Categorization*, Erlbaum, Hillsdale.
- ROSCH, E., et MERVIS, C.B.
1975 « Family resemblance : Studies in the internal structure of categories », dans *Cognitive Psychology*, vol. 7, pp. 573-605.

SMITH, E.E., et MEDIN, D.L.
1981 *Categories and Concepts*, Harvard University
Press, Cambridge.

SMITH, E.E., PATALANO, A.L., et JONIDES, J.
1996 « Alternative strategies of categorization », dans
Cognition, vol. 65, pp. 167-196.