

## Chapitre 5

# Le raisonnement prédicatif

### 5.1. Introduction

Dans un manuel de logique, à un chapitre traitant de la logique propositionnelle font suite des chapitres sur la logique prédicative. On pourrait donc penser que la même organisation s'impose pour un manuel de psychologie du raisonnement, au moins si on pense que dans un livre de psychologie les différents chapitres doivent correspondre à des chapitres différents de l'activité mentale. Toutefois, bien qu'il y ait des raisons de penser que la logique propositionnelle constitue un système psychologique en soi, une fois abandonné son domaine restreint on s'aventure en terrain inconnu. En fait, on ne connaît pas la relation entre les logiques répertoriées et les intuitions que les sujets semblent avoir quand ils raisonnent sur des problèmes qui contiennent des relations logiques non propositionnelles, ni même s'il y a un objet unique qui constitue un domaine d'étude pour le psychologue.

De plus, les problèmes et les études liés à ce que l'on appelle raisonnement prédicatif s'étendent sur un espace très vaste qui couvre, entre autres, le raisonnement avec quantificateurs, ou les relations de comparaison, en particulier celles avec des termes spatiaux et temporels. Cependant, il n'y a pas d'arguments de principe, sauf la recherche d'élégance, pour penser qu'une théorie qui explique un de ces types de problèmes soit aussi valide pour l'autre. Le point commun qui caractérise le raisonnement prédicatif est le fait que la base sur laquelle les sujets s'appuient pour trouver des conclusions nécessaires aux prémisses d'un problème doit se situer à l'intérieur de la structure des énoncés. Toutefois, cela ne signifie pas que tous ces sous-domaines aient un objet unique.

Ajoutons enfin que dans tout raisonnement humain la relation entre le message lu ou entendu et la représentation mentale que le sujet s'en fait n'est pas transparente, car on peut très difficilement séparer les activités de raisonnement des processus interprétatifs des tâches (Politzer, 1990a ; 1990b ; Roberts, Newstead & Griggs, 2001 ; Speber & Wilson, 1995) et souvent la forme logique d'un problème filtrée par les interprétations des sujets est complètement différente de sa forme de surface.

Cependant, dans ce chapitre, nous serons obligés d'opérer par simplifications radicales. D'abord, nous ferons abstraction de tous facteurs interprétatifs, comme si la forme logique d'un problème était en fait très proche de sa forme de surface. Ensuite, nous nous bornerons à analyser des théories sur le raisonnement avec quantificateurs dans des problèmes en forme de syllogisme. Ces choix, dictés par l'exigence de synthèse, peuvent quand même se justifier en vertu du fait que toutes les théories du raisonnement, y compris celles qui soulignent les éléments interprétatifs, doivent tôt ou tard se prononcer sur la forme strictement logique des structures sur lesquelles les sujets, après avoir choisi une interprétation, raisonnent (Bonatti, 1994b).

## 5.2. Problèmes pour une théorie psychologique du syllogisme

Un syllogisme est un problème composé de deux prémisses et une conclusion en forme de prédication simple (par exemple « Q A sont B », où Q est un quantificateur et A et B sont des prédicats à une place), les deux prémisses contenant un prédicat commun qui ne doit pas apparaître dans la conclusion. Cette forme de raisonnement, bien que très peu naturelle, est l'objet d'études depuis l'antiquité et son histoire se confond avec l'histoire même de la logique. Nous nous bornerons à examiner plusieurs théories du raisonnement syllogistique parmi les plus récentes, laissant de côté un certain nombre d'approches plus anciennes que l'on trouvera résumées dans les travaux d'Evans, Newstead & Byrne (1983) et de Politzer (1995). Nous chercherons en particulier des réponses aux questions suivantes :

– d'où vient le sentiment que certaines conclusions semblent s'imposer en vertu simplement de la relation entre les prédicats et les quantificateurs contenus dans les prémisses d'un problème ?

– pourquoi certains problèmes, pourtant très similaires en surface, sont-ils résolus avec des taux de réponses correctes assez différents ?

– pourquoi les sujets jugent-ils que certains problèmes, pourtant très similaires entre eux, ont des degrés de difficulté très différents ?

– pourquoi le simple ordonnancement des prédicats dans les prémisses semble-t-il influencer la forme de la conclusion que les sujets produisent ?

Comme illustration de la première question, considérons les prémisses en (i) :

- (i)            Tous les avocats sont boulangers.  
                   Tous les boulangers sont chauffeurs.

Elles nous amènent à « voir » la conclusion « Tous les avocats sont chauffeurs ». Dans ce cas, notre jugement concorde avec la logique mais, bien sûr, nous n'avons pas besoin de l'étudier pour nous faire une idée assez précise de la validité de cette conclusion. Quel processus psychologique nous fait donc « voir » la conclusion ? Une théorie du raisonnement syllogistique doit proposer une explication à ce phénomène.

Les questions 2 et 3 peuvent sembler très proches, mais elles sont différentes en définitive. Considérez encore (i). Les sujets l'estiment très facile et dans près de 100 % des cas ils acceptent, correctement, la conclusion « Tous les avocats sont chauffeurs ». Dans ce cas, l'impression subjective de facilité et la capacité à donner une réponse logiquement correcte vont donc ensemble.

Considérons maintenant (ii), problème similaire à (i) et aussi perçu comme très facile :

- (ii)            Quelques avocats sont boulangers.  
                   Quelques boulangers sont chauffeurs.

Toutefois dans ce cas les sujets souvent se trompent, en acceptant la conclusion, non valide, « Quelques avocats sont chauffeurs ». Une autre petite modification de (i) suffit pour créer (iii)

- (iii)           Tous les boulangers sont avocats.  
                   Aucun boulanger n'est chauffeur.

qui pourtant est tout de suite perçu comme très difficile et a un taux de solutions correctes qui ne dépasse pas 10 %<sup>1</sup>.

Dans certains cas donc, perception subjective et habilité logique vont ensemble, et dans d'autres elles se séparent. Pourquoi les sujets pensent-ils parfois qu'un problème est facile, et se trompent pourtant en donnant une conclusion non valide ? Et, plus généralement, pourquoi y a-t-il une telle variation dans les perceptions et les performances des sujets, comme les exemples (i) à (iii) le montrent ? Une théorie du raisonnement prédicatif est censée clarifier ces deux questions.

---

1. Une conclusion logiquement valide est, par exemple « Quelques avocats ne sont pas chauffeurs ».

Les syllogismes possèdent des propriétés très attractives. Ils permettent d'exercer un contrôle assez précis sur plusieurs paramètres tels que la longueur physique des problèmes ou le nombre de prémisses et prédicats mentionnés, mais ils suscitent quand même des taux de réussite qui varient de 0 à 100 %. Une telle variation semble dépendre de la représentation mentale des parties logiques qu'ils contiennent. C'est pourquoi l'étude des syllogismes offre un champ idéal pour le psychologue du raisonnement. De plus, la performance syllogistique est corrélée, bien que faiblement (environs 0, 2), avec la performance dans d'autres tâches de raisonnement (Stanovich & West, 1998). Il n'est donc pas absurde de penser extraire des études sur les syllogismes quelques renseignements généraux sur le raisonnement.

Les propriétés formelles des syllogismes sont aussi bien connues. En permutant la forme logique des énoncés (présence/absence de négation et type de quantificateurs) et la position des prédicats dans les prémisses et la conclusion, on obtient 512 syllogismes différents<sup>2</sup>, dont seulement 27 admettent une conclusion valide<sup>3</sup>. Une distinction traditionnelle classe les énoncés d'un syllogisme selon leur *mode*, établi par le type de quantificateurs des prémisses : mode A pour un énoncé universel affirmatif (« Tous les A sont B »), E pour un universel négatif (« Aucun des A n'est B »), I pour un existentiel affirmatif (« Quelques A sont B »), O pour un existentiel négatif (« Quelques A ne sont pas B »). Les syllogismes sont aussi classifiés selon leur *figure*, définie par l'ordre des prédicats dans leurs prémisses. Par exemple (en indiquant l'ordre des prédicats avec des lettres cursives et le mode avec des lettres d'imprimerie), le problème (i) serait du type

<i>A-B</i>	(mode A)
<i>B-C</i>	(mode A)

et le problème (iii) du type :

<i>B-A</i>	(mode A)
<i>B-C</i>	(mode E).

A l'aide de ces classifications, nous pouvons expliquer les données à la base de la question 4. Les syllogismes avec prémisses de figure *A-B*, *B-C* invitent à des conclusions de la forme *A-C*, indépendamment de la forme logique des prémisses.

2. La classification traditionnelle contenait 256 syllogismes, ne considérant pas la possibilité que la conclusion pouvait avoir deux formes différentes, selon l'ordre de ses prédicats.

3. Cette conclusion vaut seulement sous l'hypothèse que l'univers du discours contient au moins un élément. Comme nous le verrons, certaines théories (par exemple celle de Rips) proposent de représenter explicitement cette présupposition existentielle et appliquent une variante de la logique classique pour dériver ces conclusions ; d'autres théories (comme les modèles mentaux) choisissent un type de représentation qui contient implicitement cette hypothèse.

Cet effet substantiel, appelé *effet de figure*, a un aspect paradoxal : il semble contredire la nature même du processus du raisonnement. D'un côté, il semble qu'en cherchant la conclusion d'un problème, nous sommes sensibles à sa forme logique, faute de quoi la psychologie du raisonnement déductif n'existerait pas ; de l'autre côté, l'effet de figure montre que des aspects de surface d'un problème complètement extérieurs à sa forme logique guident notre processus de raisonnement. La question 4 demande aux théories du raisonnement prédicatif de résoudre la tension entre ces deux éléments.

### 5.3. La théorie des modèles mentaux

Sans doute la théorie la plus répandue chez les psychologues du raisonnement déductif est la théorie des modèles mentaux de Johnson-Laird (1983 ; 1991). Appliquée désormais à de nombreux domaines du raisonnement, déductif ou non, elle est née et a gagné sa force surtout grâce à son explication des problèmes du syllogisme. Elle s'est imposée alors comme seule véritable théorie digne de ce nom. Nous présenterons d'abord une partie de la « mécanique » des modèles mentaux prédicatifs, pour montrer comment la solution aux problèmes 1 à 4 semble découler naturellement d'elle. Nous rappellerons aussi une partie des résultats fondamentaux qui semblent confirmer la théorie, pour comprendre pourquoi les modèles mentaux se sont imposés avec une telle force chez les psychologues du raisonnement déductif. Nous passerons ensuite aux critiques qui lui ont été faites récemment et essayerons de la réévaluer en fonction d'elles.

#### 5.3.1. Modèles pour le raisonnement syllogistique

Le trait fondamental du raisonnement prédicatif, et par conséquent du raisonnement syllogistique, est la quantification. Tout traitement logique des quantificateurs demande un langage avec des symboles de type différent, notamment des constantes, qui se réfèrent à des individus, et des variables, qui peuvent se référer à tout l'univers du discours. En logique, les quantificateurs n'ont un sens que s'ils sont appliqués à des variables. Le raisonnement prédicatif semble donc dépendre de l'existence de structures mentales abstraites sur lesquelles les sujets peuvent effectuer des opérations également abstraites, de nature essentiellement algébrique, comme par exemple la substitution de valeurs appropriées pour les variables.

En dépit de cela, l'idée de base des modèles mentaux est que tout raisonnement prédicatif ne demande pas de variables ou autres structures abstraites, mais s'effectue sur des échantillons constitués d'individus ou d'exemplaires concrets représentant le domaine du discours. Comme l'affirme Johnson-Laird (1983,

p. 488), « un modèle ne contient pas de variables. Une représentation linguistique de, par exemple, « Tous les avocats sont boulangers » pourrait prendre la forme de « Pour tous les  $x$ , si  $x$  est un avocat, alors il est un boulanger ». Au lieu d'une variable, comme  $x$  dans cette expression, un modèle emploie des individus concrets (*tokens*) qui représentent un ensemble d'individus » (voir aussi (Johnson-Laird, 1991, p. 488)). Le premier présupposé de la théorie des modèles mentaux est donc que les représentations mentales sous-jacentes aux processus de raisonnement prédicatif sont des échantillons des catégories correspondant aux prédicats mentionnés dans les prémisses. Pour simplifier (mais avec une simplification fondée sur la théorie), on se représente mentalement l'énoncé « Tous les avocats sont boulangers » en pensant à Jean qui est avocat et boulanger, Marie qui est avocate et boulangère, Charles qui est avocat et boulanger, et ainsi de suite. Un deuxième présupposé de base de la théorie est que notre esprit paresseux essaie toujours d'exercer l'effort minimum nécessaire pour représenter les énoncés d'un problème. Donc, en joignant les deux présupposés, on conclut qu'à chaque énoncé devra correspondre une quantité minimale d'échantillons mais suffisante pour rendre sa signification. Si besoin est, ces représentations minimales pourront être enrichies avec d'autres échantillons qui rendront explicites certains aspects de la signification des énoncés initialement non représentés.

Voyons d'abord les représentations minimales initiales associées aux énoncés des quatre modes syllogistiques. Le modèle mental correspondant à un énoncé en mode A (« Tous les A sont B ») sera

[A]	B
[A]	B
...	

où chaque ligne représente un individu dans le modèle (et non pas un modèle différent, à la différence de la convention suivie pour les modèles propositionnels), les points signifient que d'autres individus peuvent être ajoutés au modèle, et les crochets signalent que l'on ne peut pas avoir d'individus avec la propriété A qui soient différents de ceux déjà représentés. Si on ajoute au modèle un autre individu qui est A, celui-ci doit être aussi B : on ne peut pas ajouter des individus qui sont A et non-B.

Un énoncé en mode I (« Quelques A sont B ») aura la représentation mentale

A	B
A	B
....	

L'absence de crochets laisse ouverte la possibilité d'envisager des A qui ne sont pas B, ce qui d'après la théorie des modèles devrait suffire pour distinguer la signification d'un énoncé purement existentiel de celle d'un énoncé universel.

Un énoncé du mode E (« Aucun A n'est B ») est représenté par des exemples de A qui ne sont pas B et des B qui ne sont pas A, en utilisant les crochets d'exhaustivité :

[A]  
[A]  
[B]  
[B]

...

et enfin un énoncé de mode O (« Quelques A ne sont pas B ») est représenté par des A qui pourraient être ou non être B (cette information étant laissée indéterminée), des B qui sont A et des B qui, en revanche, ne peuvent pas être A :

A  
A  
A [B]  
[B]

où la présence de deux lignes différentes contenant des B entre crochets signifie que tout autre B qui peut s'ajouter au modèle doit apparaître seul ou avec un A. Dans cette représentation donc, qui exprime les présuppositions existentielles directement, s'il y a des A qui ne sont pas B, alors il existe forcément des B qui ne sont pas A, ce qui marque une différence avec la signification logique du quantificateur existentiel sur laquelle on va revenir dans les pages suivantes.

### 5.3.2. Compréhension d'un texte et composition des prémisses : quelques exemples de raisonnements

Dans la théorie des modèles mentaux, l'activité de raisonnement partage beaucoup de points communs avec l'activité de compréhension d'un texte. D'après les modèles mentaux (voir par exemple (Garnham, 1987)), comprendre un texte signifie construire un modèle mental dans lequel les informations que le texte fournit s'accumulent, tout en gardant leur cohérence. On peut voir les prémisses d'un problème logique comme un texte, pour lequel le sujet cherche à créer un modèle (ou plusieurs, si nécessaire) qui les intègre. Une fois examinée une prémisses, le ou les modèles mentaux minimaux correspondants sont construits ; les prémisses successives en génèrent d'autres, qui sont intégrés aux modèles déjà construits, de sorte que la structure en mémoire reste la plus économique possible. *Une conclusion est simplement une description synthétique des modèles ainsi construits.* Ce qui rend un texte logique différent d'un texte quelconque est ce qui vient après la formulation d'une conclusion : une conclusion étant formulée, une recherche de modèles alternatifs des prémisses peut s'initier. Si l'on en trouve, la conclusion est considérée non valide ; dans le cas contraire, elle est retenue comme valide. On verra plus avant

comme cette vision du raisonnement peut facilement générer des prédictions sur la difficulté des syllogismes.

Comment composer les modèles générés par les différentes prémisses est une question cruciale, qui cependant n'a jamais été pleinement expliquée (voir (Bonatti, 1994a ; 1998 ; Ford, 1995)). L'idée intuitive est assez claire : si deux modèles contiennent de l'information redondante, elle ne sera pas redoublée ; si un modèle contient des informations nouvelles concernant des individus dans les modèles déjà construits, elle seront ajoutées aux individus déjà existants ; s'il contient des informations sur des individus non présents dans les modèles déjà construits, les nouveaux individus et les informations correspondantes y seront ajoutés ; finalement, s'il contient des informations incompatibles avec les modèles déjà construits, la composition sera impossible : aucun modèle ne sera retenu. Donc parfois la composition de différents modèles des prémisses requiert le même nombre de modèles émanant de chaque prémisses individuellement ; parfois elle en exige un nombre supérieur ; et parfois elle en réduit le nombre ou les élimine complètement. Les exemples suivants aideront à clarifier le mécanisme de composition.

Reprenons le problème (i). La compréhension de la première prémisses aboutira à la construction du modèle suivant :

[A]	B
[A]	B

...

et la deuxième prémisses construira le modèle :

[B]	C
[B]	C

...

Dans ce cas, la deuxième prémisses ajoute de l'information sur les individus contenus dans le modèle de la première prémisses (elle dit que les B sont C). L'information est donc directement intégrable au modèle de la première prémisses :

[[A]	B]	C
[[A]	B]	C

...

Les crochets doubles indiquent que les A sont épuisés par rapport aux B, et que les B le sont par rapport aux C.

Notons un point qui deviendra important : dans ce cas, il n'y a pas de façon alternative de composer les modèles. La conclusion « Tous les A sont C » apparaît comme une nécessité.

Imaginons maintenant que la deuxième prémisses, au lieu d'être « Tous les B sont C », soit « Tous les C sont D » ; dans ce cas, elle ne donnerait aucune information sur les individus déjà présents dans le modèle, et donc le modèle de la première prémisses serait élargi :

[A]	B
[A]	B
[C]	D
[C]	D
...	

Si en revanche elle était « Aucun A n'est B », toute composition avec le modèle déjà existant serait impossible, le premier modèle étant :

[A]	B
[A]	B
...	

et le deuxième étant :

[A]	
[A]	
[B]	
[B]	
...	

Le premier modèle oblige tous les A à être B, alors que le deuxième demanderait d'ajouter des individus qui soient A mais non pas B. Cela interdit la construction d'un modèle à partir des deux prémisses.

Les problèmes (ii) et (iii) offrent des exemples où la composition des prémisses peut ou doit générer plusieurs modèles. Le problème (ii) a une composition minimale élémentaire : les modèles des deux prémisses étant :

A	B
A	B
...	

et :

B	C
B	C
...	

une première composition possible est simplement d'ajouter au modèle de la première prémisses l'information sur les B contenue dans la deuxième prémisses, donnant le modèle :

A	B	C
A	B	C
...		

Si celui-ci était le seul modèle possible, la conclusion « Quelques A sont C » serait valide. Elle ne l'est pas, mais, comme on l'a rappelé, plusieurs sujets la proposent, ce qui pourrait être expliqué par le fait qu'ils s'arrêtent à ce niveau. S'ils étaient moins paresseux, ils pourraient voir qu'il n'y a aucune raison de penser que les B dont la première prémisses parle sont les mêmes que ceux évoqués dans la deuxième prémisses (aucun crochet d'exhaustivité ne l'exige). Ils pourraient donc voir que les deux prémisses sont validées aussi par le modèle :

A	B
A	B
B	C
B	C
...	

qui rend donc vraies les prémisses (car il y a des A qui sont B et des B qui sont C), mais pas la conclusion examinée (car il n'y a aucun A qui est C. Il pourrait y en avoir, mais dans ce modèle, il n'y en a pas). L'exemple montre pourquoi les sujets pourraient considérer valide une conclusion qui ne l'est pas : il pourraient construire un modèle des prémisses qui la vérifie, sans considérer qu'un autre modèle ne la vérifie pas.

Le problème (ii) nous montre que parfois un seul modèle suffit pour réaliser les prémisses, même si l'on peut en construire d'autres. En revanche, le problème (iii) oblige à en construire plusieurs. Pour le voir, il faut d'abord introduire un autre type d'opération nécessaire pour la composition. Dans la théorie des modèles, la notation a une signification bien précise. Ainsi, le fait que dans les lignes des modèles un A soit à la gauche ou à la droite d'un B n'est pas un accident : la théorie suppose que l'intégration des items en mémoire se fait par proximité. Si un modèle contient la ligne :

A	B
---	---

et un autre modèle contient :

B	C,
---	----

alors l'intégration peut se faire directement, car par hypothèse les lettres occupent des espaces proches en mémoire :

	A	B	B	C	
=>	A	B	<del>B</del>	C	C
=>	A	B	C		

où les flèches indiquent les processus successifs de construction d'une ligne intégrée en mémoire.

Mais si la ligne du deuxième modèle était :

C      B,

alors intégration demanderait un passage ultérieur. La juxtaposition des lettres donnerait :

A      B      C      B

et les deux B devraient être rapprochés en mémoire pour éliminer la lettre redondante. Cela se fera en inversant les lettres de la ligne du deuxième modèle (en la transformant de C B à B C). Donc l'intégration se ferait avec les passages suivants :

A	B	C	B	
=>	A	B	B	C
=>	A	B	<del>B</del>	C
=>	A	B	C	

Cette opération d'inversion a un coût : elle ajoute des difficultés ultérieures et peut conduire à des erreurs.

Revenons au modèle intégré des prémisses du syllogisme (iii). La première prémisses (« Tous les boulangers sont avocats ») demande le modèle :

[B]      A  
 [B]      A  
 ...

et la deuxième (« Aucun des boulangers n'est chauffeur ») le modèle :

[B]  
 [B]  
   [C]  
   [C]  
 ...

Les B, qui ne sont pas « dans la position correcte » pour être intégrés, doivent changer de place d'abord, par exemple en transformant l'ordre de termes dans le modèle de la première prémisses :

A      [B]  
 A      [B]  
 ...

Seulement à ce point-là le modèle pour la deuxième prémisse peut être intégré. Sa juxtaposition donnera le modèle :

A	[B]	[B]
A	[B]	[B]
	[C]	
	[C]	

...

à partir duquel les B redondants seront éliminés et les C, qui sont épuisés par rapport aux B, resteront sur des lignes séparées :

A	[B]
A	[B]
[C]	
[C]	

...

Ce modèle rend vraie la conclusion « Aucun avocat n'est chauffeur », une conclusion effectivement donnée par un certain nombre de sujets (voir (Johnson-Laird, 1984, tableau 12)). Toutefois, une telle conclusion n'est pas valide : des modèles alternatifs existent dans lesquels elle n'est plus vraie. Il suffit de remarquer que la première prémisse admet aussi un modèle « enrichi », qui peut contenir des A qui ne sont pas des B :

[B]	A
[B]	A
	A

...

et que même si les lignes contenant des C dans le deuxième modèle ne peuvent pas contenir des B, elles pourraient bien contenir des A. La composition du modèle « enrichi » de la première prémisse et du modèle de la deuxième prémisse pourrait donc donner lieu au modèle :

A	[B]	
A	[B]	
A		[C]
A		[C]

dans lequel toutes les lignes contenant des C contiennent aussi des A, ou bien au modèle :

A	[B]	
A	[B]	
A		[C]
		[C]

dans lequel seulement certaines en contiennent. Ni l'un ni l'autre ne comportent la conclusion « Aucun avocat n'est chauffeur », qui est donc invalide. L'erreur des sujets qui proposent cette conclusion est proprement un erreur de perspective : ils voient le problème sous l'angle du première modèle, le plus simple, mais ils n'arrivent pas à envisager les modèles alternatifs.

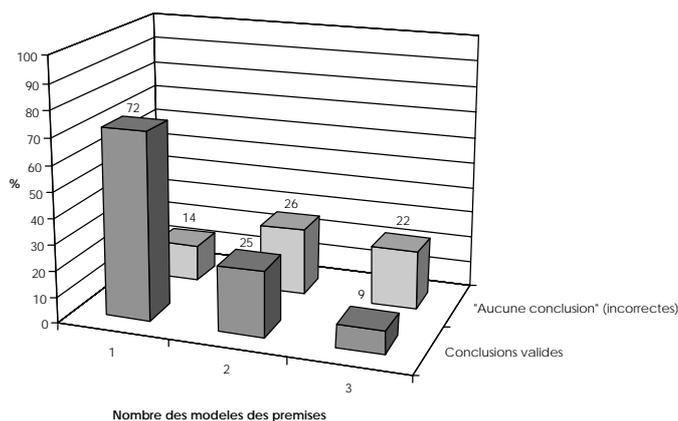
### **5.3.3. La solution des problèmes par les modèles mentaux : théorie et évidence**

Reprenons les éléments essentiels contenus dans les exemples précédents. Pour la théorie des modèles mentaux, comprendre un problème c'est construire un modèle mental minimal de ses prémisses. Dans ce processus de construction, deux facteurs fondamentaux déterminent la difficulté d'un problème. Le premier est la construction éventuelle de modèles alternatifs des prémisses, s'il y en a ; le deuxième est l'ordonnement en mémoire des individus des modèles selon un principe de proximité. Les deux processus sont coûteux et les sujets, s'ils le peuvent, s'en passeront, mais même quand ils s'y engagent, le succès n'est pas garanti : il y a des limites à ce que la mémoire peut accomplir et la construction de modèles alternatifs peut vite la saturer.

Nous avons maintenant tous les éléments pour comprendre comment la théorie de modèles mentaux peut expliquer les problèmes du raisonnement syllogistique. Pour la théorie, raisonner c'est formuler une conclusion, c'est-à-dire décrire les modèles intégrés des prémisses, et valider la conclusion en construisant des modèles alternatifs, des contre-exemples potentiels. Pourquoi donc certaines conclusions s'imposent-elles à notre esprit comme valides (Question 1) ? Parce qu'elles sont une description fidèle des prémisses intégrées ; si nous ne trouvons pas de modèles alternatifs, elles nous apparaîtront nécessaires. Pourquoi certains problèmes similaires en surface sont-ils résolus avec des taux de réponses correctes assez différents (Question 2) ? Parce qu'ils demandent de construire un nombre différent de modèles, et plus ce nombre est élevé, plus grand sera le taux d'erreurs. Pourquoi les sujets jugent-ils que certains problèmes très similaires entre eux ont des degrés de difficulté très différents (Question 3) ? Toujours, parce qu'ils demandent un nombre de modèles différents, et plus ce nombre est élevé, plus le problème apparaîtra difficile si les sujets s'aperçoivent qu'il demande effectivement plusieurs modèles : s'il ne s'en aperçoivent pas, ils pourront croire le problème facile et donner une conclusion erronée, comme dans le cas du syllogisme (ii). Finalement, pourquoi l'ordonnement des prédicats dans les prémisses influence-t-il la forme de la conclusion produite (Question 4) ? Parce qu'il conduit à construire des structures en mémoire qui se prêtent mieux à être décrites par des conclusions dont les prédicats sont dans un certain ordre.

La théorie des modèles mentaux semble donc avoir toutes les ressources pour pouvoir facilement résoudre les problèmes posés par les syllogismes. Nombre

d'études ont testé ses prédictions. D'abord, les théoriciens des modèles mentaux ont cherché une relation entre quantité de modèles demandée par les syllogismes et taux de solutions correctes. La théorie prédit que plus ce nombre est grand, plus le taux de conclusions erronées sera élevé. De même, plus un problème demande de modèles, plus les sujets tendront à répondre qu'il n'y a pas de conclusion valide, car s'ils trouvent que la première conclusion est réfutée dans un des modèles qu'ils ont construits, ils entreprendront rarement l'effort de recommencer le processus de recherche et de test d'une autre conclusion. De manière générale, plusieurs expériences semblent confirmer ces prédictions. La figure 1 montre que sur l'ensemble des syllogismes testés par (Johnson-Laird, 1984) quand le nombre des modèles croît, le taux de conclusions correctes décroît de 70 % à 10 %, alors que le taux de conclusions (non correctes) de la forme « Il n'y a pas de conclusions » augmente (pour autres résultats essentiellement identiques, voir (Bara *et al.*, 1995), (Bucciarelli & Johnson-Laird, 1999)).

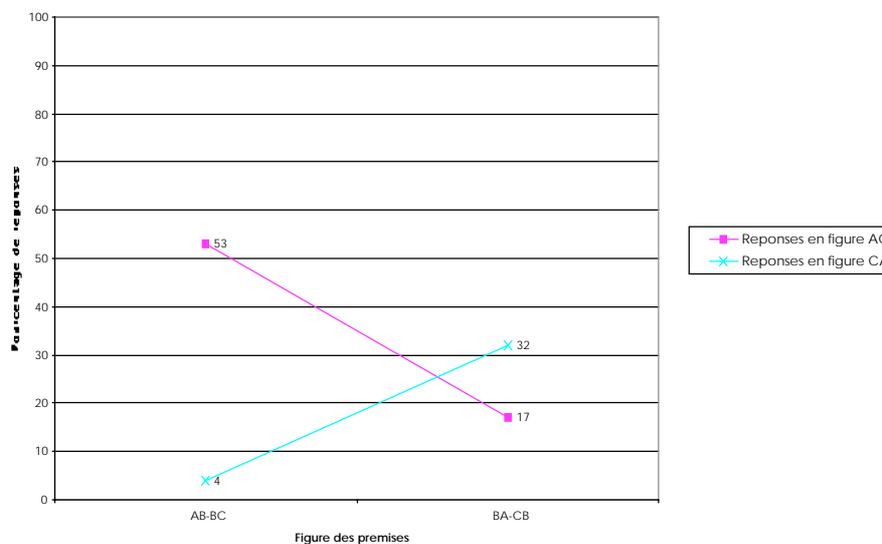


**Figure 5.1.** Pourcentages de conclusions valides et de réponses « Aucune conclusion » en fonction du nombre de modèles de prémisses (d'après Johnson-Laird & Bara, 1984)

Une troisième prédiction concerne les figures des syllogismes, qui sont censées produire des niveaux de difficulté différents. Considérons les opérations nécessaires pour intégrer les modèles des prémisses. Dans la figure *A-B, B-C* les items sont déjà dans le « bon ordre » pour être intégrés ; cette figure sera donc la plus facile. Si l'on fait aussi l'hypothèse que la mémoire fonctionne selon le principe « premier entré, premier sorti », on pourra aussi prédire que les items du modèle intégré seront « lus » selon l'ordre *A-B-C* et donc qu'en général pour les sujets il sera plus facile de trouver des conclusions en forme *A-C*.

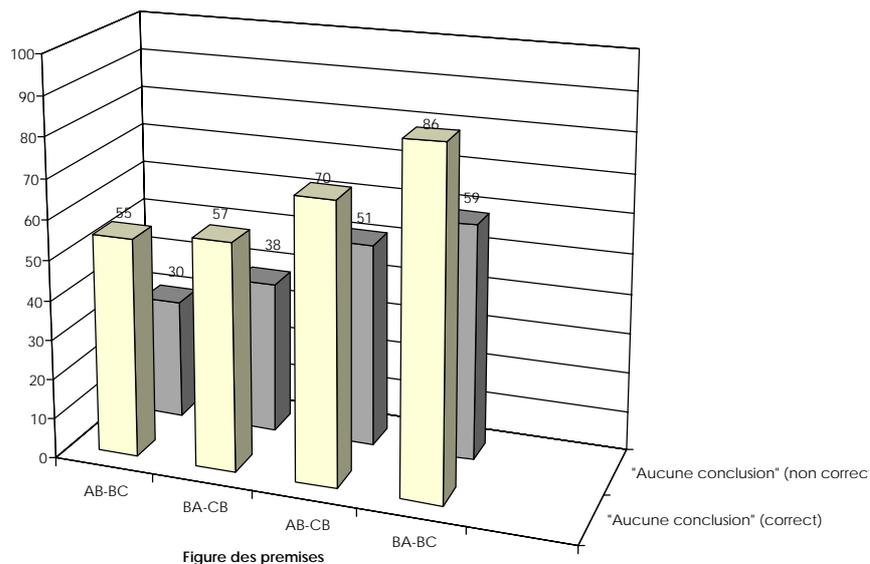
A la différence de la figure  $A-B, B-C$ , les modèles pour les prémisses de figure  $B-A, C-B$ , ne peuvent pas être intégrés directement ; elle sera donc plus difficile. Cependant, cette figure peut être facilement transformée dans une structure intégrable en inversant l'ordre de construction des deux prémisses : si l'on commence avec la deuxième prémisses, la figure  $B-A, C-B$  devient une figure  $C-B, B-A$  et les termes moyens sont chargés en mémoire avec la relation de proximité correcte, sans besoin d'opérations ultérieures ; une fois faite l'inversion des prémisses, la seule différence avec la figure  $A-B, B-C$  sera que la forme de la conclusion la plus facile à trouver sera  $C-A$ . Les deux autres figures ( $AB-CB$  et  $BA-BC$ ) qui ont le terme  $B$  « du même côté » dans les deux prémisses, ne sont pas aussi simples. Changer l'ordre de construction des prémisses ne suffit pas : il faut au moins inverser les termes d'une des deux. Elles seront donc les figures les plus difficiles, d'autant plus qu'elles ne facilitent aucune forme de conclusion.

Plusieurs expériences, à partir de Johnson-Laird & Bara (1984), ont confirmé cette série de prédictions. La figure 2 révèle un effet inverse pour les figures  $A-B, B-C$  et  $BA-CB$  comme prédit : les conclusions de forme  $A-C$  sont plus fréquentes que celles de forme  $C-A$  pour la première figure et cette relation s'inverse exactement pour la deuxième figure.



**Figure 5.2.** Représentation de la conclusion en fonction de la figure des prémisses (d'après Johnson-Laird & Bara, 1984)

La figure 5.3 montre aussi que les réponses du type « Il n'y a pas de conclusion », qu'elles soient correctes ou non, augmentent en suivant l'ordre de difficulté prédit par l'explication donnée.



**Figure 5.3.** Pourcentage de réponses « Aucune conclusion » selon la figure des prémisses (d'après Johnson-Laird & Bara, 1984)

Le fait que ces prédictions précises et originales aient été présentées et vérifiées ensemble dans un seul article (Johnson-Laird & Bara, 1984) fait bien comprendre la raison de son impact sur l'ensemble de la psychologie du raisonnement. Dans les années qui ont suivi le travail de Johnson-Laird et Bara, d'autres sortes de données que nous ne pouvons pas discuter ici se sont accumulées en faveur de la théorie, qui a vite obtenu un consensus presque général de la part des psychologues.

#### 5.3.4. Critiques et problèmes de la théorie des modèles

Dans la section précédente nous avons cherché à montrer comment la théorie des modèles mentaux semble pouvoir offrir des solutions élégantes à tous les problèmes du syllogisme, ce qui explique aussi sa place dans la psychologie du raisonnement. Toutefois, nous allons voir que la théorie prédictive des modèles mentaux prête le flanc à des critiques. Nous allons voir pourquoi.

En somme, la théorie suppose que la recherche d'une conclusion valide à un problème prédicatif demande trois étapes :

- intégration des prémisses dans un modèle initial ;
- formulation d'une conclusion ;
- validation de la conclusion par la construction de modèles alternatifs.

Or, la bonne formulation et la plausibilité de toutes ces étapes, ainsi que les résultats expérimentaux censés les confirmer, ont été mis en question.

Commençons par la dernière. Pour la théorie des modèles, raisonner c'est construire des contre-exemples au modèle initial des prémisses. Toutefois, certains auteurs ont remarqué que les sujets ne semblent pas s'engager souvent dans la recherche de contre-exemples, si jamais ils le font (voir par exemple (Polk & Newell, 1995)). L'observation de base est simple : dans le cas où les prémisses admettent un seul modèle, on devrait s'attendre à une performance presque parfaite, car les sujets ne peuvent que vérifier la conclusion dans ce modèle : un procédé relativement simple. Si un problème admet plusieurs modèles, on peut envisager que les sujets n'arrivent pas à les construire tous, et donc, qu'ils fassent des erreurs. Toutefois, quand les problèmes demandent seulement deux modèles, on devrait au moins voir une forte tendance à construire le deuxième modèle, car si même dans ce cas aucun modèle outre le premier n'est construit, soutenir que le raisonnement est une recherche de contre-exemples n'a aucun sens. En fait, les données semblent suggérer exactement cela. Ainsi, Newstead, Handley & Buck (1999) ont voulu étudier si les conclusions que les sujets donnent quand ils cherchent à résoudre des syllogismes, ou les diagrammes qu'ils dessinent, indiquent qu'ils considèrent des modèles alternatifs au premier. Ils n'en ont trouvé aucun signe. Il est possible bien évidemment de manipuler les expériences de sorte que les sujets envisagent des situations alternatives à leur interprétation initiale des prémisses (Buccarelli, 1999), mais ce résultat n'offre qu'une maigre consolation. Avant tout, une théorie du raisonnement devrait expliquer les processus fondamentaux de raisonnement que les sujets mettent en place de façon immédiate et naturelle ; si les données suggèrent que la construction de modèles alternatifs n'est pas une stratégie naturelle, cela veut dire que la théorie des modèles mentaux échoue comme théorie du raisonnement déductif naturel.

Des critiques encore plus fondamentales concernent la première et la deuxième étape, et notamment les mécanismes mêmes de construction des modèles et d'intégration des prémisses. Comprendre les modalités exactes de l'application des règles de composition des prémisses relève souvent moins de la science que de l'art. Surtout Ford (1995) a mis à nu les problèmes qu'elles soulèvent pour la compréhension du raisonnement prédicatif. Ford a d'abord remarqué des incohérences dans la construction des modèles pour certains énoncés et syllogismes. Considérons le syllogisme suivant :

Tous les boulangers sont avocats.

Tous les boulangers sont chauffeurs.

Les deux prémisses devraient donner lieu aux deux modèles initiaux

[B] A  
 [B] A  
 ...

et :

[B] C  
 [B] C  
 ...

et après inversion des termes du premier modèle ils devraient être intégrés dans le modèle suivant :

A [B] C  
 A [B] C  
 ...

Une fois éliminé le terme moyen, la simple conclusion « Quelques avocats sont chauffeurs » devrait se présenter car le modèle intégré est exactement celui de cet énoncé. Pourtant, dans les données de (Johnson-Laird & Bara, 1984) personne ne répond correctement à ce syllogisme. Comment expliquer dans la théorie des modèles que ce syllogisme ait un seul modèle et pourtant un taux de résolution de 0 % ? Sans doute, une partie de la difficulté pourrait dépendre de sa figure, mais la taille de l'effet est trop nette pour qu'un effet de figure puisse en rendre compte ; en fait, les deux autres syllogismes de la même figure à un modèle étudiés par (Johnson-Laird & Bara, 1984) avaient un taux de réponses correctes d'environ 50 %. Or, ce syllogisme « déviant » considéré par eux comme syllogisme à un modèle dans se retrouve classé comme syllogisme à trois modèles quelques années après (Johnson-Laird & Byrne, 1991, p. 110). Avec ce changement de classification il s'aligne avec les résultats sur les syllogismes à trois modèles, mais il témoigne aussi de la difficulté de comptage des modèles.

Il ne s'agit pas d'un cas isolé. Comme Ford (1995) le note, un autre cas dans lequel il est difficile de savoir compter correctement les modèles, potentiellement plus grave, concerne les énonces du type O (« Quelques A ne sont pas B »). La représentation proposée par (Johnson-Laird et Bara, 1984) et (Johnson-Laird & Byrne, 1991) est la suivante :

A  
 A  
 A [B]  
 [B]  
 ...

Toutefois elle ne peut pas être correcte, car elle rend vraie aussi l'énoncé symétrique « Quelques B ne sont pas A », qui n'est pas une de ses conséquences logiques. Or, un énoncé du type O admet aussi le modèle :

```
A
A
A      [B]
A      [B]
...
```

qui le rend vrai sans pour autant rendre vrai aussi l'autre énoncé, car dans le modèle il n'y a pas de B qui ne sont pas A. Quelle représentation initiale devrait-on choisir, la première ou la deuxième ? Du point de vue de la charge en mémoire, elles devraient être presque équivalentes. Mais si l'on choisit la première, alors le compte des modèles pour les prémisses composées change. Considérons par exemple le syllogisme :

Quelques A ne sont pas B.  
 Quelques B sont C

Si l'on opte pour la deuxième représentation, on obtient les deux modèles :

```
A
A
A      [B]          B      C
A      [B]          B      C
...
```

Dans leur composition, il est indifférent que des C soient ajoutés à la troisième ou à la quatrième ligne du premier modèle : les deux solutions suivantes sont équivalentes :

```
A
A
[A      [B]]      C
[A      [B]]

A
A
[A      [B]]
[A      [B]]      C
```

Mais si l'on choisit la deuxième représentation pour les énoncés du type O, alors le résultat de la composition change selon l'endroit où on l'ajoute le terme C :

```
A
A
A      [B]      C
...

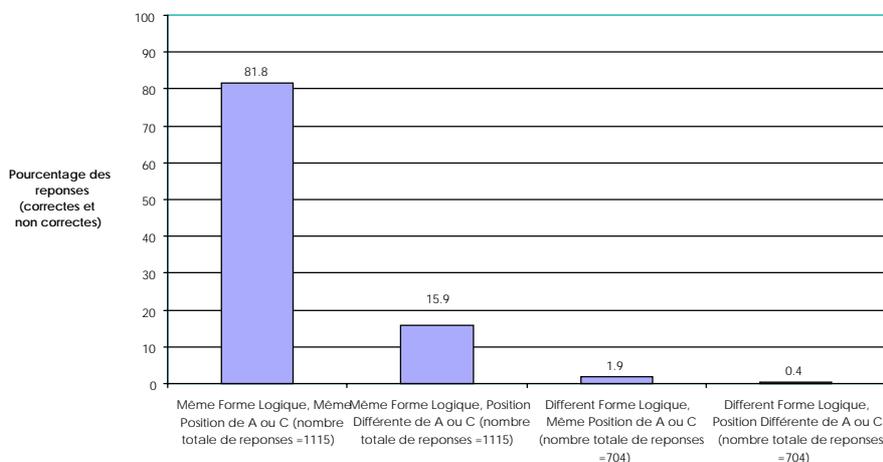
A
A
A      [B]
...      [B]      C
```

Donc avec la première lecture les prémisses ont un modèle minimal, alors qu'avec la deuxième ils en ont au moins deux. Bien évidemment, les prédictions sur tout syllogisme contenant une prémisses O changent.

En fait, souvent on ne sait pas comment compter les modèles prédictifs, et en conséquence, on ignore quelles sont les prédictions de la théorie prédictive des modèles : elle semble avoir trop de degrés de liberté pour pouvoir se prêter à des tests rigoureux de confirmation ou de réfutation.

### 5.3.5. Avons-nous besoin de la théorie des modèles mentaux ?

Ces critiques touchent les points fondamentaux de la théorie des modèles mais comment faut-il évaluer leur profondeur ? Nous pourrions penser qu'elles obligent simplement à la modifier, comme le pensent par exemple Barrouillet & Lecas (1998), ou bien qu'il faut chercher une théorie du raisonnement alternatif. Pour décider, il faut d'abord poser la question suivante : pouvons nous expliquer le raisonnement syllogistique sans faire référence à la théorie des modèles mentaux ou avons-nous besoin d'elle pour rendre compte des très convaincantes données de Johnson-Laird et ses collaborateurs ? Nous suivrons encore l'analyse de Ford.



**Figure 5.4.** Relation entre la forme de la conclusion et des prémisses, et entre la position des termes A et C dans la conclusion et dans les prémisses (d'après Ford, 1995)

Ford remarque d'abord que la réduction de performance entre syllogismes à un, deux et trois modèles (voir figure 5.1) pourrait dépendre d'autres propriétés non liées au nombre des modèles qu'ils demandent de construire. Notamment, Ford remarque que la conclusion valide de tous les syllogismes à un modèle testés par (Johnson-Laird & Bara, 1984) (sauf le syllogisme que l'on vient d'examiner, dont le nombre de modèles a été changé de un à trois) a la forme logique d'une des prémisses et les termes A et C dans la même position que les prémisses. Ensuite, elle observe aussi que la conclusion de tous les syllogismes à deux modèles a la forme logique d'une des prémisses, mais sa forme est toujours « Quelques A ne sont pas C ». Enfin, elle note que la conclusion de tous les syllogismes à trois modèles, comme la conclusion de ceux à deux modèles, a la forme « Quelques A ne sont pas C », mais, à différence des syllogismes à deux modèles, aucune des prémisses n'a cette forme. Si donc les sujets ont tout simplement une tendance à chercher des conclusions qui ont la même forme logique que l'une des prémisses et les termes dans la même position que dans l'une des prémisses, alors on pourrait prédire le même ordre de difficulté des syllogismes trouvé par (Johnson-Laird & Bara, 1984) sans faire aucune référence aux modèles mentaux : les problèmes dont la conclusion a la même forme logique que l'une des prémisses seront les plus faciles, suivis par ceux qui ont une conclusion avec la même forme logique que l'une des prémisses mais les termes A ou C en position différente par rapport aux prémisses, et ensuite par les problèmes dont la conclusion a une forme logique différente des deux prémisses et les termes A ou C en position différente par rapport aux prémisses.

Cette explication alternative est-elle plausible ? D'abord, en réexaminant les données de (Johnson-Laird & Bara, 1984), Ford montre qu'en effet les sujets ont une très forte tendance à donner des conclusions avec la même forme que l'une des prémisses. De plus, il leur est manifestement difficile de trouver des conclusions (qu'elles soient correctes ou erronées) qui ont les termes dans des positions différentes par rapport aux prémisses. La figure 5.4 montre que le taux de solutions correctes suit rigoureusement l'ordre prédit par la proposition de Ford. On peut donc expliquer autrement la forte chute en performance entre les différents problèmes testés par (Johnson-Laird & Bara, 1984) ainsi que l'effet de figure, en considérant non pas les modèles, mais leur structure logico-syntaxique.

Ensuite, grâce à une des rares analyses des protocoles dans le domaine du raisonnement déductif, Ford a essayé de détecter les types de représentations mentales que les sujets emploient, au moins consciemment, pendant la résolution des problèmes syllogistiques. Elle a pu montrer qu'on peut partager les sujets en deux groupes caractérisés par différentes stratégies de résolution. Certains sujets raisonnent avec des stratégies fondamentalement verbales, que Ford caractérise comme algébriques car elles semblent faire appel à des principes de substitution de valeurs dans des variables. Par exemple, face au syllogisme :

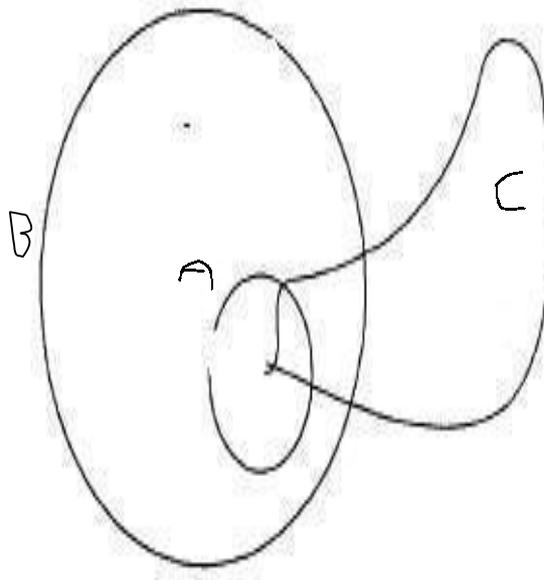
Tous les boulangers sont avocats  
Quelques boulangers ne sont pas chauffeurs

un sujet (Eric) raisonne ainsi :

« Ok, (...) j'aime quand ils commencent avec ces choses car, tu sais, c'est beaucoup plus facile de penser au reste car on peut dire que, donc, quelques uns de ces boulangers avocats ne sont pas chauffeurs et on peut tout simplement le mettre a sa place, comme si l'on faisait de l'algèbre, tu vois ce que je veut dire, on peut dire que quelques avocats ne sont pas des chauffeurs » (Ford, 1995, p. 17).

D'autres sujets utilisent par contre une stratégie plus graphique : ils cherchent à visualiser les relations de sous-ensembles déterminés par les prémisses. Ainsi, en raisonnant sur le syllogisme

Tous les avocats sont boulangers.  
Quelques chauffeurs ne sont pas boulangers.  
Mike dessine des cercles qui s'entrecroisent :



et il commente : « Ok, donc nous avons une classe de boulangers et les avocats y sont dedans, donc tous les avocats sont boulangers, ils le sont tous, tous les avocats sont boulangers, quelques chauffeurs ne sont pas boulangers, donc cela signifie que nous avons une classe de chauffeurs, au moins quelques chauffeurs, qui sont en

dehors de la classe des avocats, donc cela signifie que quelques chauffeurs ne sont pas avocats » (Ford, 1995, p. 15) ; traduction adaptée à notre exemple).

Les sujets qui utilisent cette stratégie semblent employer des structures graphiques telles que des diagrammes de Venn, qui représentent les ensembles et leurs relations, et non pas directement des individus, comme la théorie des modèles le voudrait. La distinction entre les deux types de stratégie est frappante. La figure 5.5 présente le nombre total de fois où les sujets testés par Ford utilisent au moins une fois un comportement « algébrique » ou un comportement « graphique » dans la solution des 27 syllogismes ; on peut voir clairement qu'il n'y a presque aucun chevauchement entre les sujets. Cela suggère qu'ils déploient une stratégie soit algébrique, soit graphique, mais pas les deux ensemble. Pour fonder encore mieux cette différence entre groupes, Ford a aussi montré que les sujets « graphiques » résolvent mieux certains syllogismes, tandis que les sujets « algébriques » ont une performance meilleure avec d'autres types de syllogismes. Pour la discussion présente, le point fondamental est que, si l'on accepte l'analyse de Ford, il y a deux types de sujets qui emploient des stratégies différentes pour résoudre des syllogismes ; ni l'une ni l'autre de ces stratégies ne peut être décrite suivant les modèles mentaux décrits par Johnson-Laird (mais voir (Johnson-Laird *et al.*, 2000b) pour une réplique au point de vue de Ford).

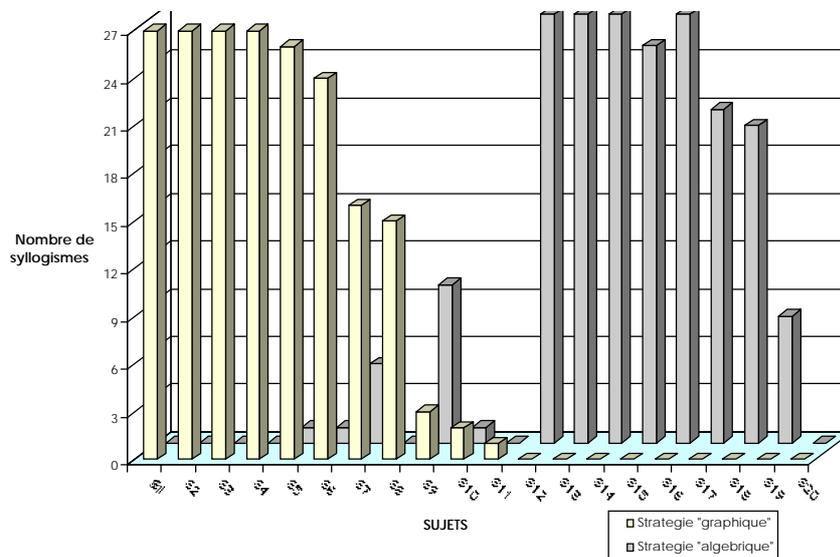


Figure 5.5. Nombre de syllogismes résolus avec stratégie graphique ou algébrique, par sujets (d'après Ford, 1995)

En conclusion, l'analyse de Ford suggère que la théorie des modèles mentaux confond deux types de populations différentes. Une fois que l'on partage les deux groupes, il n'y a plus besoin des modèles mentaux. Bien que la théorie des modèles mentaux reste le point de référence principal dans la littérature sur le raisonnement prédicatif, la recherche des théories alternatives s'impose. Des théories alternatives existent bien – en fait, plus que l'on en souhaiterait. En prenant les modèles mentaux comme point de référence, on peut les partager en deux grands champs. D'une part, il y a des théories qui partagent avec les modèles mentaux la thèse que le raisonnement prédicatif ne s'appuie pas sur des représentations logico-linguistiques, mais elles se différencient des modèles mentaux par la nature des processus et des stratégies qu'elles attribuent aux sujets (Chater & Oakford, 1999 ; Bechtel & Abrahamsen, 1991 ; Evans *et al.*, 1989). Pour ces théories, le raisonnement est plus un ensemble de stratégies heuristiques non systématiques, plutôt qu'un processus de déduction au sens propre. D'autres théories partagent avec les modèles mentaux la thèse que les sujets raisonnent logiquement, mais elles soutiennent par ailleurs que les représentations sur lesquelles le raisonnement est fondé sont plus proches de la forme logique comme les philosophes et les sémanticiens la décrivent que d'une simple inspection d'échantillons individuels représentant isomorphes des situations réelles (Rips, 1984 ; Braine, 1998 ; Sperber & Wilson, 1995 ; Politzer, 1995). Dans les pages qui suivent, nous présenterons les éléments essentiels de ces deux types de théories.

#### **5.4. Le modèle des heuristiques probabilistes du raisonnement logique**

##### **5.4.1. Description du modèle**

Dans la littérature récente, l'étude des capacités de raisonnement sur des problèmes non déductifs, demandant des estimations de probabilité, est devenue un argument central. Parmi les myriades de positions prises dans ce domaine, un point de vue nous intéresse plus particulièrement en raison de sa radicalité et de ses conséquences pour les théories psychologiques du raisonnement déductif. Chater & Oaksford (1999) ont soutenu que tout raisonnement, y compris le raisonnement apparemment déductif, est probabiliste. Etant donné que « les gens ne traitent même pas le raisonnement syllogistique comme une tâche logique, mais ils l'assimilent à leurs stratégies quotidiennes de raisonnement probabiliste » (p. 200), les auteurs ont essayé de développer une théorie entièrement probabiliste du raisonnement syllogistique.

D'après les auteurs, les théories du raisonnement psychologique, qui font référence à des modèles logiques, commettent une erreur fondamentale : « Le raisonnement de tous les jours ne peut pas être fondé sur la logique » (Chater & Oaksford, 1999, p. 193) car il est toujours en état d'incertitude. Au lieu de

procédures logiques, Chater & Oaksford soutiennent que les sujets utilisent des stratégies probabilistes, des heuristiques « rapides et frugales » qui leur permettent de tirer des conclusions sans posséder des capacités logiques. Pour les syllogismes, ils en proposent trois. L'heuristique « MIN » dicte de choisir comme quantificateur pour la conclusion le quantificateur contenu dans la prémisse la moins informative ; par exemple, « MIN » amène à choisir une conclusion avec le quantificateur « Quelques », si les deux prémisses contiennent respectivement les quantificateurs « Tous » et « Quelques », car un énoncé existentiel est moins informatif qu'un énoncé universel. L'heuristique « *conséquence probabiliste* » sélectionne, si MIN ne suffit pas, le deuxième meilleur candidat pour une conclusion ; celui-ci serait alors un énoncé impliqué probabilistiquement par la prémisse la moins informative. L'heuristique « *attachement* » dicte de choisir comme sujet de la conclusion le sujet de la prémisse la moins informative, si celui-ci est A ou C (c'est-à-dire, s'il n'est pas le terme à éliminer).

En outre, Chater et Oaksford supposent que deux autres heuristiques « rapides et frugales » agissent comme filtres sur les conclusions suggérées par les trois heuristiques mentionnées ci-dessus, en évaluant leur « bonne santé probabiliste ». L'une est l'heuristique « MAX » ; elle dicte d'accepter la conclusion en fonction du degré d'information que celle-ci offre par rapport à la prémisse la plus informative. L'autre est l'heuristique « O » ; elle dicte d'éviter les conclusions de forme « O » qui, étant très peu informatives, ont peu de chance d'être d'une utilité quelconque pour le sujet.

Dans une méta-analyse de plusieurs expériences sur les syllogismes, Chater & Oaksford montrent que l'ensemble de ces heuristiques (notamment MIN) prédit assez bien les réponses (valides et non valides) des sujets. Si l'on accepte de nombreuses hypothèses que nous ne pouvons pas discuter ici, le modèle d'heuristiques probabilistes proposé par les auteurs rend compte de 80 % de la variance dans les réponses des sujets.

Les heuristiques proposées, si « rapides et frugales » qu'elles soient, apparaissent néanmoins comme une collection plutôt arbitraire de principes. Pour justifier leur origine et leur existence, Chater et Oaksford développent un argument sémantico-computationnel. D'abord, ils proposent une analyse en termes probabilistes des énoncés quantifiés contenus dans les syllogismes, fondée sur la notion de validité probabiliste (P-validité) ; ensuite ils cherchent à montrer que des considérations computationnelles fondées sur la P-validité justifient l'existence des heuristiques proposées.

En simplifiant considérablement, l'idée de l'analyse sémantique proposée est de représenter les prémisses d'un syllogisme avec des valeurs de probabilités (conditionnelles ou conjointes) entre leurs prédicats, selon la relation spécifiée par

les quantificateurs. Ainsi, « Tous les A sont B » peut être exprimé comme la probabilité conditionnelle que si quelque chose est A, elle soit aussi B ; le quantificateur universel force cette probabilité à être maximale ( $P(B/A) = 1$ ). Les énoncés des autres modes peuvent aussi être traduits en valeurs de probabilité : « Quelques A sont B » exprime le fait que la probabilité que un objet soit A et B est non nulle ( $P(A, B) > 0$ ) (et que il y a des objets qui sont A) ; « Quelques A ne sont pas B » exprime le fait que la probabilité que si quelque chose est A elle soit aussi B n'est pas maximale ( $P(B/A) < 1$ ) (et qu'il y a des objets qui sont A mais qui ne sont pas B) ; et « Aucun A n'est B » exprime le fait que la probabilité que quelque chose soit A et B est nulle ( $P(A, B) = 0$ ).

Si l'on suppose que les termes A et C d'une conclusion sont indépendants, on peut déterminer quelles contraintes les probabilités des prémisses imposent sur la probabilité de la conclusion ; c'est-à-dire, on peut calculer comment les probabilités de A par rapport à B et de B par rapport à C jouent sur la probabilité de A par rapport à C. La conclusion d'un syllogisme sera impliquée probabilistiquement (P-impliquée) par ses prémisses si la probabilité qu'elle exprime est conséquence des probabilités conjointes des prémisses.

Par exemple, comme « Tous les A sont B » et « Tous les B sont C » correspondent aux probabilités  $P(B/A) = 1$  et  $P(C/B) = 1$ , nous pouvons déterminer que aussi  $P(C/A) = 1$ . Donc, « Tous les A sont C » est une conséquence probabiliste des prémisses, car la traduction probabiliste de cet énoncé est précisément  $P(C/A) = 1$ . Par contre, « Quelques A ne sont pas C » ne l'est pas, car son contenu probabiliste,  $P(C/A) < 1$ , est incompatible avec les contraintes que les deux prémisses conjointes imposent sur la probabilité conditionnelle de C si A. En revanche, « Quelques A sont C » est aussi P-impliqué par les deux prémisses de ce syllogisme, car son contenu probabiliste est que  $P(C/A)$  soit supérieur à 0, ce qui est évidemment impliqué par le fait que  $P(C/A) = 1$ . Ce dernier exemple montre aussi que des prémisses peuvent P-impliquer plusieurs conclusions à la fois.

Pour développer leur argument, Chater & Oaksford calculent les énoncés qui sont P-conséquences des prémisses des différents syllogismes, et ils montrent que les heuristiques qu'ils proposent s'accordent bien avec la notion de conséquence probabiliste : elles amènent à choisir des conclusions qui sont aussi P-valides. Par exemple, dans une analyse des conclusions P-valides, Chater & Oaksford montrent que dans vingt-deux cas sur trente-et-un les conclusions suggérées par l'heuristique MIN sont aussi P-valides. Donc, ils en concluent que les heuristiques proposées sont après tout justifiées car elle donnent aux sujets des outils qui se révèlent optimaux du pur point de vue d'une théorie probabiliste de la signification.

#### 5.4.2. *La plausibilité de l'approche probabiliste*

Pour évaluer cette théorie, il faudrait rechercher la vraie origine de son excellent pouvoir prédictif. Une première observation vient à l'esprit. Si l'on considère les heuristiques MIN et *attachment*, une fois dénuées du jargon probabiliste, le lecteur pourra facilement retrouver des idées fort similaires à celles de Ford : les conclusions préférées ont la forme logique de l'une des prémisses (une forme plus faible de MIN, mais dans la majorité des cas équivalente), et elles ont les termes A et C dans la même position que dans les prémisses (ce que implique *attachment*). Donc, dans ce modèle probabiliste, une des parties fondamentales d'une théorie psychologique des syllogismes – comment les sujets aboutissent à leurs conclusions – ne semble pas trop faire appel à des notions probabilistes. Les heuristiques de filtrage, MAX et O ne semblent pas non plus faire un emploi direct de la probabilité, mais elles se fondent plutôt sur la notion de contenu informatif. Le rôle de la probabilité dans la théorie reste donc encore à trouver.

Essayons donc d'en trouver un. L'appel à des notions probabilistes apparaît à deux endroits : dans la définition de la notion de P-conséquence, invoquée par l'heuristique *conséquence probabiliste*, et dans la définition du contenu informatif, employée dans MIN, MAX, O et *attachement*.

Toutefois, la notion de conséquence probabiliste n'a pas de rôle pour expliquer ou prédire les réponses des sujets. Chater & Oaksford soutiennent, non pas que les sujets effectuent les calculs de probabilité qui seraient nécessaires pour déterminer si une conclusion est P-Valide ou non, mais simplement que les heuristiques « rapides et frugales » peuvent être justifiées indirectement par la notion de P-Validité, car les conclusions qu'elles suggèrent sont dans la plupart des cas aussi P-Valides. Mais cet argument est à double tranchant : si P-validité et heuristiques s'accordent si bien, cela peut être une raison d'abandonner l'une des deux. Et dans ce cas, c'est la notion de conséquence P-Valide qui doit s'effacer. La notion de conséquence probabiliste est une extension propre de la notion de conséquence logique classique ; donc toutes les conclusions logiques à partir des prémisses syllogistiques sont aussi P-valides ; donc, par définition, la notion de conséquence probabiliste ne peut pas expliquer les raisonnements des sujets mieux que la notion de conséquence logique classique qui, on le sait bien, est complètement inadéquate. Par conséquent, dans l'explication proposée par Chater & Oaksford, ce n'est pas la notion de probabilité qui permet de prédire le comportement des sujets ; ce sont plutôt les heuristiques proposées, lesquelles, vraies ou fausses en elles mêmes, ne demandent pas de notions probabilistes.

En ce qui concerne le contenu informatif, Chater & Oaksford renvoient à la théorie classique de l'information, qui effectivement traite le contenu d'un énoncé comme une fonction inverse de sa probabilité. Toutefois, on sait depuis longtemps

que cette notion de contenu informatif n'est ni la seule disponible pour le langage ni la plus appropriée (Sperber & Wilson, 1995). Un énoncé comme, par exemple, « Quelques hommes sont forts » peut être considéré comme moins informatif que « Tous les hommes sont forts » car, comme le suggère le modèle probabiliste, la vérité du premier est plus probable que la vérité du deuxième. Mais on peut aussi, sans parler de probabilité, le regarder comme moins informatif parce que, dans des domaines non vides, le premier est tout simplement logiquement impliqué par le deuxième. Plus généralement, si une théorie doit rester proche du « raisonnement de tous les jours », comme Chater & Oaksford le déclarent, alors il vaut mieux reconnaître que l'informativité d'un énoncé n'est pas une propriété fixe de sa forme logique, mais se détermine en fonction de sa pertinence. Ainsi l'énoncé « Quelques condamnés à la peine de mort sont innocents » est bien en ce mode O qui, pour Chater & Oaksford, a un contenu informatif insignifiant (ce qui fonde l'heuristique O), mais il est hautement informatif. Ironiquement, l'analyse probabiliste de Chater & Oaksford est encore plus éloignée du « raisonnement de tous les jours » que les théories qu'elle est censée remplacer.

Pour terminer, rappelons que Chater & Oaksford admettent que « des procédures de test apparemment logiques semblent effectivement avoir une certaine influence, à côté des procédures de génération, dans la détermination des réponses des sujets » (Chater & Oaksford, 1999, p. 214). La présence de ces procédures, que le modèle d'heuristiques probabilistes ne peut pas expliquer par hypothèse, montre que même en acceptant le modèle probabiliste, le problème de chercher une théorie psychologique des procédures logiques dont les sujets disposent pour raisonner reste encore entier. Mieux vaut donc prendre directement le taureau par les cornes.

## **5.5. Les théories des règles mentales pour le raisonnement prédicatif**

### **5.5.1. Principes généraux**

Les théories alternatives aux modèles mentaux qui maintiennent que les sujets raisonnent logiquement, mais abandonnent l'idée qu'ils construisent des modèles mentaux, sont essentiellement des variations de l'hypothèse dite des règles mentales. Rips (1994) a montré comment son modèle computationnel, conçu pour la logique propositionnelle, peut être élargi à la logique prédicative ; ainsi, Braine (1998) et ses collaborateurs (Yang *et al.*, 1998) ont récemment présenté une tentative ambitieuse d'élargir la théorie de la logique mentale propositionnelle au cas du raisonnement prédicatif. Avant de détailler les deux théories et leurs différences, rappelons certains points et problèmes qu'elles partagent.

D'abord, les théories des règles mentales partagent la vision du domaine d'une théorie du raisonnement déductif. Elles ne se proposent pas comme des théories globales de la pensée ou de la signification (comme veulent l'être les théories des

modèles mentaux : voir (Johnson-Laird, 1983 ; Garnham, 1987)). Elles n'ont pas l'ambition d'expliquer la signification des messages linguistiques, au-delà de leur partie proprement logique, ou la construction des structures de mémoire. Elles laissent volontiers la place à la sémantique pour étudier la signification et à la pragmatique pour déterminer comment un message (linguistique ou non) assume une signification particulière selon le contexte d'émission et la tâche des sujets. Elles affirment, en revanche, qu'une fois que sémantique et pragmatique ont sélectionné un certain message, les sujets en sélectionnent une forme logique (ou une série d'énoncés écrits dans le langage de la pensée, comme Fodor (1975) l'appellerait), sur laquelle des règles de dérivation en forme de déduction naturelle peuvent s'appliquer. Ces règles, censées être innées et universelles (au moins dans la version de Braine), appliquées par une procédure de démonstration définie, génèrent des démonstrations qui construisent une « route mentale » amenant d'un ensemble des prémisses à des conclusions. L'ambition de la théorie est de décrire cette route.

Donc, une théorie de règles mentales demande de spécifier trois éléments : la syntaxe dans laquelle les représentations des messages sont formulées ; les règles de raisonnement qui entrent dans les démonstrations ; la procédure d'application des règles. Plus précisément, une théorie des règles mentales est censée développer les deux derniers aspects, et c'est dans les détails de ces développements que les propositions de Braine et Rips se différencient. Cependant, les théories s'accordent sur la stratégie prédictive fondamentale : le niveau de difficulté et les erreurs des sujets doivent être liés à la structure de la démonstration.

Dire que les sujets construisent une démonstration quand ils raisonnent ne signifie pas du tout affirmer que les sujets ne devraient jamais arriver à des conclusions erronées, comme l'ont soutenu plusieurs auteurs (par exemple, (Johnson-Laird, Savary, 1999)). Il y a démonstration et démonstration ; certaines emploient seulement des règles élémentaires et créent des structures linéaires, d'autres demandent de construire des structures complexes avec des sous-démonstrations emboîtées dans la démonstration principale, un peu comme quand en mathématiques on démontre un lemme pour retourner, la démonstration du lemme terminée, à l'énoncé de départ. Dans ce deuxième type de démonstration, qui demande d'ouvrir des « sous-domaines », il faut non seulement garder la trace des passages ponctuels, mais surtout du processus de raisonnement dans son ensemble. Les sujets, qui sont sensibles à la structure des raisonnements (Marcus, 1982 ; Rips, 1994, chap. 5), peuvent facilement s'y perdre. Il y a donc ample possibilité dans ces théories d'envisager que les sujets puissent ne pas arriver à compléter une démonstration, ou perdre trace de sa structure. Dans ces cas, les sujets peuvent bien arriver à des conclusions logiquement fausses, ou même contradictoires si la structure de la démonstration le permet.

Les théories des règles mentales partagent aussi une vision de la relation entre logique mentale et logique standard. Elles ne soutiennent pas que la logique que le sujet emploie en raisonnant est la logique classique ; au contraire, leur but est de déterminer exactement laquelle, parmi la classe infinie des options différentes, est effectivement employée dans les processus psychologiques. Les théories de Braine et Rips concordent quand même sur le fait que la classe des conclusions psychologiquement valides est un sous-ensemble de celle définie par la logique classique, et donc ne contient pas de contradictions. Comme on l'a vu, cette thèse n'est pas incompatible avec le fait que les sujets font souvent des erreurs ; il n'est donc pas nécessaire de supposer que les erreurs de raisonnement signalent l'absence d'une logique mentale, ou la présence d'une logique mentale non cohérente (comme suggéré par exemple par (Johnson-Laird, Legrenzi *et al.*, 2000a)).

Les théories ne supposent pas non plus que la logique prédicative standard décrive correctement la syntaxe des formes logiques psychologiquement réelles : la nature même de la syntaxe de la pensée est une question empirique. En revanche, à la différence des modèles mentaux, elles partagent avec la logique standard la thèse que ces formes logiques contiennent des structures abstraites de nature algébrique : des variables de types différents, des quantificateurs et des opérations de substitution d'individus à variables.

Cependant, ce dernier point fait apparaître une difficulté commune à toute théorie des règles mentales prédicatives. Ce qui rend le cas prédicatif beaucoup plus difficile que le cas propositionnel est la spécification des formes logiques sur lesquelles le raisonnement prédicatif est défini. Le problème n'est pas seulement propre aux théories des règles mentales, mais dans une large mesure il concerne la sémantique du langage naturel dans son ensemble. S'il est excessif de demander à une théorie du raisonnement de résoudre tous les nombreux problèmes qui occupent les sémanticiens de la langue naturelle (Cherchia & McConnell-Ginet, 2000), au moins deux questions doivent être posées. Une théorie de la syntaxe des représentations prédicatives adéquate pour une théorie du raisonnement doit spécifier le type de quantificateurs sur lequel les règles sont définies et leur position à l'intérieur d'un énoncé en forme logique.

Le premier problème dérive du fait que le langage naturel contient de nombreux types de quantificateurs, qui ont aussi des fonctions différentes. Par exemple, un énoncé comme « Les chiens ont quatre pattes » est utilisé pour exprimer quelque vérité sur notre conception des chiens en général : nous n'abandonnerions pas cet énoncé si nous rencontrions un chien sans pattes. Mais un énoncé comme « Tous les chiens ont quatre pattes » est ambigu : certains ne considèrent pas qu'un chien sans pattes le réfute, d'autres le pensent. Pour sortir de cette ambiguïté, qui concerne la pensée exprimée avec un énoncé quantifié universellement, réduisons la portée d'une théorie de la logique naturelle aux *quantificateurs en lecture extensionnelle* :

quantificateurs qui sont censés varier sur tous les objets d'un domaine. Ce choix nous laisse encore deux possibilités : nous pouvons concevoir les quantificateurs comme *non restreints*, ou comme *restreints* au domaine du discours spécifié par des indices linguistiques ou pragmatiques. Pour suivre l'exemple précédent, nous pouvons penser à « Tous les chiens ont quatre pattes » comme à un énoncé qui parle de tous les chiens existants ; dans ce cas, le quantificateur serait non restreint, équivalent à la représentation logique standard « Pour tous les  $x$ , si  $x$  est un chien,  $x$  a quatre pattes ». Mais on pourrait considérer l'énoncé comme une ellipse pour « Tous les chiens dans cette pièce ont quatre pattes » ou « Tous les chiens dans cette ferme ont quatre pattes », selon que l'univers du discours est cette pièce ou cette ferme. L'énoncé en lui-même ne spécifie pas l'univers du discours, mais son contexte d'émission le fera ; dans ce cas le quantificateur aurait une lecture restreinte. Plusieurs arguments font penser que les quantificateurs du langage naturel se comportent comme des quantificateurs restreints (Bonatti, 1998).

De cette double option vient la possibilité de développer la représentation logique sur laquelle les règles d'une logique naturelle sont définies de deux façons différentes, comme l'ont fait Rips et Braine. Les deux théoriciens concordent sur le traitement extensionnel des quantificateurs, mais choisissent des options différentes pour leur lecture restreinte ou générale.

Le problème de la position des quantificateurs vient d'un décalage entre représentations logiques et langues naturelles. Dans une logique standard, les quantificateurs sont externes aux énoncés et séparés des prédicats qu'ils quantifient. Toutefois, aucune langue naturelle étudiée ne présente cette configuration (Ioup, 1975) : les quantificateurs sont toujours proches de leurs prédicats de référence. Cette régularité semble être trop forte pour être un pur hasard ; donc pour représenter les quantificateurs, une forme logique naturelle devrait trouver une solution qui respecte cette contrainte positionnelle. Les théories de Braine et Rips concordent sur le fait qu'une forme logique satisfaisante doit satisfaire cette contrainte, mais choisissent des solutions différentes.

Nous pouvons maintenant examiner les deux théories de façon plus détaillée.

### **5.5.2. PSYCOP : un modèle computationnel du raisonnement prédicatif**

#### *5.5.2.1. Les hypothèses de PSYCOP*

PSYCOP, le système de raisonnement par déduction naturelle proposé par Rips (1994), élargit le travail de (Rips, 1983), veut rendre psychologiquement plausible l'application au domaine du raisonnement d'une série de techniques de démonstration bien connues en intelligence artificielle. Cela est possible grâce à une série d'hypothèses, à la fois computationnelles et psychologiques, sur la forme des

règles, les structures de mémoire et les structures de contrôle, que nous présenterons brièvement.

PSYCOP suppose que les règles de raisonnement ont la forme classique condition-action des systèmes de production : chaque règle a toujours la structure « Si (liste de conditions à satisfaire) alors (liste d'actions à entreprendre) ». Les règles sont de deux types : proactives (*forward*) et rétroactives (*backward*). Les règles proactives s'appliquent à des énoncés déjà présents dans le domaine d'une démonstration et permettent d'en ajouter d'autres dans le même domaine. Par exemple une application élémentaire d'un *modus ponens* (élimination de l'implication) contrôle si les deux énoncés « Si A alors B » et « A » sont déjà dans la démonstration, et en cas affirmative ajoute « B » à la liste des énoncés acceptés. Donc les conditions de cette règle sont : 1. Contrôler que « Si A alors B » soit présent dans un domaine D ; 2. Contrôler que « A » soit présent dans le même domaine D ; et 3. Contrôler que « B » ne soit pas déjà dans D ; et la liste d'actions contient la seule action « Ajouter B au domaine D ».

Les règles rétroactives réalisent une stratégie de démonstration différente. Elles « regardent en arrière », essayent de déterminer de quels énoncés on aurait besoin pour garantir la dérivation d'une certaine conclusion, et cherchent à assurer qu'ils soient démontrés s'ils peuvent l'être. Pour continuer l'exemple du *modus ponens*, la règle rétroactive d'élimination de l'implication part de la conclusion « B » et demande de contrôler si « Si A alors B » est déjà présente dans le domaine ; si oui, elle met en place une procédure pour tenter de démontrer « A », ce qui garantirait B en cas de réussite. Dans ce cas, la liste des conditions contient : 1. Fixer le domaine D comme domaine du but courant ; 2. Fixer B comme but courant ; 3. Contrôler que « Si A alors B » est présent dans le domaine D ; 4. Contrôler que « A » n'est pas déjà un sous-but dans D (c'est-à-dire, qu'il n'y a pas déjà un sous-domaine de D où on est en train de chercher une démonstration de A). Si les conditions 3 et 4 sont vérifiées, l'action suivante sera entreprise : « Ajouter A à la liste des sous-buts » (ou, de façon équivalente, ouvrir un sous-domaine dans D où chercher une démonstration de A).

Déjà dans la formulation même de la différence, il paraît clair que les règles proactives placent des contraintes différentes sur le système cognitif. Cette différence en poids se reflète dans les présuppositions sur la structure de la mémoire, qui n'est pas considérée comme un réceptacle indifférencié d'énoncés, mais comme un (ensemble d') espace(s) dans le(s) quel(s) les énoncés sont organisés par liens de *déduction* et de *dépendance*. Imaginons que les prémisses d'un raisonnement ouvrent un espace de mémoire, dans lequel la démonstration est enregistrée. L'application des règles proactives lie les énoncés concernés en relations de déduction ; les règles proactives demandent donc que la mémoire garde trace de la construction de la démonstration, sans pour autant l'obliger à sortir de cet espace. En

revanche, une règle rétroactive oblige à ouvrir un sous-espace de mémoire séparé dans lequel le démonstrateur cherche à démontrer les énoncés dont la règle a besoin pour garantir la conclusion dans l'espace d'origine. La charge cognitive devient donc tout de suite plus importante, du fait que la mémoire doit non seulement garder trace des liaisons (de déduction) dans le premier espace, mais aussi des liaisons de dépendance entre espaces différents. De plus, un sous-espace peut en ouvrir d'autres, ce qui peut compliquer encore plus une démonstration.

Pour ces raisons, entre autres, PSYCOP suppose aussi que l'application des règles est sujette à un contrôle spécial. La procédure de démonstration appliquera toujours d'abord les règles proactives, qui peuvent donc se déclencher de façon pratiquement automatique, et n'aura recours aux règles rétroactives qu'en deuxième instance ; à leur tour elles seront appliquées selon un ordre de difficulté déterminé par la quantité de sous-domaines dont elles ont besoin. Récursivement, dans chaque sous-domaine la procédure appliquera d'abord les règles proactives, ensuite si nécessaire les rétroactives, et ainsi de suite. Ces hypothèses sur le contrôle peuvent être aussi mises en relation avec la différence psychologique entre génération spontanée de conclusions et évaluation de conclusions données. Quand les sujets doivent générer des conclusions de façon spontanée, ils n'ont pas de buts explicites, tandis que quand ils doivent évaluer une conclusion, le but est déjà donné : celui d'évaluer la conclusion. Ainsi, on peut rapprocher la tâche de génération spontanée avec l'application de règles proactives, qui ne demandent pas de buts donnés. En revanche, dans une tâche d'évaluation le fait même que le but est posé peut inviter les sujets à utiliser des règles rétroactives non disponibles autrement. PSYCOP peut donc prédire quelles conclusions tendront à être générées spontanément et quelles conclusions données seront acceptées, dans la logique prédicative comme dans celle propositionnelle.

#### 5.5.2.2. *L'extension de PSYCOP au raisonnement prédicatif : élimination des quantificateurs*

Venons-en maintenant à l'aspect plus proprement prédicatif de PSYCOP. L'idée de base de Rips est de capitaliser au maximum sur l'aspect propositionnel de PSYCOP, évitant d'avoir des règles de dérivation définies sur les quantificateurs, et ajoutant le minimum de règles nouvelles définies sur le raisonnement proprement prédicatif<sup>4</sup>. Pour obtenir cela, deux problèmes doivent être résolus. Le premier est

---

4. Les seules règles nouvelles proprement prédicatives du système de Rips sont trois règles définies sur énoncés avec variables libres, qui expriment les propriétés de transitivité, exclusivité et conversion. Respectivement, elles sont :

– transitivité : si deux énoncés de forme « Si A(x) alors B(x) » et « Si B(y) alors C(y) » sont valides dans un domaine D, et si l'énoncé « Si A(z) alors C(z) » ne l'est pas encore, alors ajoute-le dans D ;

général : comment éliminer les quantificateurs. Le deuxième est plus technique : comment lier les énoncés sans quantificateurs avec les règles pour les connecteurs propositionnels en sorte qu'elles donnent des résultats corrects. Le premier problème peut être résolu facilement grâce à une astuce logique connue sous le nom de Skolemisation, d'après le logicien Skolem qui l'a introduite. Toutes les fois que l'on rencontre un quantificateur universel, on l'élimine en laissant la variable libre, et toutes les fois que l'on rencontre un quantificateur existentiel, on l'élimine et on substitue à la variable correspondante un nom (une constante) temporaire ; il faut simplement faire attention à ne pas mélanger les noms et les variables, en introduisant de nouveaux noms pour chaque quantificateur existentiel éliminé et de nouvelles variables pour chaque quantificateur universel éliminé. Ainsi, un énoncé comme « Tous les hommes sont bons » est représenté en logique comme «  $\forall x ((\text{Homme}(x) \Rightarrow \text{Bon}(x)))$  », et en PSYCOP il est transformé tout simplement en « Si Homme(x) alors Bon(x) » ; ainsi, « Quelques hommes et quelques femmes sont sympa » s'exprime en logique avec «  $\exists x \exists y ((\text{Homme}(x) \& \text{Sympa}(x)) \& (\text{Femme}(y) \& \text{Sympa}(y)) \& (x \neq y))$  » autant que dans PSYCOP il est représenté avec deux noms quelconques à la place des existentiels, ainsi : « (Homme(a) & Sympa(a)) & (Femme(b) & Sympa(b) & (a  $\neq$  b)) ».

L'élimination des quantificateurs a aussi l'effet secondaire de maintenir les prédicats proches de leur place d'origine dans l'énoncé correspondant de la langue naturelle, ce qui réalise un des desiderata d'une forme logique psychologiquement plausible. Donc dans PSYCOP les quantificateurs sont invisibles, proches de leurs prédicats, et ils ont une portée non restreinte. En revanche, elle crée un deuxième problème, que nous ne pouvons pas expliquer dans le détail ; il s'agit de s'assurer qu'avec l'introduction des noms temporaires les règles s'appliquent correctement même avec des passages entre domaines et sous-domaines, sans que des substitutions incorrectes entre variables libres, noms temporaires et noms permanents entraînent le démonstrateur vers des routes logiquement incorrectes. Le problème ne concerne pas les règles proactives, mais il surgit dans toutes les règles rétroactives, que Rips doit donc modifier pour éliminer l'émergence de possibles erreurs de substitution.

En somme, à l'aide de quelques adaptations techniques, Rips obtient une extension de PSYCOP au raisonnement prédicatif sans en modifier la construction de base.

---

– exclusivité : si deux énoncés de forme « Si A(x) alors B(x) » et « Non (B(y) et C(y)) » sont valides dans un domaine D, et si l'énoncé « Non (A(z) et C(z)) » ne l'est pas encore, alors ajoutez-le dans D ;

– conversion : si un énoncé de forme « Non (A(x) et B(x)) » vaut dans un domaine D, et si énoncé « Non (B(y) et A(y)) » ne l'est pas encore, alors ajoutez-le dans D.

### 5.5.2.3. *Le traitement des syllogismes et les données empiriques en faveur de PSYCOP*

PSYCOP peut démontrer un sous-ensemble des énoncés démontrables dans une logique classique, et ne peut démontrer aucun énoncé non démontrable classiquement. Or, plusieurs inférences prédicatives ne sont pas logiquement valides, comme par exemple l'inférence de « Tous les A sont B » à « Quelques A sont B », valide seulement si on accepte des présuppositions existentielles des prédicats, ou l'inférence, logiquement invalide mais pragmatiquement correcte, de « Quelques A sont B » à « Quelques A ne sont pas B ». Toutefois, nous ne sommes pas là devant un problème majeur pour PSYCOP, comme par ailleurs pour d'autres théories logiques. Si l'on rend explicites les présuppositions existentielles et pragmatiques contenues dans les énoncés quantifiés, PSYCOP peut traiter le raisonnement syllogistique sans modifications ultérieures. Rendre explicites ces présuppositions signifie parfois devoir représenter les énoncés de forme A, E, I, O non pas avec un seul énoncé en forme logique, mais avec deux. Ainsi, pour rendre explicites les présuppositions existentielles, « Tous les A sont B » peut être représenté par le couple « Si A(x), alors B(x) » et « A(a) & B(a) », et « Aucun A n'est B » peut être représenté par le couple « Non (A(x) & B(x)) » et « A(a) & Non (B(a)) ». De même, pour expliciter les présuppositions pragmatiques, « Quelques A sont B » pourra être traduit par le couple « A(b) & B(b) » et « A(a) & non (B(a)) » ; et « Quelques A ne sont pas B » par le couple « A(b) & non (B(b)) » et « A(a) & B(a) ». Cette solution n'affirme pas que tous les sujets représenteront invariablement les présuppositions des énoncés A, E, I, O sous forme explicite, mais que, avec une certaine probabilité, ils pourront le faire.

Nous pouvons maintenant présenter quelques arguments que Rips a avancés en faveur de sa théorie. Un des points centraux de PSYCOP est la différence entre règles proactives et rétroactives, avec le corollaire psychologique que les règles proactives (qui ne demandent pas de créer des sous-buts) sont plus simples et plus facilement disponibles. Donc, les syllogismes résolubles avec la seule application de règles proactives devraient générer plus de réponses correctes par rapport aux syllogismes qui demandent des règles rétroactives. Dans des expériences originales (Rips, 1994, chap. 7) et à travers une analyse des données existantes, Rips a montré que les données sont cohérentes avec cette prédiction. PSYCOP peut déduire quinze syllogismes ; dans une tâche de validation de conclusions, ceux avec règles proactives suscitent 89 % de réponses correctes, tandis que pour ceux qui demandent des règles rétroactives le pourcentage de réponses exactes est de 80 %. Mais dans une tâche de validation de conclusion qui présente déjà aux sujets un but, il est probable que l'activation de règles rétroactives soit facilitée ; avec une tâche de génération de conclusions dans laquelle le sujet devrait employer surtout des règles proactives qui ne réclament pas de buts, les différences devraient être plus extrêmes. En fait, une réanalyse des résultats de production spontanée de conclusions collectés

par (Johnson-Laird & Bara, 1984) révèle qu'avec les syllogismes qui demandent les seules règles proactives pour générer une conclusion valide les conclusions que les sujets proposent sont correctes dans 75 % des cas, alors qu'elles se réduisent à 32 % pour les syllogismes qui génèrent des conclusions valides seulement en appliquant les règles rétroactives.

Un deuxième type d'argument proposé par Rips, que nous présentons sous forme simplifiée, est très intéressant même s'il est très spéculatif. Il s'agit de lier les réponses des sujets à une estimation de la probabilité de trouver une démonstration dans PSYCOP, avec certaines assumptions sur la disponibilité des règles de base<sup>5</sup>. D'après un modèle par règles, quand les sujets évaluent la conclusion d'un problème, ils cherchent une démonstration de la conclusion à partir des prémisses. Dans ce type de modèles, la probabilité de trouver une démonstration doit être fonction de la probabilité d'appliquer ensemble toutes les règles nécessaires pour obtenir la démonstration. En estimant la probabilité d'appliquer chaque règle (estimations que Rips obtient à partir des données), on pourra calculer la probabilité de trouver une démonstration comme étant égale à la probabilité conjointe des  $n$  règles qui doivent être appliquées pour l'obtenir.

Toutefois, d'autres facteurs contribuent à la probabilité de trouver une démonstration : d'abord, le fait qu'à partir d'un groupe de prémisses on puisse trouver plusieurs démonstrations ; ensuite, le fait que les sujets puissent accepter les implications existentielles ou pragmatiques, dans la conclusion comme dans les prémisses, ce qui modifie la structure logique des problèmes ; enfin, le fait que les sujets puissent donner une réponse correcte simplement parce qu'ils devinent la réponse, et non parce qu'ils ont trouvé une démonstration. Heureusement, on peut aussi estimer ces facteurs avec PSYCOP et, en les intégrant dans le modèle, nous pourrions finalement estimer la probabilité que le sujet réponde qu'une conclusion est nécessaire comme la somme de la probabilité qu'il trouve une des démonstrations et de la probabilité qu'il réponde « Oui » quand il ne la trouve pas mais tire au hasard la réponse correcte. Une fois pris en compte tous ces facteurs, Rips montre que la corrélation entre les réponses « Oui » observées pour les 256 syllogismes et les probabilités de réponses « Oui » estimées par le modèle est très solide ( $r = 0,9$ ), ce qui montre que PSYCOP peut se prêter à la formulation de prédictions assez élaborées sur les syllogismes.

Quelles réponse PSYCOP peut-il donc proposer aux quatre questions sur les syllogismes posées au début de ce chapitre ? La réponse à la question 1 est en

---

5. Il ne faut pas confondre cette utilisation de la probabilité avec celle de Chater & Oaksford. Pour Rips, quand un sujet répond correctement à un problème, il le fait par un raisonnement entièrement déductif et non probabiliste, même si l'on peut évaluer la probabilité qu'il trouve une telle réponse correcte. En revanche, pour Chater & Oaksford, c'est le processus de raisonnement même qui est probabiliste.

quelque sorte construite directement à l'intérieur d'une théorie des règles mentales. Le sentiment que certaines conclusions s'imposent en vertu de la relation entre les quantificateurs contenus dans les prémisses reflète le fait que, pour une classe identifiable de syllogismes, les sujets peuvent construire une démonstration logique de la conclusion, largement grâce aux mêmes procédures du raisonnement propositionnel, une fois les quantificateurs éliminés comme suggéré par PSYCOP. Les questions 2 et 3 peuvent être résolues en rappelant deux facteurs : les différences entre règles proactives et rétroactives, et les différents taux de disponibilité des règles. Par exemple, pour le problème (i) (« Tous les A sont B » ; « Tous les B sont C » ; donc, « Tous les A sont C ») PSYCOP estime que la probabilité de trouver la conclusion est de 89 %, tandis que pour le problème (iii) (« Tous les B sont A », « Aucun B n'est C » ; donc, « Quelques A ne sont pas C ») la probabilité baisse à 15 %. Ajoutons que le problème (i) demande seulement d'éliminer les quantificateurs et d'appliquer des règles proactives, alors que pour résoudre (iii) il faut employer des règles rétroactives. Ainsi, le problème (i) sera à la fois mieux résolu que (iii) et perçu aussi comme plus facile. En fait, dans les données de Rips, alignées sur celles trouvées par d'autres chercheurs, (i) présente un taux de réponses correctes de 90 %, alors que les réponses correctes à (iii) ne sont que de 10 %.

Il nous reste à répondre à la question 4, l'effet de figure. Pour cela, nous pouvons généraliser une observation de Ford (1995) : à bien voir, l'explication proposée par Johnson-Laird & Bara (1984) d'après laquelle l'effet est généré par la façon dont l'information est mise et récupérée en mémoire, ne demande pas de modèles mentaux. Elle peut s'appliquer à toutes théories qui postulent que les représentations mentales sur lesquelles le raisonnement s'effectue sont de nature linguistique, car les prédicats contenus dans les énoncés sont aussi chargés et récupérés en mémoire pendant l'application des règles. Si l'explication est la bonne pour les modèles, elle le sera aussi pour des représentations linguistiques.

En conclusion, PSYCOP semble donc pouvoir proposer un modèle pour le raisonnement prédicatif aussi valide que celui du raisonnement propositionnel. La théorie de Rips peut expliquer la majorité des données connues sur le syllogisme et de plus elle génère des prédictions originales. Soigneusement définie dans ses propriétés formelles, les limites de PSYCOP doivent plutôt être recherchées dans un certain manque de motivation psychologique dans les opérations formelles qui en sont à la base. Ainsi, l'élimination des quantificateurs en faveur des variables libres est une solution technique dont les bases psychologiques doivent être fournies. Les règles de base mêmes sont en quelque sorte psychologiquement arbitraires ; par ailleurs, la validité psychologique des règles de base n'est pas un des objectifs affichés dans le programme de Rips, qui ne se préoccupe donc pas d'en trouver une confirmation expérimentale. Même pour le fondement psychologique de la distinction entre règles rétroactives et proactives, Rips ne propose que des preuves indirectes, mais il faut encore déterminer si cette distinction concorde simplement

avec les données, ou si elle correspond à des réels processus psychologiques entrepris par les sujets. Par exemple, en dépit de la bonne concordance générale entre prédictions de PSYCOP et résultats expérimentaux sur les syllogismes, plusieurs arguments intuitifs apparemment faciles devraient être classés comme difficiles d'après PSYCOP car ils demandent des règles indirectes. Prenons par exemple, dans le domaine propositionnel, « Il y a un A ; il y a un B ; il y a un C. S'il y a un A, un B et un C, alors il y a un D : que peut-on conclure » ? Pour trouver une conclusion, il faut employer des règles rétroactives qui ne peuvent pas s'activer sans un but défini ; pourtant les sujets trouvent assez facilement la conclusion (Braine, 1998). Cependant, même en considérant ses limites, PSYCOP reste un modèle de la manière dont on peut développer une théorie psychologique de la logique tout en respectant rigueur formelle et contraintes computationnelles.

### 5.5.3. La théorie de la logique mentale prédicative de Braine et O'Brien

La théorie de Rips n'est pas la seule version existante de règles mentales. Braine (1998), O'Brien et ses collaborateurs (1998) ont récemment essayé d'élargir leur théorie de la logique mentale propositionnelle au raisonnement prédicatif. Les présupposés de la théorie de Braine restent les mêmes dans le cas prédicatif comme dans le cas propositionnel, et en grande partie sont partagées aussi par PSYCOP ; nous nous bornerons donc à présenter les points où la nouvelle théorie prédicative s'éloigne de celle de Rips.

On peut identifier trois différences fondamentales entre la proposition de Braine et celle de Rips. La première concerne la syntaxe des représentations logiques ; la deuxième les règles de raisonnement proposées et leur statut ; la troisième concerne la procédure de génération des démonstrations.

#### 5.5.3.1. *Éléments de représentations sémantiques*

D'abord la syntaxe. Braine (1998) rappelle que le système de représentation de la logique standard, largement accepté dans le modèle de Rips, reste très loin du langage naturel. L'observation n'est pas nouvelle ; il existe déjà plusieurs tentatives de créer des systèmes de représentation plus proches du langage naturel et de sa logique.

En général, ces propositions supposent qu'au-delà de l'expression d'un énoncé dans une langue naturelle et de son interprétation, un niveau ultérieur de représentation est construit, qui est souvent appelé *représentation sémantique*. Toutefois, en l'état actuel, les théories de la représentation sémantique proposées ont été plutôt motivées par des considérations linguistiques : aucune n'est psychologiquement motivée. L'ambition de Braine est de fournir des représentations

sémantiques à la fois proches du langage naturel, privées des ambiguïtés quantificationnelles qu'il contient, et motivées psychologiquement.

Une telle représentation est censée respecter le principe de Fodor, d'après lequel « les énoncés que nous sommes capables de comprendre ne sont pas tellement différents des formules qui les représentent au niveau interne » (Fodor, 1975, p. 156). Conformément à ce principe, le système représentationnel devra contenir toute structure syntaxique qui semble apparaître dans toutes les langues du monde. Parmi ces structures, nous trouvons la prédication, qui comme l'avait remarqué déjà Frege a dans le langage naturel le même rôle que l'application d'une fonction à un argument en mathématiques, et avec elle les différences de types logiques qui en découlent. Mais comme structures probablement universelles nous pouvons aussi compter les constructions relatives et génitives, qui d'un point de vue logique servent à restreindre l'univers du discours à des sous-ensembles (de « Tous les amis » on peut passer à « Tous les amis qui ont un chapeau noir », ou à « Tous les amis de Maria »), ou les formes plurielles, qui indiquent qu'un prédicat est satisfait non pas par un seul individu (comme l'est le quantificateur existentiel de la logique standard), mais au minimum par une pluralité, ou surtout les quantificateurs.

La représentation des quantificateurs dans une forme logique qui satisfait les desiderata présentés par Braine pose de nombreux problèmes intriqués, pour lesquels Braine (1998) essaie de proposer des solutions psychologiquement plausibles. Nous rappelons seulement que dans la représentation sémantique proposée par Braine les quantificateurs sont exprimés explicitement (à la différence de PSYCOP), qu'ils restent proches de leur lieu d'origine (à la différence de la logique standard), et qu'ils sont interprétés dans un domaine restreint (à la différence de PSYCOP et de la logique standard).

Pour obtenir que les quantificateurs restent sur place sans générer d'erreur de portée, la représentation comprend un opérateur (PRO) qui agit comme un pronom, ou un indice de coréférence. Par exemple, la phrase « Toutes les billes sont vertes ou rouges » est ambiguë. Elle peut signifier que les billes sont toutes vertes ou toutes rouges (et donc si elle sont toutes rouges il n'y en a aucune qui est verte, et *vice versa*), ou bien que quelques-unes sont bleues vertes, quelques autres sont rouges, et qu'il n'y en a pas d'autres couleurs. Pour éliminer l'ambiguïté entre les deux lectures, tout en gardant les quantificateurs proches de leurs prédicats, le système de Braine permet de construire deux représentations séparées, l'une d'entre elles contenant une référence-PRO au quantificateur introduit dans la première phrase. Ainsi, une représentation possible devient :

Vertes [Toutes les billes] ou Rouges [Toutes les billes],

dans laquelle les quantificateurs sont indépendants et donc permettent la lecture d'après laquelle les billes sont toutes vertes (aucune bille rouge) ou toutes rouges (aucune bille verte) ; l'autre représentation possible sera :

Vertes [Toutes les billes] ou Rouges [PRO-Toutes les billes],

qui laisse les deux quantificateurs explicitement, mais force la lecture d'après laquelle quelques billes peuvent être vertes et d'autres peuvent être rouges.

D'autres restrictions sur la portée des quantificateurs peuvent aussi être obtenues avec un opérateur de relativisation, représenté par le symbole « : », qui permet de représenter la fonction logique fondamentale assurée par la relativisation dans les langues naturelles. Ainsi la structure :

[Tous les P : S[PRO]]

signifie « Tous les P qui sont aussi S », et une formule comme « Transparentes [Toutes les billes : Vertes[PRO]] » signifie que toutes les billes vertes sont transparentes. La possibilité d'exprimer la relativisation est importante pour lier les structures prédicatives aux structures propositionnelles, car un énoncé quantifié comme « Toutes les billes vertes sont transparentes » est logiquement équivalent à un énoncé quantifié qui contient une implication, comme « Toutes les billes, si elles sont vertes, sont transparentes ».

Si l'on veut garder les quantificateurs proches de leur place d'origine, il faut aussi résoudre des problèmes de portée présentés par la négation. Comment peut-on distinguer l'énoncé « Marie n'aime aucun des hommes », qui nous permet de dériver que Marie n'aime pas Jean, de « Marie n'aime pas tous les hommes », à partir duquel on ne peut pas savoir si Marie aime ou n'aime pas Jean ? Dans le système de Braine, la négation (exprimée avec NEG) reste toujours proche de son prédicat, mais on signale les quantificateurs qui sont à l'intérieur de sa portée en les incluant entre deux tildes. Ainsi, le premier énoncé est rendu comme

NEG Aime [Marie] [Tous les Hommes]

(que l'on pourrait interpréter comme « Tous les hommes sont tels que Marie ne les aime pas ») et les règles spécifieront que le quantificateur universel peut admettre toutes substitutions d'hommes, y compris celle de Jean :

NEG Aime [Marie] [Jean]

alors que le deuxième énoncé devient :

NEG Aime [Marie] [~Tous les Hommes~]

(que l'on interprétera comme « Ce n'est pas le cas que tous les hommes sont tels que Marie les aime »), et le quantificateur encadré dans les deux tildes ne permet pas de substitution.

Avec ces éléments de base de la syntaxe du système, nous pouvons maintenant voir comment les énoncés en forme A, E, I, O sont réalisés dans la représentation sémantique proposée par Braine. « Tous les A sont B » est tout simplement rendu comme « B [Tous les A] » ; « Quelques A sont B » devient « B [Quelques A] » ; « Aucun A n'est B » sera « NEG B [ ~Quelque A~ ] » et « Quelques A ne sont pas B » sera « NEG B [Quelques A] ». Ces représentations étant plutôt élémentaires, on peut se demander si l'effort de développer un nouveau système représentationnel vaut la peine. Toutefois, il ne faut pas perdre de vue le fait que nous traitons le syllogisme comme un exemple de raisonnement prédicatif. Toutes les règles du système de logique naturelle s'appliqueront aux représentations de raisonnements avec énoncés A, E, I, O, mais elles iront bien au delà du syllogisme. Une fois le système représentationnel en place, d'autres types de raisonnement pour le moment en dehors de toutes investigations psychologiques seront à la portée de la théorie. Pour présenter un exemple un peu extrême, un énoncé comme « Tous les hommes qui sont amis de quelque amie de Marie sont ennemis de quelqu'un », dont la représentation en termes, par exemple, de modèles mentaux ne serait même pas envisageable, peuvent être exprimés dans la syntaxe de la théorie sans difficulté.

### 5.5.3.2. Règles et programme de dérivation : les règles prédictives et la thématization

Comme on s'y attend, le système de logique mentale prédictive dans la théorie de Braine et O'Brien est une extension du système propositionnel correspondant, avec deux adjonctions importantes : l'une concerne le groupe de règles, l'autre le programme pour leur application. Outre les règles propositionnelles, le système contient aussi des versions prédictives des mêmes règles, nécessaires pour en lier l'application aux prédicats corrects<sup>6</sup>, et d'autres règles qui n'ont pas de

---

6. Ainsi, si le système propositionnel contient la règle :

P ou Q ;  
Si P, alors R  
Si Q, alors R

-----

R,

le système prédicatif contiendra une version qui ancre l'inférence à la structure prédictive. Dans ce cas, l'inférence prédictive correspondante est :

S1 [Tous les X] OU S2 [PRO-Tous les X]  
S3 [Tous les X : S1 [PRO]]  
S3 [Tous les X : S2 [PRO]]

-----

correspondant propositionnel. Parmi ces dernières, Braine inclut une forme de généralisation universelle (« Si un X quelconque est S et s'il y a des X qui sont S, alors tous les X sont S ») et de généralisation existentielle (« Si *a* est S, et si *a* est un des X, alors quelques X sont S »), ainsi qu'une série assez riche de règles de conversion entre formes logiques différentes des mêmes énoncés.

Mais la différence peut-être la plus importante entre le système propositionnel et le système prédicatif est la notion de thématization, avec les conséquences qu'elle apporte pour l'application des règles. Comme pour le cas propositionnel, la théorie de la logique mentale prédicative suppose que les sujets sont doués d'une routine d'application élémentaire et essentiellement automatique des règles, appelée « routine de raisonnement direct », qui explique les capacités primaires de raisonnement logique, tandis que d'autres stratégies non nécessairement disponibles à tout moment et pour tout le monde, constituant ce que l'on appelle les « capacités secondaires de raisonnement », expliquent les différences individuelles. Cependant, la différence entre logique propositionnelle et prédicative amène à changer ce que l'on entend par raisonnement primaire et secondaire. Il s'agit d'adapter le programme de raisonnement à la plus grande liberté permise par la présence des structures prédicatives et relationnelles. Dans un raisonnement prédicatif, plusieurs objets de discours peuvent apparaître, et se focaliser sur l'un ou sur l'autre peut amener les sujets à développer une ligne de raisonnement plutôt qu'une autre : si l'on parle de Jean et de Marie, nos pensées pourraient suivre des chemins différents (et donc aboutir à des conclusions différentes) selon que l'on se concentre sur l'information que le discours met à notre disposition sur Jean ou bien sur Marie. L'idée de la thématization est de permettre le développement de ces différentes stratégies de recherche de démonstration dépendantes des acteurs introduits par les prémisses d'un raisonnement prédicatif. Or, la capacité de trouver le point de thématization qui permet de développer une stratégie démonstrative efficace est partiellement dictée par le problème même, et partiellement dépendante des habilités individuelles.

Ainsi, comme Braine avait distingué les raisonnements primaire et secondaire pour le cas propositionnel, dans l'extension prédicative de la théorie il distingue la thématization primaire, censée être universellement disponible et en large mesure

---

### S3 [Tous les X]

que l'on pourrait lire comme « Si 1) tous les X sont S1 ou sont S2, et 2) si tous les X qui sont S1 sont S3, et 3) si tous les X qui sont S2 sont aussi S3, alors on peut conclure que tous les X sont S3 ». Par exemple, si toutes les billes sont rouges ou vertes, et si toutes les billes qui sont rouges sont transparentes, et si aussi toutes les billes qui sont vertes sont transparentes, alors toutes les billes sont transparentes. La raison pour laquelle une inférence conditionnelle prend cette forme dans la logique prédicative de Braine apparaît claire si l'on tient compte de la relation entre énoncés relatifs quantifiés et forme conditionnelle.

automatique, de la thématization secondaire. Quand un problème présente une conclusion donnée, avec un sujet donné, le point thématique du discours sera le sujet de la conclusion, ou, en l'absence d'une conclusion donnée, le sujet de sa première prémisses. Ceci est la stratégie de la thématization primaire qui, étant donné qu'un objet de discours est nécessaire, ne demande aucun effort pour se déployer. C'est pourquoi ce type de thématization est censé faire partie des habilités de raisonnement primaire. Toutefois, concentrer l'attention sur un point thématique du discours (ou dans ce contexte, essayer de lui appliquer les règles du système) peut aussi conduire sur de fausses pistes. Changer de point thématique peut permettre de « voir » comment utiliser la règle qui nous permet de faire le bon pas d'inférence, mais ce changement n'a rien d'automatique.

Choisir un point thématique secondaire est donc question de différences individuelles, plutôt que de stratégies universellement disponibles qui sont l'objet d'étude de la théorie de la logique mentale ; cependant, l'analyse d'une de ces stratégies peut nous permettre de comprendre comment la théorie peut expliquer les variations de performances dans les problèmes syllogistiques.

Braine propose l'idée que la difficulté à trouver des solutions à des syllogismes peut bien dépendre du fait que le point thématique primaire, déclenché par exemple par le sujet de la conclusion, bloque la possibilité d'appliquer des règles disponibles dans le système logique, mais non applicables à ce point thématique. C'est le cas, par exemple, du problème (iii) :

Tous les boulangers sont avocats.  
Aucun boulanger n'est chauffeur.

Réfléchissons sur ce que l'on peut dire des boulangers : rien ne semble venir à l'esprit. La stratégie de thématization primaire impose comme point thématique les boulangers, et en fait aucune règle du système prédicatif ne s'applique à ce point thématique. Il n'est pas surprenant, donc, de trouver un taux de solutions correctes qui ne dépasse pas les 10 %.

Une stratégie de thématization plus efficace est de déplacer le point thématique du sujet de la prémisses (ou si elle est donnée, de la conclusion) vers le sous-ensemble des individus déterminé par le sujet de la conclusion auquel le terme moyen s'applique. Cette stratégie de thématization secondaire demande de relativiser un terme quantifié, c'est-à-dire qu'elle amène à choisir un point thématique avec la forme logique [... les A ... : B [PRO]] (où les points de suspension indiquent un quantificateur). Avec un point thématique de cette forme, il sera plus naturel de chercher à appliquer les règles qui valent pour un quantificateur relativisé, règles qui ne pouvaient pas s'appliquer au point thématique primaire, qui n'était pas une forme PRO. Dans le cas de notre exemple, nous ne savons pas dire grand chose des boulangers, mais que dirons-nous des avocats qui sont boulangers

(c'est-à-dire, du point thématique « [Tous les avocats : boulangers [PRO]] »)? Immédiatement, nous pouvons voir que les Avocats qui sont boulangers ne peuvent pas être chauffeurs, car la deuxième prémisse nous dit qu'il n'y a pas de boulangers qui sont chauffeurs. Donc : les avocats qui sont boulangers ne sont pas chauffeurs, et donc, « Quelques avocats ne sont pas chauffeurs ». Voilà une conclusion valide : à partir d'un point thématique secondaire, elle apparaît presque naturelle, mais rien ne garantit que nous serons capables de voir ce point secondaire.

### 5.5.3.3. Preuves empiriques pour la théorie et ses limites

A la différence des théories que l'on a présentées jusqu'à maintenant, la théorie de Braine et O'Brien n'a pas spécialement cherché de confirmation par des analyses de performance sur les syllogismes. Bien qu'en principe elle puisse, comme on l'a vu, fournir des éléments intéressants et nouveaux d'analyse des syllogismes, son envergure a plutôt porté ses auteurs à explorer des raisonnements de type différent. Ainsi, Yang *et al.* (1998) ont choisi d'appliquer le modèle prédicatif à des raisonnements avec des relations unaires et binaires, de type non étudié auparavant, tous résolubles à partir de la routine prédicative du raisonnement direct et avec thématization primaire.

Sur la trace des études sur le raisonnement propositionnel (Braine, Connel & Romain 1998 ; Yang, Braine, O'Brien, 1998) ont réalisé plusieurs études avec différentes classes de problèmes, libres dans la mesure du possible d'implications conversationnelles et de significations particulières, concernant des relations entre certaines pièces d'un jeu. Les sujets avaient deux tâches : évaluer si une conclusion donnée était une conséquence des prémisses, et estimer la difficulté subjective de chaque problème. Voici un exemple de problème avec prédicats unaires :

Il n'y a pas de pièces rouges faites en bois  
 Toutes les pièces triangulaires sont faites en bois  
 Toutes les pièces sont soit rouges, soit vertes  
 -----  
 Quelques pièces triangulaires ne sont pas rouges ?

et voici un exemple de problème avec prédicats binaires :

Les filles ont toutes eu des pièces soit bleues soit rouges  
 Les filles qui ont eu les pièces rouges ont joué avec Marc  
 Les filles qui ont eu les pièces bleues ont joué avec Carl  
 -----  
 Quelques filles n'ont joué ni avec Marc ni avec Carl ?

Bien que les problèmes comprennent en général plusieurs prémisses et apparaissent plus compliqués que des syllogismes, dans les trois études présentées (Yang, Braine & O'Brien, 1998) ont trouvé un taux de réponses correctes qui

dépasse les 95 %. De même, la corrélation entre difficulté prédite par la théorie et difficulté estimée par les sujets varie entre 0,75 et 0,90.

Ces résultats, pour lesquels d'autres théories n'ont pas une explication évidente, font de la nouvelle théorie de Braine & O'Brien un des développements les plus intéressants de ces dernières années. Cependant, étant donné sa complexité, elle n'est pas encore assez développée tant dans sa structure, qu'en ce qui concerne sa base expérimentale, pour pouvoir l'évaluer avec précision. D'autres travaux sont nécessaires pour déterminer si elle sera à la hauteur des attentes.

## 5.6. Conclusions

Si l'on porte un regard d'ensemble sur les théories que nous avons examinées, nous ne pouvons pas nous empêcher d'éprouver un sentiment de déception. Nous avons introduit une série de problèmes bien définis, pour lesquels on serait en droit d'attendre des solutions claires de la part d'une discipline mature. Les solutions ne manquent pas, car presque chaque article publié en propose une. Toutefois, pas une seule ne peut se réclamer gagnante. Vues d'une certaine distance, toutes les théories semblent en mesure de pouvoir expliquer les phénomènes qu'on a examinés, mais un examen plus approfondi montre une autre réalité. La théorie acceptée par la plupart des psychologues reste la théorie des modèles mentaux de Johnson-Laird, mais il est difficile de voir comment elle peut résister aux objections que nous avons formulées. Toutefois, les alternatives aux modèles mentaux ne sont pas satisfaisantes. Certaines propositions, comme la tentative de réduire le raisonnement déductif à un phénomène statistique, semblent plus être des exercices de virtuosité que des positions tenables. D'autres, bien plus motivées et développées sur le plan formel et empirique, comme la théorie de Rips ou celle de Braine & O'Brien, manquent encore d'une solide fondation psychologique et ne couvrent pas encore tout le spectre de phénomènes auxquels il faut chercher une explication.

De façon un peu obsessionnelle, presque toutes les théories se sont surtout concentrées sur les phénomènes du syllogisme, alors que le raisonnement syllogistique n'est que la partie émergée de l'iceberg des capacités de raisonnement déductif prédicatif. Beaucoup de travail a été fait dans les vingt dernières années, mais pour explorer la totalité de l'iceberg, il reste encore beaucoup à faire, sans par ailleurs que l'on sache si l'on est sur la bonne voie. Vingt-trois siècles après Aristote, la découverte des bases psychologiques du raisonnement prédicatif oscille encore entre science et mystère.

L'auteur remercie Martine Gueguen pour son aide dans la rédaction française de ce texte.

## 5.7. Bibliographie

- Bara, B.G., Bucciarelli, M., Johnson-Laird, P.N. (1995). Development of syllogistic reasoning. *American Journal of Psychology*, 108, 157-193.
- Barrouillet, P. & Lecas, J.F. (1998). How can mental models theory account for content effects in conditional reasoning? A developmental perspective. *Cognition*, 67, 209-253.
- Bechtel, W. & Abrahamsen, A. (1991). *Connectionism and the mind*. Oxford : Basil Blackwell.
- Bonatti, L. (1994a) Propositional reasoning by model? *Psychological Review*, 101, 725-733.
- Bonatti, L. (1994b). Why should we abandon the mental logic hypothesis? *Cognition*, 50, 17-39.
- Bonatti, L. (1998). Possibilities and real possibilities for a theory of reasoning. In Z. Pylyshyn (Ed.) *Essays on representation* (pp. 85-119). Londres : Ablex.
- Bonomi, A. (1998). Indices and contexts of discourse. *Lingua e stile*, XXXIII, 471-485.
- Braine, M.D.S., Reiser, B.J., Rumin, B. (1998). Some empirical justifications for a theory of natural propositional logic. In Braine, M.D.S., O'Brien, D.P. (Eds.) *Mental Logic* (pp. 91-144). Mahwah, New Jersey : Lawrence Erlbaum.
- Braine, M.D.S. (1998). Steps toward a mental predicate logic. In Braine, M.D.S., O'Brien, D.P. (Eds.) *Mental Logic* (pp. 273-332). Mahwah, New Jersey : Lawrence Erlbaum.
- Bucciarelli, M. & Johnson-Laird, P.N. (1999). Strategies in syllogistic reasoning. *Cognitive Science*, 23, 247-303.
- Chater, N. & Oaksford, M. (1999). The probability heuristics model of syllogistic reasoning. *Cognitive Psychology*, 38, 191-258.
- Cherchia, G. & McConnell-Ginet, S. (2000). *Meaning and grammar*. Cambridge : MIT Press.
- Evans, J. St. B.T., Newstead, S.E., Byrne, R.M.J. (1983). *Human reasoning. The psychology of deduction*. Hove : Lawrence Erlbaum.
- Evans, J. St. B.T. (1989). *Bias in human reasoning*. Hove : Lawrence Erlbaum.
- Fodor, J.A. (1975). *The Language of Thought*. New York : Thomas Y. Crowell.
- Ford, M. (1995). Two modes of mental representation and problem solution in syllogistic reasoning. *Cognition*, 54, 1-71.
- Garnham, A. (1987). *Mental models as representations of discourse and text*. Chichester : Ellis Horwood.
- Ioup, G. (1975). Some universals of quantifier scope. In Kimball, J. (Ed.) *Syntax and Semantics* (Vol. 4). Bloomington : Indiana University Press.
- Johnson-Laird, P.N. (1983). *Mental models*. Cambridge, MA : Harvard University Press.
- Johnson-Laird, P.N. & Bara, B. (1984). Syllogistic inference. *Cognition*, 16, 1-61.
- Johnson-Laird, P.N., Byrne, R.M.J. (1991). *Deduction*. Hove : Lawrence Erlbaum.

- Johnson-Laird, P.N. & Savary, F. (1999). Illusory inferences: A novel class of erroneous deductions. *Cognition*, 71, 191-229.
- Johnson-Laird, P.N., Legrenzi, P., Girotto, V., Legrenzi, M.S. (2000). Illusory inferences. *Science*, 288 (5465), 531-532.
- Johnson-Laird, P.N., Savary, F., Bucciarelli, M. (2000). Strategies and tactics in reasoning. In Schaeken W., De Vooght, G., D'Ydewalle, G. (Eds.) *Deductive reasoning and strategies* (pp. 209-240). Mahwah : L. Erlbaum.
- Marcus, S.L. (1982). Recall of logical argument lines. *Journal of Verbal Learning and Verbal Behavior*, 21, 549-562.
- Newstead, S.E., Handley, S.J., Buck, E. (1999). Falsifying mental models: Testing the predictions of theories of syllogistic reasoning. *Memory and Cognition*, 27, 344-354.
- Politzer, G. (1990). Immediate deduction between quantified sentences. In Gilhooly, K.J., Keane, M.T., Logie, R.H., Erdos, G. (Eds.) *Lines of Thinking* (Vol. 1., pp. 85-97). Chichester : Wiley.
- Politzer, G. (1990). Non-logical solving of categorical syllogisms. In M.G. Caverni, J.-P., Fabre, J.-M., Gonzales (Eds.) *Cognitive biases*(pp. 193-204). Amsterdam : North-Holland.
- Politzer, G. (1995). Les raisonnements orientés vers la preuve. In R. Ghiglione et J.-F. Richard (Eds.) *Cours de Psychologie* (Vol. 6, pp. 315-341). Paris : Dunod.
- Polk, T.A. & Newell, A. (1995). Deduction as verbal reasoning. *Psychological Review*, 102, 533-566.
- Rips, L.J. (1983). Cognitive processes in propositional reasoning. *Psychological Review*, 90, 38-71.
- Rips, L.J. (1994). *The psychology of proof: Deductive reasoning in human thinking*. Cambridge, MA : MIT Press.
- Roberts, M.J., Newstead, S.E., Griggs, R.A. (2001). Quantifier interpretation and syllogistic reasoning. *Thinking and Reasoning*, 7, 173-204.
- Sperber, D. & Wilson, D. (1995). *Relevance. Communication and cognition*. Oxford : Basil Blackwell, 2<sup>e</sup> édition.
- Stanovich, K.E. & West, R.F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General*, 127, 161-188.
- Yang, Y., Braine, M.D.S., O'Brien, D.P. (1998). Some empirical justification of the mental-predicate-logic model. In Braine, M.D.S., O'Brien, D.P. (Eds.) *Mental Logic* (pp. 333-365). Mahwah : Lawrence Erlbaum.