

ABSTRACTION : DES EXEMPLAIRES AUX REGLES

*Sabine Gelaes
Jean-Pierre Thibaut*

A quoi servent les concepts ? Ils sont une représentation du monde construite par les individus. Ils organisent un matériau qui sans eux ne serait qu'un chaos (Smith et Medin, 1981 ; Murphy, 2002 ; Cordier, ce volume). Classiquement, cette organisation repose sur les notions de similarité et de taxonomie. Similarité, car les concepts sont des représentations mentales qui spécifient en quels termes, selon quelles dimensions, les entités, les événements d'une même catégorie sont semblables. Taxonomie, car ils reflètent les relations d'inclusion reliant ou non les catégories. Par exemple, un caniche est une sous-catégorie de la catégorie des chiens, elle-même incluse dans celle des mammifères, cette dernière partie de la classe des animaux. Quels rapports entretiennent les concepts avec les entités qu'ils permettent de décrire, de classer ? Pour mieux cerner cette question, partons d'une situation où un enfant doit comprendre ce qu'est un chien. Au total, cette compréhension se marquera par sa capacité à distinguer les chiens du reste du monde, particulièrement des autres mammifères quadrupèdes, et par l'acquisition de connaissances pertinentes associées à ces animaux. A une extrémité, en termes de cet exemple, on postulera que le concept de chien est encodé sous la forme des exemplaires individuels, des expériences particulières au sujet des chiens. A l'autre extrémité, on considère que le concept de chien transcende les situations dont il est extrait pour devenir une représentation abstraite, voire amodale des chiens (Anderson, 1978 ; Pylyshyn, 1981).

Concepts et abstraction

Dans de nombreux modèles, les concepts sont décrits comme des ensembles structurés de composants, les traits ou dimensions (Smith et

Medin, 1981; Lamberts et Shanks, 1997; Murphy, 2002 pour des revues de la littérature). Ils permettent les inférences et les combinaisons conceptuelles. Dans les modèles dits classiques, les concepts sont des résumés dans lesquels l'abstraction est maximale. Les traits qui composent ces concepts sont considérés nécessaires et suffisants. Les traits constituent des règles qui, une fois vérifiées, permettent de classer tout nouveau stimulus dans une catégorie, sans ambiguïté. Dans les modèles structurés autour de la notion de prototype, les traits sont des abstractions de propriétés perceptives, contextuelles, et fonctionnelles. Chaque trait est affecté d'une probabilité qui définit dans quelle mesure un trait est présent dans les membres de la catégorie. Par exemple, le trait « vole » est affecté d'une probabilité importante dans le concept d'oiseau, le trait « chante » est affecté d'une autre probabilité. Même si dans le cadre de ces modèles, les concepts ne sont plus définis par des traits nécessaires et suffisants et que l'appartenance à une catégorie peut être matière ambiguë, les traits qui les composent sont des propriétés abstraites qui transcendent les concepts. Par exemple, pour étudier la composition des concepts, on demande à de sujets de lister les traits qui composent tel ou tel concept. Si les sujets citent « a des ailes » pour oiseau et avion, cela signifie que ces deux classes d'entités partagent un trait (Rosch, 1978; Cordier, ce volume).

Dans les modèles à exemplaires, les concepts sont représentés sous la forme des exemplaires rencontrés par le sujet. Le concept d'oiseau n'est rien d'autre que la représentation des oiseaux, ou un sous-ensemble de ceux-ci, rencontrés et encodés par le sujet. Ces modèles ne sont pas dépourvus d'abstraction, notamment pour les dimensions qui décrivent chaque exemplaire. En effet, un exemplaire ne peut pas être représenté selon toutes ses dimensions, surtout dans le cas des stimuli de la réalité qui peuvent se décrire en termes d'une infinité de dimensions. On voit mal comment l'individu pourrait encoder et stocker les différents aspects de tous les exemplaires de chien rencontrés. Dans la plupart des expériences, les exemplaires proposés se définissent sur un ensemble limité de dimensions. Par exemple, des stimuli définis par une forme (carré ou triangle) une taille (grand ou petit), une texture, une couleur, ou des visages simplifiés où chaque trait prend quelques valeurs seulement.

La plupart des modèles de représentations des concepts adoptent une perspective compositionnelle (ou décompositionnelle) dans laquelle les concepts sont représentés sous la forme de traits, de dimensions. Ici aussi, plusieurs conceptions des traits ont été développées. Elles se situent entre deux pôles. D'un côté, et c'est la

perspective la plus répandue, une conception globale, abstraite dans laquelle un trait est un symbole « général » utilisable dans tous les contextes. Les avantages d'une telle conception sont multiples, la simplicité, la généralité, l'indépendance contextuelle. La simplicité provient de l'utilisation de la même unité dans toutes les représentations où elle intervient, ce qui lui donne aussi une grande généralité et une grande indépendance par rapport au contexte. Par exemple, des pieds de table, de chaise, d'animal, d'humain sont des instances d'une même entité abstraite, le pied, qui est utilisable dans tous ces contextes différents. Remarquons que ces traits généraux, abstraits sont également au centre des modèles à exemplaires (Smith et Medin, 1981). Un exemplaire d'oiseau serait représenté comme une configuration particulière de dimensions (par exemple, par des propriétés communes à tous les oiseaux, comme avoir des ailes, mais aussi des propriétés comme « est petit », « est rouge », etc.).

A l'opposé, la conception localiste considère qu'une même propriété est représentée différemment dans les différentes entités auxquelles elle participe. Dans l'exemple donné ci-dessus, les diverses représentations du pied (de table, de chaise, d'humain) sont différentes et ne sont donc plus interchangeables (Medin et Shoben, 1988; Wisniewski, 1998). Au lieu d'un seul symbole qui représenterait une propriété apparaissant dans une multitude de situations différentes (par exemple, toutes les entités qui comportent un pied), on aurait un grand nombre de symboles qui représente chacun une forme spécifique de la propriété (par exemple, les différentes formes de pieds). Dans ce cadre localiste, on reconnaît que la segmentation des entités par le langage ne correspond pas nécessairement aux représentations construites par les sujets. Ainsi si le mot « pied » désigne de manière commode un trait important des sièges, il n'est pas évident que le sujet ne possède qu'une seule représentation abstraite de la propriété.

Dans un travail visant à départager ces deux positions, Solomon et Barsalou (2001) prédisent que si toute propriété est représentée de manière générale, sa vérification pour un concept devrait faciliter la vérification de cette propriété dans tout autre concept, alors que si elle est représentée localement, elle ne devrait pas faciliter sa vérification sous d'autres formes. Les auteurs utilisent le paradigme d'amorçage dans une tâche de vérification de propriétés. Ils demandent à des sujets, par exemple, de vérifier le plus rapidement possible si un poney a une crinière. Juste avant cette vérification, ils ont demandé au sujet de vérifier si soit un cheval soit un lion a une crinière. Les auteurs ont ensuite comparé le temps de réaction et le nombre d'erreurs pour la

vérification de poney-crinière selon qu'elle est précédée de la vérification cheval-crinière ou par la vérification lion-crinière. Selon les auteurs, si les propriétés sont représentées localement, la vérification de « crinière » pour cheval devrait amorcer davantage la vérification de « crinière » pour poney que la vérification préalable de « crinière » pour lion, puisque la crinière d'un cheval est très semblable à celle d'un poney. Lorsque le sujet doit vérifier la propriété crinière pour poney, la représentation spécifique activée pour cheval est utilisable pour crinière-poney mais pas pour crinière-lion. Au contraire, si les propriétés sont représentées globalement, toute activation préalable d'une crinière (d'un cheval ou d'un lion) devrait faciliter la vérification ultérieure de la présence de cette propriété dans n'importe quelle situation. Les résultats confirment l'hypothèse localiste avec un nombre d'erreurs et des temps de réaction plus élevés lorsque la vérification poney-crinière était précédée par la vérification lion-crinière.

Règles et exemplaires : Abstraction et instantiation

Dans de nombreux travaux sur la catégorisation, particulièrement sur l'apprentissage conceptuel, on présente une série de stimuli, généralement multidimensionnels, que le sujet doit apprendre à catégoriser. Le sujet cherche les traits qui unifient les catégories, qu'il s'agisse de traits nécessaires et suffisants ou, plus souvent, d'un ensemble de traits qui donnent aux stimuli un air de famille. Dans ce dernier cas, un exemplaire particulier peut appartenir à la catégorie par certains traits alors qu'un autre en fait partie par des traits en partie différents.

On peut aussi donner la règle de catégorisation au sujet, puis analyser comment l'utilisation répétée de cette règle en rend l'utilisation indépendante des aspects idiosyncrasiques des stimuli. L'énoncé de cette règle est abstrait dans le sens où les traits sont formulés de manière générale. Dans le cas des stimuli de la Figure 1, la règle pourrait être définie sur les dimensions « taille du corps », « présence ou absence de points sur le corps » « forme du corps ». Par exemple, si l'animal a deux des trois propriétés suivantes, « a un corps de petite taille, a des points sur le corps, a six pattes » il faut le classer en A, sinon il faut le classer en B. La règle s'applique à toutes les formes de points ou de corps de petite taille, à tous les groupes de six pattes, quelles que soient les variations introduites (à condition que ces caractéristiques soient perceptibles, bien entendu). En d'autres termes, si le sujet doit apprendre à classer les animaux de la Figure 1 le plus

rapidement possible selon une règle simple et énoncée de manière abstraite, il devrait rapidement apprendre à l'appliquer de manière automatique. Après une période d'apprentissage durant laquelle il apprend à rechercher les dimensions cibles dans les stimuli et à les reconnaître, les performances devraient être identiques quels que soient les exemplaires présentés. L'hypothèse opposée serait que les caractéristiques des exemplaires rencontrés pendant l'apprentissage influencent la catégorisation d'un nouvel item. Dans ce cas, les dimensions apprises ne sont pas totalement abstraites et la classification d'un item particulier est influencée par les exemplaires rencontrés durant l'apprentissage.

Pour étudier cette question, Allen et Brooks (1991 ; voir aussi Regelir et Brooks, 1993) ont conçu un paradigme intéressant. Ils montrent que la classification de nouveaux items (dits de transfert) suivant une règle automatisée est influencée par la similarité de ces nouveaux items avec les items ayant servi à l'apprentissage de la règle de classification. Pendant la phase d'entraînement, les sujets apprennent à classer des stimuli selon une règle simple et parfaitement prédictive (par exemple, pour les stimuli de la Figure 1, une règle était : « Si l'animal possède au moins deux parmi les trois caractéristiques suivantes – corps arrondi, six pattes, avoir des points –, il s'agit d'un X, sinon il s'agit d'un Y »). Durant la phase de test qui suit cette phase d'apprentissage, les sujets doivent classer les stimuli d'apprentissage et de nouveaux stimuli selon la même règle. Les nouveaux stimuli de transfert sont identiques à ceux vus durant l'apprentissage à une différence près. Sur une des trois dimensions, ils prennent systématiquement la valeur opposée à celle de leur « jumeau » (par exemple, un stimulus d'entraînement avec des points sur le corps sera présenté sans ces points à la phase de transfert, et vice-versa). Par exemple, dans la Figure 1, les stimuli de transfert 1B et 2B sont le résultat de la transformation « points » des stimuli d'apprentissage 1A et 2A. Cette transformation des stimuli d'entraînement donne lieu à deux types de stimuli de transfert, les « bons » et les « mauvais » stimuli de transfert. Les « bons » stimuli de transfert appartiennent à la même catégorie que leur « jumeau » d'entraînement en termes de la même catégorisation alors que les « mauvais » stimuli de transfert, selon cette règle de catégorisation, appartiennent à la catégorie opposée à celle de leur « jumeau » d'entraînement. Dans la Figure 1, le stimulus 1B est un « bon » transfert alors que le stimulus 2B est un mauvais transfert. En résumé, les bons et mauvais stimuli de transfert ont la même relation à l'item d'entraînement dont ils dérivent en termes de similarité (ils sont

identiques à leur item jumeau à un trait près), mais ils se différencient en termes d'appartenance catégorie (l'un appartient à la même catégorie que son jumeau, l'autre pas).

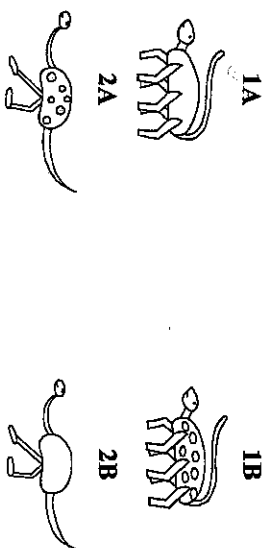


Figure 1 : stimuli utilisés par Regier et Brooks (1993), Thibaut et Gelaes.

Puisque la règle donnée aux sujets est parfaitement prédictive, on ne devrait pas observer de différences dans la classification des deux types de stimuli de transfert. Cependant, Brooks et collègues ont montré à plusieurs reprises que des sujets commettent plus d'erreurs et présentent des temps de réaction plus longs pour les « mauvais » stimuli de transfert que pour les « bons ». Il s'agit d'un effet d'exemplaire, puisque les propriétés individuelles des stimuli rencontrés durant l'apprentissage influencent la catégorisation des items de transfert qui, par définition, n'ont jamais été rencontrés auparavant.

Thibaut et Gelaes (soumis) ont étudié l'évolution de l'effet des exemplaires dans le cadre du paradigme de Brooks et collaborateurs, en augmentant le nombre d'essais d'entraînement et le nombre de stimuli présentés durant l'apprentissage. L'augmentation du nombre d'essais d'apprentissage peut, *a priori*, entraîner deux effets opposés. Premièrement, avec la pratique, la règle s'automatise et l'utilisation des traits pertinents devient de plus en plus efficace. En conséquence, l'influence de la similarité devrait diminuer à mesure que l'expertise avec la règle s'accroît. À l'inverse, l'augmentation du nombre d'essais pourrait renforcer l'influence des exemplaires car les caractéristiques idiosyncrasiques des exemplaires devraient être encodées de plus en plus distinctement avec les essais. L'effet de l'augmentation du nombre d'essais d'apprentissage pourrait être modulé par le nombre d'exemplaires présentés à l'apprentissage (restreint vs. important). L'emploi d'un petit nombre d'exemplaires pourrait favoriser l'influence des exemplaires. Au contraire, l'utilisation d'un grand

nombre d'exemplaires durant l'apprentissage devrait avoir un impact négatif sur l'influence de la similarité car le nombre et la variabilité des stimuli devraient les rendre moins discriminables et permettre une expression « pure » de la règle.

Dans les différentes conditions que nous avons étudiées (8 stimuli d'entraînement présentés 5, 12 et 30 fois et 20 stimuli d'entraînement présentés 2, 5 et 12 fois), l'analyse des résultats ne révèle, contrairement à ce qu'avait obtenu Brooks et ses collaborateurs, qu'un effet d'exemplaires très restreint lorsque les stimuli d'apprentissage sont peu nombreux et présentés 30 fois. Cette discordance entre nos résultats et ceux obtenus par Brooks et collaborateurs s'explique par la présence d'un biais systématique non contrôlé dans les expériences de ces derniers.

En outre, selon Regier et Brooks (1993), l'effet des exemplaires ne se manifeste que pour des stimuli fortement individualisés, c'est-à-dire qui diffèrent fortement les uns des autres de manière globale. Cependant, chacune des dimensions qui composent les stimuli peut contribuer individuellement à l'effet des exemplaires, les sujets encodant probablement les aspects idiosyncrasiques de chaque instance des dimensions composant un stimulus. Contrairement à la position de Regier et Brooks, les exemplaires individuels des dimensions pourraient influencer la classification même si les stimuli sont présentés sous la forme de traits séparés les uns des autres. Afin de tester cette hypothèse, nous avons présenté les stimuli avec leurs dimensions constitutives isolées les unes des autres (la tête séparée du corps, séparé des pattes, etc.). Si les dimensions sont encodées séparément, les sujets qui reconnaissent les traits individuels dans la phase de transfert devraient classer les « bons » stimuli de transfert plus rapidement et/ou avec plus d'exactitude que les « mauvais » stimuli de transfert. L'analyse des erreurs montre qu'elles sont plus nombreuses pour les mauvais transferts que pour les bons. Ce résultat suggère que l'effet des exemplaires est pris en charge, en partie, par les dimensions isolées.

En conclusion, non seulement les sujets apprennent à utiliser la règle de classification de plus en plus efficacement, mais encodent aussi les traits individuels des stimuli. Cependant, contrairement aux hypothèses de Regier et Brooks, l'influence des exemplaires d'apprentissage s'exerce par l'intermédiaire des dimensions des stimuli prises isolément plutôt que par l'intermédiaire des stimuli pris globalement. Nous pensons, et des données préliminaires le suggèrent, que l'influence de la similarité globale entre exemplaires d'apprentissage et de test ne se manifeste que lorsque la dimension de

catégorisation modifiée entre l'apprentissage et le test affecte peu leur ressemblance. Dans ce cas, un exemplaire d'apprentissage et un « bon » test sont difficiles à distinguer sur la base de leur similarité globale alors qu'ils appartiennent à la même catégorie. Un « mauvais » exemplaire de test est, lui aussi, globalement très semblable à l'exemplaire d'apprentissage dont il dérive alors qu'il en est très différent en termes de la règle puisqu'il n'appartient pas à la même catégorie.

Les propriétés corrélées à la règle et leur influence sur son utilisation

Lorsqu'une règle de classification simple est disponible, d'autres dimensions peuvent être corrélées plus ou moins parfaitement à cette règle de classification. Par exemple, une règle de classification stipule que tous les stimuli de couleur rouge appartiennent à une catégorie donnée et la majorité des stimuli rouges sont aussi de forme carrée. Dans quelle mesure cette corrélation entre les propriétés va-t-elle influencer la catégorisation ? Pour le savoir, il faudrait comparer les résultats obtenus, lors d'un test, pour de nouveaux items qui respectent cette corrélation et des items qui ne la respectent pas.

Une seconde question porte sur la prise de conscience que le sujet a de cette corrélation. Peut-il être influencé par cette corrélation sans en prendre conscience ? Peut-il l'apprendre sans la noter et celle-ci peut-elle influencer sa catégorisation de nouveaux stimuli ? Expérimentalement, cela signifierait, à nouveau, comparer les performances obtenues pour des stimuli qui confirment ou non la corrélation présente durant l'apprentissage. Dans l'affirmative, il s'agit d'un apprentissage implicite, c'est à dire l'apprentissage d'une régularité sans intention d'apprendre et sans prise de conscience de ce qui a été appris (voir French et Cleemans, 2002). Thibaut, Lemaire, Quadri (1998) ont réalisé de telles expériences. Dans une première phase, les sujets doivent découvrir une règle de classification simple. Par exemple les stimuli de la Figure 2 ont été conçus pour que la répartition des « pattes » dans l'espace fournisse une règle de classification parfaite (1 patte + 3 pattes pour la catégorie A, 2 pattes + 2 pattes pour la catégorie B). Pendant la phase de recherche de la règle de classification proprement dite, seule la répartition des pattes peut servir de critère de classification. Les autres régularités possibles sont réparties dans les deux catégories de manière équivalente et ne peuvent être utilisées comme règle de classification. Suit une phase d'automatisation pendant laquelle on associe aux stimuli de chaque catégorie une propriété de manière systématique, mais sans le signaler

aux participants. Par exemple, tous les stimuli de la catégorie A (1+3) ont été arrondis (propriété A') alors que ceux de la catégorie B (2+2) ont été allongés et écrasés verticalement (propriété B'). (Cette propriété, arrondi vs. allongement, si elle avait été détectée pendant la phase d'apprentissage, ne pouvait être utilisée pour catégoriser car elle était répartie dans les deux catégories.) Pendant l'automatisation, les sujets voient les stimuli plusieurs fois et ils doivent les classer le plus rapidement possible en termes de la règle de classification apprise à la phase précédente. Durant la phase de test qui suit, on présente des stimuli nouveaux qui respectent la corrélation proposée durant la phase d'automatisation (soit les associations A-A' et B-B' : ce sont les stimuli congruents) ou non (soit les associations A-B' et B-A' : ce sont les stimuli contradictoires). Après l'expérience, on vérifie à l'aide d'un questionnaire si les sujets ont noté l'association présente durant toute la phase d'automatisation entre les traits explicites de catégorisation (1+3 et 2+2) et les traits associés A' et B'. Les résultats indiquent que les temps de réaction aux stimuli congruents A-A' et B-B' sont plus courts que les temps pour les stimuli contradictoires A-B' et B-A'. Cette comparaison entre les deux types de stimuli est réalisée uniquement pour les sujets qui n'ont pas noté l'association entre A et A' et/ou B et B' durant la phase d'automatisation. Ces résultats indiquent que dans une situation où la règle de classification est très simple, parfaitement prédictive, et automatisée, des régularités associées à cette règle influencent la performance.

Lors du traitement et de l'automatisation de la règle de catégorisation, d'autres aspects des stimuli sont traités et encodés par les sujets. Dans le cas présent, les dimensions pertinentes de classification sont, en quelque sorte, coulées dans les dimensions associées. En effet, les propriétés associées utilisées ici sont globales dans le sens où elles affectent toutes les parties du stimulus. Ainsi, les pattes d'un stimulus 1+3 ne peuvent être traitées que comme des pattes arrondies et les pattes 2+2 que comme des pattes allongées. A mesure que l'automatisation se poursuit, la recherche de l'information « 1+3 = catégorie A et 2+2 = catégorie B » est de plus en plus associée respectivement aux propriétés « arrondie » et « allongée ». Lors du test, la rupture de cette association crée un nouveau contexte de recherche des propriétés, moins familier, qui entraîne un ralentissement du temps de réaction. Au total, en dehors d'une prise de conscience du sujet, d'autres propriétés que celles consciemment visées par le sujet, finissent par contrôler la prise de décision, et ce même pour des stimuli très simples, dont les règles de classification sont surapprises.

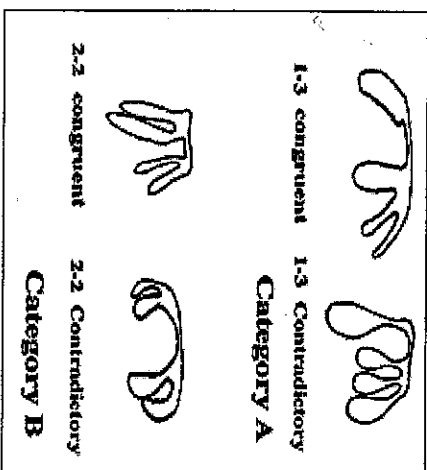


Figure 2: stimuli utilisés par Thibaut *et al.* (1998). Stimuli congruents et contradictoires pour les catégories A et B.

Nos résultats ont des implications sur le statut des tâches de listing de propriétés souvent utilisées en psychologie. Dans ces tâches, on demande aux sujets de citer les propriétés qui, selon eux, définissent une catégorie particulière. Outre les difficultés associées à l'utilisation de telles techniques pour étudier les concepts (Smith, 1989), nos résultats soulignent l'importance de propriétés dont les sujets n'ont pas conscience et qui pourtant influencent leurs classifications.

Plus généralement, si le langage permet une description des concepts que nous utilisons, ceux-ci ne se réduisent pas à cette seule description. Des propriétés non verbalisées et/ou difficilement verbalisables (e.g., propriétés holistiques émergentes) contribuent à la classification. Le langage est (permet) une description des concepts parmi d'autres.

Généralisation des règles et variabilité des instances

Par définition, une règle générale, un concept doit être productif dans le sens où il doit pouvoir être appliqué à des situations autres que celles dans lesquelles il a été abstrait. Sans cela il ne serait que la simple description sans grande généralité d'un ensemble de circonstances, d'entités, rencontrées dans le passé de l'individu. Cette nécessité de généralisation nous renvoie à variabilité des instances rencontrées durant l'apprentissage. Dans la plupart des cas, l'apprentissage se fait

sur un nombre limité d'instances qui ne peuvent représenter l'éventail de toutes les instances. Ainsi, le jeune enfant apprend de nouveaux mots, selon bien des auteurs, sur base d'une présentation d'un seul exemplaire. Par exemple, il doit appliquer le mot « chien » à tous les chiens alors que, dans bien des cas, la seule association entre « chien » et la catégorie correspondante se limite au seul chien de la maison. Il doit donc abstraire à partir de cette seule présentation. Les caractéristiques des entités qu'il devra dénommer « chien ». De nombreuses expériences montrent que l'enfant est capable de réaliser cette association très rapidement (Carey, 1978; Bloom, 2000). Cependant de nombreuses observations indiquent qu'il commet des erreurs de généralisation, soit en surgénéralisant (en dénommant « chien » d'autres animaux) ou en sous généralisant (en ne dénommant pas « chien » des entités que l'adulte appelle « chien »). Cette absence de généralisation n'est pas typique de l'enfant. Dans le cadre d'un test, de nombreux étudiants, confrontés à un nouveau concept (e.g., le concept de schizophrène) seront incapables de l'appliquer à de nouveaux cas. La raison n'est pas que l'enfant ou l'étudiant n'a pas compris ce qui lui a été présenté mais tout simplement qu'il est incapable de « voir la même chose dans quelque chose de différent ».

Thibaut (2000) a étudié des facteurs qui peuvent influencer la généralisation d'une règle de classification chez des enfants de 4 et 6 ans. L'expérience se faisait en deux phases, la première d'apprentissage, la seconde de généralisation. Deux conditions de généralisation étaient comparées, avec ou sans feedback. Le principe général de l'expérience est d'apprendre aux enfants à classer des stimuli, tels que ceux de la Figure 3, selon une règle de classification simple, à nouveau 1+3 ou 2+2 (voir stimuli Figure 3A). Ensuite, on procède à un nouvel apprentissage de la règle sur du matériel plus variable, comportant des traits saillants non pertinents (voir Figure 3B). Ce nouvel apprentissage est réalisé avec feedback ou sans feedback après chaque stimulus. Nous pensions que le feedback pourrait être nécessaire pour apprendre comment appliquer la règle connue à des stimuli dont la structure générale a été modifiée. Cela montrerait que la règle abstraite sur les items d'apprentissage n'a aucune généralité. Des expériences préliminaires ont montré que si on veut réaliser l'apprentissage de la règle directement sur les stimuli complexes (Figure 3B), les enfants échouent jusqu'à l'âge de 12 ans environ. Si le passage par les stimuli simples permet l'apprentissage de la règle complexe, on a trouvé une manière de favoriser l'apprentissage d'une notion dans ces acceptions plus complexes.

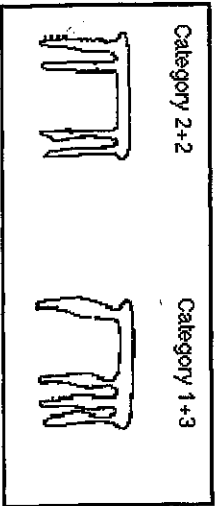


Figure 3a : deux stimuli simples

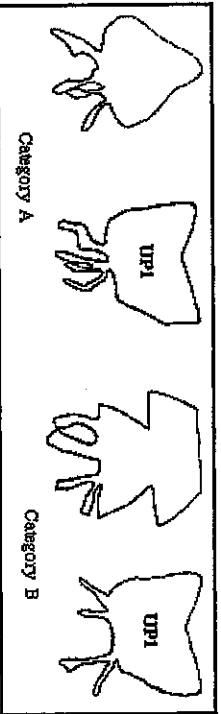


Figure 3b : Quatre stimuli complexes des catégories 1+3 et 2+2

On présente d'abord plusieurs fois les stimuli simples, un par un et on donne un feedback à chaque essai sur l'exactitude de la classification. Lorsque le sujet classe correctement les stimuli simples avec la règle, il passe à la phase de généralisation, avec ou sans feedback, sur les stimuli complexes. Dans le cas de l'absence de feedback, les stimuli sont présentés un par un et l'enfant ne reçoit pas de feedback sur la classification qu'il a proposée. Les résultats montrent que les enfants de six ans réussissent mieux le passage des items simples aux complexes. Ils montrent également que les enfants de 4 ans échouent davantage en l'absence de feedback qu'en présence d'un feedback. Cette différence n'existe pas à 6 ans. Ceci indique que la généralisation d'instances d'apprentissage à de nouveaux stimuli plus complexes ne se fait pas automatiquement chez les plus jeunes, particulièrement lorsque le feedback est absent. Chez ces derniers, ce feedback est nécessaire pour obtenir une généralisation de la règle. Dans nos termes, une fois celle-ci apprise, les enfants doivent apprendre à l'appliquer de manière systématique à des stimuli nouveaux. Si les enfants de six ans parviennent à généraliser la règle plus fréquemment y compris en l'absence de feedback, c'est probablement parce qu'ils parviennent à dépasser, ignorer les traits saillants non pertinents qui constituent les stimuli complexes.

Dans une seconde expérience, nous avons donné directement une règle de catégorisation aux enfants, règle qui pouvait s'appliquer (congruente) ou non (non-congruente) aux stimuli de généralisation. Les stimuli d'apprentissage pouvaient être décrits selon deux types de traits de catégorisation. Dans la condition règle congruente, on expliquait la règle 1+3 vs. 2+2 aux enfants pour les stimuli de la Figure 4 (voir stimuli A1 et A2), puis on leur présentait les stimuli B1 et B2 de la même Figure qui peuvent également être catégorisés selon cette règle. Dans la condition non-congruente, la règle qui était donnée dans la première phase ne pouvait pas s'appliquer aux stimuli de la deuxième phase. Par exemple, on décrivait les stimuli A1 et A2 comme des « fourchettes tordues » et des « croix », description non généralisable aux stimuli B1 et B2. Notez que les mêmes caractéristiques physiques des stimuli d'apprentissage sont interprétées différemment dans les deux conditions. En effet, lors du passage des stimuli d'apprentissage aux stimuli de généralisation, on dit aux enfants qu'ils vont voir de nouveaux exemplaires des catégories précédentes. Le but de l'expérience était de vérifier si des enfants qui apprennent à classer des stimuli dans la condition congruente pouvaient appliquer la règle aux nouveaux stimuli et, complémentarément, de vérifier si ceux qui ont appris à classer ces mêmes stimuli d'apprentissage dans la condition non-congruente seraient capables d'appliquer la règle.

Les résultats montrent que les enfants de 4 ans échouent le plus souvent à apprendre à classer les nouveaux stimuli de la Figure 4 (B1 et B2) dans la condition non congruente alors leur performance est meilleure dans la condition congruente. Dans le premier cas, lorsqu'on leur décrit les stimuli d'apprentissage comme des « croix » et des « fourchettes », ils sont probablement incapables de voir que ces stimuli pouvaient aussi se décrire comme 2+2 et 1+3. Ensuite, ils sont incapables de générer une nouvelle description des stimuli de transfert (stimuli B1 et B2) lorsque la description qui leur a été donnée durant l'apprentissage s'avère inutilisable pour ces stimuli. On notera que la plupart des enfants de 4 ans sont capables d'apprendre à classer des stimuli comme B1 et B2 lorsqu'ils sont présentés seuls, c'est-à-dire sans apprentissage préalable d'autres stimuli. L'incapacité à générer une nouvelle description des stimuli de transfert dans la condition non congruente s'explique sans doute par les faibles capacités de flexibilité mentale ou l'inhibition des enfants (Dempster, 1992; Kipp-Harnishfeger, 1995).

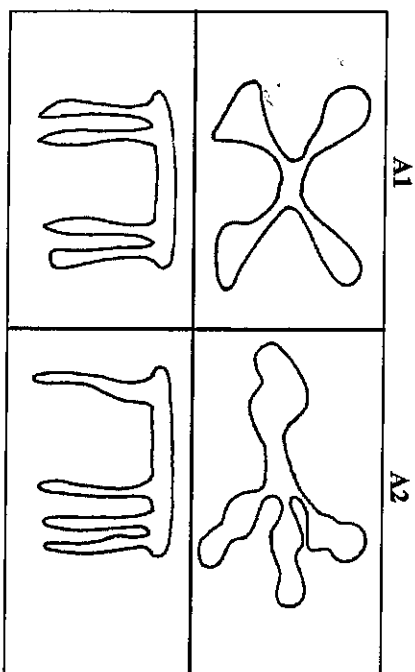


Figure 4 : les stimuli A1 et A2 servent à l'apprentissage et les stimuli B1 et B2 à la généralisation.

Conclusions

Les exemples que nous avons décrits montrent que la recherche d'une règle se fait souvent dans un espace limité. Par là, nous voulons dire que la règle utilisée n'est pas totalement abstraite. Nous avons vu avec Brooks et collègues que, même après une période d'automatisation relativement longue, la classification de nouveaux exemplaires dépend de leur similarité avec les exemplaires d'apprentissage. De la même manière, une règle de catégorisation surapprise est influencée par des dimensions corrélées alors que les sujets non pas conscience de cette corrélation. Dans notre interprétation, ces dimensions corrélées délimitent l'espace dans lequel la règle est traitée de façon optimale. Modifier cet espace affecte l'efficacité du traitement du stimulus. En d'autres termes, le traitement de la règle se fait nécessairement sur des stimuli particuliers. Les règles de catégorisation, même automatisées, se « souviennent » des situations (des exemplaires) dans lesquelles elles ont été apprises.

La catégorisation est influencée par des traits qui ne font pas partie des traits explicites par les individus. Les processus cognitifs impliqués dans l'apprentissage et l'utilisation d'une règle « créent » d'autres sources de catégorisation pertinentes que celles que le sujet pense

utiliser lors de ses catégorisations. Dès lors peut-on encore parler de règles de catégorisation, de concepts sous la forme d'un ensemble ou d'une liste de traits, comme c'est généralement le cas dans cette littérature. La réponse est positive dans la mesure où dans tous les exemples cités, les classifications des sujets étaient guidées par des règles. La réponse est négative si l'on ignore que ces traits n'épuisent pas toutes les sources des catégorisations effectuées.

Les règles de catégorisation, les concepts, sont des ensembles de traits qui permettent de calculer une similarité entre une entité à classer et des catégories potentielles (« sachant qu'une entité possède telles caractéristiques, elles peut appartenir à telle ou telle catégorie »). Les règles, les concepts définissent des fonctions pour calculer la similarité (e.g., via des exemplaires si on conçoit que les concepts sont représentés sous la forme d'exemplaires, via des prototypes si les concepts sont représentés selon des théories naïves, voir Cordier, ce volume) utilisées pour classer de nouvelles instances (voir cependant la notion de dissociation entre similarité et catégorisation, selon laquelle, dans certaines situations, une entité peut être jugée plus semblable à la catégorie A qu'à la catégorie B tout en étant classée dans la catégorie B plutôt que dans la catégorie A ; voir Rips, 1989, Ahn et Dennis, 2001, Thibaut, Dupont et Anselme, 2002). Ces fonctions de similarité ne se limitent pourtant pas aux traits, dimensions qui les définissent explicitement.

L'abstraction telle qu'elle est généralement conçue, n'est qu'une redescription commode d'une réalité. Au total, si on part de ces descriptions abstraites, il reste à savoir comment les individus les interprètent, les projettent sur la réalité. L'enfant qui apprend une règle de catégorisation, nous l'avons vu, ne verra pas nécessairement qu'elle s'applique aussi à des stimuli dont la structure est différente de celle qui prévalait lors de l'apprentissage ; l'étudiant en psychologie qui apprend un concept de ce domaine à partir d'une liste de traits ne parviendra peut-être pas à reconnaître des entités réelles auxquelles ce concept s'applique. Si les règles apprises sont souvent influencées par les contextes dans lesquels elles sont nées, de la même manière et complémentarément, il faut apprendre à traduire des règles abstraites dans la réalité.

Bibliographie

- Ahn, W. K., & Dennis, M. J., 2001: Dissociation between categorization and similarity judgement: Differential effect of causal status on feature weights in U. Hahn, & M. Ramscar eds, *Similarity and categorization*, 87-107, New-York, Oxford University Press.
- Allen, S. W., & Brooks, L. R., 1991: Specializing the operation of an explicit rule, *Journal of Experimental Psychology: General*, 120, 3-19.
- Anderson, J. R., 1978: Arguments concerning representations for mental imagery, *Psychological Review*, 85, 249-277.
- Barsalou, L., 1999: Perceptual symbol systems, *Behavioral and Brain Sciences*, 22, 577-660.
- Bloom, P., 2000: *How children learn the meanings of words*, Cambridge, MIT Press.
- Carey, S. & Bartlett, 1978: Acquiring a single new word, *Papers and Reports on Child Language Development*, Department of linguistics, Stanford University, 15, 17-29.
- Dempster, F. N., 1992: The rise and fall of inhibitory mechanism: Toward a unified theory of cognitive development and aging, *Developmental Review*, 12, 45-75.
- French, R. M. & Cleeremans, A., 2002: *Implicit learning and consciousness*, Brighton, Psychology Press.
- Kipp-Harminsfeger, K., 1995: The development of cognitive inhibition: Theories, definitions, and research evidence in F. Dempster & C. Bramerd eds, *Interference and inhibition in cognition*, 175-204, San Diego, Academic Press.
- Lamberts, K. & Shanks, D., eds, 1997: *Knowledge, concepts, and categories*, Cambridge, MIT Press
- Medin, D. L. & Shoben, E., 1988: Context and structure in conceptual combination, *Cognitive Psychology*, 20, 158-190.
- Murphy, G. L., 2002: *The big book of concepts*, Cambridge, MIT Press.
- Pylyshyn, Z. W., 1981: The imagery debate: Analogue media versus tacit knowledge, *Psychological Review*, 88, 16-45.
- Regier, G., & Brooks, L. R., 1993: Perceptual manifestations of an analytic structure: The priority of holistic individuation, *Journal of Experimental Psychology: General*, 122, 92-114.
- Rips, L. J. 1989 Similarity, typicality, and categorization in S. Vosniadou & A. Ortony eds, *Similarity and analogical reasoning*, 21-59, Cambridge, Cambridge University Press.
- Rosch, E. H., 1978: Principles of categorization in E. H. Rosch et B. Lloyd eds, *Cognition and categorization*, Hillsdale, Lawrence Erlbaum.
- Smith, E., 1989: Concepts and induction in M. Posner ed., *Foundations of cognitive science*, 501-526, Cambridge, MIT Press.
- Smith, E. & Medin, D., 1981: *Categories and concepts*, Cambridge, Harvard University Press.

- Solomon, K. O. & Barsalou, L., 2001: Representing properties locally, *Cognitive Psychology*, 43, 129-169.
- Thibaut, J.-P., 2000: Learning and generalizing new concepts, Proceedings of the Twenty Second Annual Conference of the *Cognitive Science Society*, 745-750, Hillsdale, Lawrence Erlbaum.
- Thibaut, J.-P. & Gelaes, S. (soumis) Learning and automating a new categorization rule: the interaction between the number of exemplars and practice.
- Thibaut, J.-P., Dupont, M., & Anselme, P., 2002: Dissociations between similarity and categorization: why and when?, *Memory & Cognition*, 30, 647-656.
- Thibaut, J.-P., Lemaire, F. & Quadri, J., 1998: Categorization under the influence in M. A. Gernsbacher & S. J. Derry, eds, Proceedings of the Twentieth Annual Conference of the Cognitive Science Society, 822-827, Hillsdale, Lawrence Erlbaum.
- Wisniewski, E. J., 1998: Property instantiation in conceptual combination, *Memory & Cognition*, 26, 1330-1347.