

PSYCHOLOGIE ET SCIENCES DE LA PENSÉE

Collection dirigée par

Olivier Houdé

Comité d'évaluation scientifique

Stanislas Dehaene, Pierre Jacob, Denis Miéville,
Jacques Pitrat, Joëlle Proust et François Rastier

0110MAR 97 0924

1

POLYMORPHISME DU RAISONNEMENT HUMAIN

Une approche de la flexibilité
de l'activité inférentielle

Christian GEORGE



Presses Universitaires de France



DL 10 MAR. 97 09240

Je dédie cet ouvrage à toutes les personnes, principalement des étudiants, qui ont accepté de participer aux expériences que j'ai réalisées sur le raisonnement.

Je remercie tous mes collègues du laboratoire de Psychologie cognitive de l'Université de Paris VIII qui, par leurs propres travaux, leurs commentaires et leurs critiques, ont contribué à l'élaboration des réflexions présentées ici, en particulier Jean-François Richard, Guy Politzer et Rui Manuel Da Silva Neves.

ISBN 2 13 048046 2

ISSN 1248-8976

Dépôt légal — 1^{re} édition : 1997, février

© Presses Universitaires de France, 1997
108, boulevard Saint-Germain, 75006 Paris

Sommaire

Introduction, 1

PREMIÈRE PARTIE LE RAISONNEMENT AVEC DES ARGUMENTS DÉDUCTIFS

Chapitre Premier – Le raisonnement propositionnel, 21

- 1.1. Syntaxe et sémantique de la logique des propositions, 22
 - 1.1.1. *L'approche syntaxique (l'approche de la théorie de la preuve)*, 22
 - 1.1.2. *L'approche sémantique (l'approche de la théorie des modèles)*, 23
 - 1.1.3. *Relations entre les approches sémantique et syntaxique*, 26
- 1.2. L'interprétation des connecteurs propositionnels dans le langage naturel, 26
- 1.3. Les syllogismes propositionnels, 29
- 1.4. La production de la conclusion et les contraintes pragmatiques, 32
- 1.5. Principales réactions à deux questions, 34
- 1.6. La théorie des règles formelles d'inférence, 36
- 1.7. La théorie des modèles mentaux, 40

Chapitre 2 – Le raisonnement conditionnel, 47

- 2.1. Assertabilité et vérifonctionnalité des énoncés conditionnels, 47
- 2.2. Les interprétations du conditionnel, 51
- 2.3. Les syllogismes conditionnels, 54
- 2.4. Règles logiques ou règles pragmatiques ?, 56
- 2.5. A propos de la tâche de sélection de Wason, 58

VI *Polymorphisme du raisonnement humain*

- 2.6. Règles formelles et modèles mentaux, 60
- 2.7. La théorie des schémas pragmatiques, 61
- 2.8. Conditionnels et instructions procédurales, 63

Chapitre 3 – Le raisonnement avec des quantificateurs et des variables, 65

- 3.1. Les quantificateurs dans la logique d'Aristote, 66
- 3.2. Prédicats et variables dans la logique moderne, 67
- 3.3. Les inférences immédiates avec des énoncés quantifiés, 69
- 3.4. Les syllogismes catégoriques, 72
- 3.5. Deux biais : atmosphère et croyance, 75
- 3.6. Reasonner avec des variables et des constantes, 77
- 3.7. Théorie des modèles mentaux et théorie des règles formelles, 80

DEUXIÈME PARTIE

LE RAISONNEMENT

AVEC DES ARGUMENTS NON DÉDUCTIFS

Chapitre 4 – L'induction classique, 87

- 4.1. Variétés de raisonnement et variétés d'induction, 88
- 4.2. Degré de confirmation et force d'un argument, 89
- 4.3. Facteurs sémantiques dans l'évaluation de la force des arguments inductifs, 92
- 4.4. Facteurs syntaxiques et variations concomitantes selon J. S. Mill, 95
- 4.5. Les difficultés de l'identification d'une régularité, 99
- 4.6. Les biais dans la formulation d'une conclusion inductive, 101
- 4.7. L'inférence causale, 102
- 4.8. Modèles et théories de l'induction, 109

Chapitre 5 – L'abduction et l'explication, 113

- 5.1. L'abduction selon Peirce, 114
- 5.2. Les processus en jeu dans l'abduction, 115
- 5.3. Le paradigme de l'attribution causale de Kelley, 118
- 5.4. Diagnosticité et typicalité, 120

- 5.5. Le diagnostic médical, 121
- 5.6. Les inférences des jurés, 123
- 5.7. Un modèle symbolique du raisonnement abductif dans un « micro-monde », 125
- 5.8. Un modèle connexionniste de la cohérence explicative, 126

Chapitre 6 – Le raisonnement avec des prémisses incertaines, 129

- 6.1. L'expérience princeps de Byrne, 130
- 6.2. L'origine de l'incertitude associée à un énoncé, 132
- 6.3. L'effet de l'incertitude d'une prémisse sur la conclusion d'un argument déductif, 134
- 6.4. L'effet du type d'argument, 135
- 6.5. L'effet de la force attribuée à un argument, 136
- 6.6. L'effet combiné de deux prémisses incertaines, 138
- 6.7. Autres moyens d'évaluer l'incertitude d'une proposition, 141
- 6.8. Aspects théoriques, 143
- 6.9. Raisonnements fondés sur des croyances ou sur des suppositions, 144

TROISIÈME PARTIE

LE MAINTIEN DE LA COHÉRENCE

Chapitre 7 – Le raisonnement non monotone, 151

- 7.1. Les situations avec intervention d'un raisonnement non monotone, 152
- 7.2. Précautions pour minimiser le risque d'infirmité des inférences par défaut, 154
- 7.3. Le choix entre plusieurs sortes de révision, 155
- 7.4. Oscillations de la conclusion et hiérarchie de défauts, 157
- 7.5. La pluralité des extensions, 158
- 7.6. Quelle logique pour le raisonnement non monotone ?, 159
- 7.7. Le conditionnel comme règle par défaut, 162

Chapitre 8 – La macrostructure du raisonnement, 165

- 8.1. La transitivité des inférences, 166
- 8.2. Les stratégies dans les raisonnements déductifs longs, 167

VIII *Polymorphisme du raisonnement humain*

- 8.3. Le raisonnement par héritage des défauts, 169
- 8.4. La transmission et le cumul de l'incertitude, 170
- 8.5. L'utilisation combinée de la déduction et de l'induction dans un raisonnement, 173
- 8.6. L'induction comme démarche alternative ou parasite dans l'identification d'une régularité, 176
- 8.7. L'architecture du raisonnement, 178

Chapitre 9 – La rationalité humaine, 183

- 9.1. Les formes de la rationalité, 183
- 9.2. Les principales thèses sur la rationalité du raisonnement, 184
- 9.3. Rationalité limitée et compétence limitée, 188
- 9.4. La rationalité instrumentale, 189
- 9.5. La rationalité des croyances, 192

Conclusion, 197

Bibliographie, 213

Introduction

« Faites-moi plaisir et allez-vous-en.
« Je me demandai quelle était la valeur de ce "et", s'il reliait les termes d'une équivalence ou d'une chronologie, choix qui impliquait de ma part deux attitudes nettement différentes. Suzanne s'approcha et m'ôta mes habits et mes doutes. »

MICHEL RIO, *Alizés*.

Le raisonnement est présent dans de nombreuses activités mentales, y compris dans des activités où sa présence n'est pas manifeste. Il intervient bien sûr dans l'argumentation, la résolution de problème, la planification, les décisions. Il intervient également dans la compréhension du discours oral ou écrit, dans la catégorisation, le jugement. Quiconque lit le récit de la prise de Troie comprend immédiatement que si Ulysse et ses compagnons se cachent à l'intérieur du fameux cheval, et que celui-ci est tracté à l'intérieur des remparts, alors Ulysse et sa troupe seront à l'intérieur de la ville. De même quiconque apprend que Zeus et Léo ont eu pour enfants Apollon et Artémis sait qu'Apollon est le frère d'Artémis. Non seulement l'individu fait constamment de telles inférences pour son propre compte, mais il en prête de semblables à ses congénères afin de donner un sens à leur conduite et d'interagir avec eux. De fait, il est impossible à un être humain ou à un système informatique de communiquer avec autrui, et à plus forte raison d'enseigner quelque chose ou d'anticiper les réactions de quelqu'un, s'il ignore ce que le destinataire peut inférer à partir du message qui lui est adressé. Relevons enfin que le raisonnement intervient de façon insidieuse dans la révision des croyances. Il y a donc de nombreuses raisons de s'intéresser aux formes que peuvent revêtir les raisonnements humains, aux facteurs qui les affectent, et aux mécanismes sous-jacents qui les produisent.

L'étude du raisonnement humain est d'abord l'affaire de la psychologie même si d'autres disciplines lui ont apporté des contributions fondamentales. Il est intéressant de noter que les études sur le raisonnement effectif de l'adulte se comptent sur les doigts d'une main

(ou de deux, pour tenir compte de l'existence de travaux oubliés) pendant la première moitié de ce siècle, alors qu'elles se multiplient après. Il y a une explication fort simple à cette situation. Pendant longtemps on a pensé comme Descartes que l'homme était un être doué de raison, en tout cas l'homme adulte à condition d'échapper à la folie, à l'ivresse et aux passions. Cette propriété a longtemps rapproché l'homme des dieux, et l'a distingué de l'animal : dans la culture occidentale l'homme et les dieux communiquent entre eux mais non avec l'animal. Non seulement la logique fournissait les critères normatifs du bien raisonner, mais l'homme adulte était censé se conformer de façon plutôt satisfaisante à ses prescriptions. Une des questions dont débattaient les savants était de savoir si c'est la pensée qui est le miroir de la logique, ou l'inverse (cf. Piaget, 1947). Dans ces conditions l'étude du raisonnement adulte ne semblait pas présenter un grand intérêt puisqu'on supposait qu'elle révélerait les principes déjà décrits par les logiciens. Il était certes connu que l'homme adulte commettait occasionnellement des erreurs de raisonnement, mais on les considérait comme des accidents de parcours qui ne justifiaient pas une remise en cause de sa compétence logique. Par contre on pouvait s'attendre à découvrir des phénomènes beaucoup plus curieux en étudiant le raisonnement de l'enfant avant qu'il n'accède à la pensée logique de l'adulte, et les chercheurs se sont tournés en plus grand nombre vers cette étude. L'œuvre de Piaget témoigne de cet intérêt. Tout a changé dans les années soixante lorsque quelques psychologues, principalement britanniques, ont mis en évidence l'existence chez l'adulte de déviations trop systématiques par rapport aux prescriptions de la logique pour qu'on puisse les ignorer (voir l'ouvrage de référence d'Evans, Newstead et Byrne, 1993). Depuis lors la logicité de la pensée adulte ne constitue plus une certitude première. L'idée d'une correspondance profonde entre raisonnement humain et logique sera encore plus profondément ébranlée ultérieurement lorsque divers travaux suggéreront qu'il est possible de bien raisonner sans se référer à des règles logiques d'inférence (par ex. Johnson-Laird, 1983), voire sans aucune sorte de référence à la logique (par ex. l'approche connexionniste).

Pendant que ce type de changement survenait en psychologie, des changements non moins importants survenaient dans les disciplines connexes concernées par l'étude du raisonnement. La logique, classiquement définie comme science du raisonnement correct ou valide, a subi elle-même de profondes transformations depuis que sa formalisation au tournant du vingtième siècle en a fait un *calcul*, c'est-à-dire

une suite d'opérations sur des symboles. La logique moderne n'est plus la logique d'Aristote. En distinguant deux faces dans tout système logique, une face syntaxique et une face sémantique, auxquelles correspondent respectivement la relation de conséquence prouvable et celle de conséquence (tauto)logique, on pouvait évaluer celui-ci par rapport à des propriétés comme la consistance, la complétude, la fiabilité (*soundness*) (par ex. Church, 1956 ; Kleene, 1967 ; voir le § 1.1). En complétant ou en modifiant les axiomes des systèmes classiques on a engendré de nouveaux systèmes. Cela a entraîné l'apparition de logiques « étendues » ou « déviantes » (les logiques étendues complètent la logique classique, comme par ex. la logique modale, tandis que les logiques « déviantes » reposent sur des principes rivaux, par ex. le refus de la bivalence entre le *Vrai* et le *Faux*). Une autre forme d'extension de la logique classique concerne les inférences inductives, en particulier l'attribution d'un *degré de confirmation* à une hypothèse lorsque des faits connus sont favorables à celle-ci sans permettre d'affirmer avec certitude qu'elle est vraie. Ainsi à côté d'une logique de la *démonstration* se constitue une logique de la *confirmation* (Carnap et Jeffrey, 1971 ; Hempel, 1943). Plus récemment les logiques modales ont connu un grand développement grâce à la possibilité de les interpréter avec la sémantique des « mondes possibles ». En bref, *la logique* entendue comme discipline a élaboré *des logiques* entendues comme systèmes d'inférence. Dans ces conditions parler de raisonnement logique ou de pensée logique nécessiterait en droit qu'on spécifie à quel système logique on se réfère. En l'absence de spécification il s'agit généralement de la logique classique (logique du calcul des prédicats du premier ordre ou plus modestement logique propositionnelle), mais il est clair que celle-ci n'a plus le monopole du label « logique ».

Au cours de ces dernières décennies d'autres disciplines se sont également intéressées au raisonnement. En linguistique de nombreuses analyses ont été consacrées à l'interprétation et à l'utilisation des connecteurs *si... alors, ou, et...* Une contribution importante a concerné l'identification des fonctions des énoncés et des principes régissant la communication. Divers auteurs se sont intéressés à ce qui est compris sans avoir été explicitement dit. On a analysé par exemples les différences entre les présupposés et les sous-entendus (Ducrot, 1972). Surtout, on a mis en évidence la présence d'une multiplicité d'inférences faites par l'auditeur lorsque celui-ci suppose que le locuteur respecte le « principe de coopération » et les « maximes » qui en découlent (Grice, 1975). Dans les situations où l'étude du raisonne-

ment repose sur une transmission verbale d'informations entre l'expérimentateur et le sujet, il convient de prendre en compte ces facteurs de pragmatique linguistique puisqu'ils affectent la façon dont les énoncés sont compris par l'interlocuteur (Politzer, 1986, 1993).

La philosophie s'est également intéressée aux rapports entre raisonnement et logique à propos de questions concernant le sens d'un énoncé ou l'épistémologie des sciences. De grands débats ont eu lieu principalement dans les pays anglo-saxons dans le cadre ou le prolongement des courants dits « empirisme logique » et « philosophie de l'esprit » (pour une introduction en français voir Engel, 1994 ; Jacob, 1980 ; Malherbe, 1981). Certains d'entre eux concernent des questions débattues par des psychologues. Ils ont rencontré néanmoins un faible écho dans la psychologie anglo-saxonne, et un écho plus faible encore dans la psychologie de langue française. Certes ces écrits concernent souvent des spéculations qui se prêtent difficilement à une mise à l'épreuve expérimentale, mais certaines d'entre elles peuvent s'avérer très fécondes pour formuler de nouvelles hypothèses (par ex. Cherniak, 1986 ; Harman, 1986 ; Hintikka, 1962 ; Margolis, 1982).

Ces dernières décennies ont été également marquées par le développement spectaculaire de l'intelligence artificielle. L'IA élabore des programmes dits systèmes experts capables de réaliser diverses tâches. Les premiers systèmes construits furent destinés à prouver des théorèmes divers. Les systèmes ultérieurs sont destinés à se substituer à des experts dans des tâches pratiques comme le diagnostic médical ou la prospection géologique. Pour cela ils doivent tenir compte des contraintes inhérentes aux situations réelles, et en particulier de l'obligation d'émettre un avis même si les informations disponibles sont incertaines ou incomplètes. Pour émettre un avis minimisant les risques d'erreurs il a fallu non seulement recourir aux logiques « étendues » déjà existantes, mais aussi en élaborer de nouvelles qui soient appropriées aux contraintes et aux objectifs (voir par ex. Haton *et al.*, 1991 ; Lea Sombé, 1989). L'apport le plus connu de l'IA sur ce point concerne le raisonnement non monotone. L'élaboration des systèmes experts constitue une démonstration de la puissance et des limitations des diverses logiques utilisées, limitations qui ont suscité quelques prises de position contre le « logicisme » initial de cette discipline (Kayser, 1993).

Ainsi les questions des psychologues sur le raisonnement ont tout d'abord pris la forme d'une interrogation sur le rôle effectif de la logique dans celui-ci, interrogation qui rejoint celle qui est apparue dans d'autres disciplines. Cette remise en cause est plus ou moins

radicale et a bien sûr une incidence sur les théories puisque celles-ci se définissent d'abord par les mécanismes inférentiels invoqués. Elle se manifeste à un premier degré par le succès d'expressions comme « logique naturelle ». Parler de logique *naturelle*, c'est reconnaître que si logique de la pensée il y a, celle-ci ne coïncide pas exactement avec celle qui est formalisée par les logiciens. A défaut d'identifier dans la pensée des éléments comparables aux axiomes des systèmes logiques, on lui a attribué des règles d'inférence jouant un rôle semblable pour dériver des conclusions. Or l'analyse des inférences effectuées par les profanes montre que celles-ci ne coïncident que partiellement avec les inférences du logicien. En particulier une partie non négligeable des inférences valides en logique semblent étranges ou inacceptables à la presque totalité des individus. Ainsi on ne rencontre guère d'individus qui infèrent « il pleut et il pleut », ou « il pleut ou la mer est salée » à partir de la proposition « il pleut », bien que ces deux inférences soient valides en logique classique. Deux façons principales de définir une logique naturelle ont été envisagées, l'une qui consiste à ne prendre qu'un sous-ensemble des règles d'inférence de la logique classique, l'autre qui consiste à prendre un sous-ensemble de celui-ci complété par des règles distinctes spécifiques à un domaine. La première approche a été adoptée par les partisans de la théorie des règles formelles d'inférence ou théorie de la « logique mentale » (par ex. Braine, voir le chap. 1), la seconde par des auteurs proches du courant néopiagétien comme Grize pour l'argumentation (cf. Borel, Grize et Miéville, 1983).

La problématique de l'étude psychologique du raisonnement humain a subi un premier grand changement sous l'effet de l'apparition de la nouvelle approche due à la théorie des modèles mentaux de Johnson-Laird (1983; Johnson-Laird et Byrne, 1991). Celle-ci considère qu'aucune règle logique d'inférence ne fait partie de l'équipement de la pensée humaine, mais que celui-ci peut néanmoins parvenir à des conclusions valides en se conformant au principe logique qu'une conclusion est valide si elle vraie dans toutes les interprétations possibles des prémisses. Les individus raisonneraient conformément à ce principe en élaborant plusieurs représentations mentales, appelées « modèles mentaux », des différents cas de figure correspondant aux prémisses (voir le chap. 1). Par ailleurs cette théorie admet que la signification attribuée aux connecteurs propositionnels « et », « ou », « si... alors... », est conforme aux tables de vérité des logiciens. Elle fait bien l'économie d'une logique naturelle, mais elle n'est pas antilogiciste.

Une nouvelle modification de la problématique est en train de survenir à cause du développement de l'approche connexionniste. Celle-ci constitue actuellement un sérieux challenger pour les théories « établies ». Cette approche modélise le raisonnement non par une succession de transformations opérées sur des propositions ou des représentations selon des règles définies comme dans les modèles symboliques, mais comme une diffusion de l'activation produite par un input à l'intérieur d'un réseau d'unités structurées en différentes « couches » entre l'entrée et la sortie. C'est la force des liaisons associatives entre les unités qui détermine comment se répartit l'activation dans le réseau et de ce fait la forme de la conclusion, et non pas la structure logique de la relation entre les prémisses (Rumelhart, 1989). Lorsque les réponses sont conformes à la logique, leur logicité n'est qu'une propriété émergente de ce processus, non une propriété du réseau. Si l'approche connexionniste ne permet pas (encore ?) de modéliser de façon satisfaisante les résultats obtenus avec des arguments déductifs, elle conduit à des résultats très prometteurs avec des arguments inductifs (voir les chap. 4 et 5). Une conséquence de l'apparition du connexionnisme est que les deux théories des règles formelles et des modèles mentaux n'apparaissent plus comme les pôles extrêmes d'une échelle, mais comme deux points relativement proches puisque relevant toutes deux du paradigme de l'approche symbolique classique, l'autre pôle étant constitué par le nouveau paradigme connexionniste. Une des premières manifestations de cette prise de conscience d'une possible restructuration du domaine est constituée par l'opposition établie par Rips (1990) entre deux sortes de théories dans la dernière revue de question d'ensemble consacrée au raisonnement par l'*Annual Review of Psychology*. Il oppose une approche « stricte » et une approche « lâche », la première considérant que le processus inférentiel est constitué par une succession d'étapes discrètes et repose sur l'utilisation d'indices spéciaux tels que les connecteurs « ou », « si...alors », contrairement à la seconde.

On peut considérer que l'élaboration d'une psychologie du raisonnement comporte deux tâches principales, l'une consistant à établir les contours de la compétence à raisonner des humains, l'autre consistant à expliquer par quel(s) mécanisme(s) cette compétence est assurée. La première tâche consiste en quelque sorte à établir la carte géographique des inférences acceptées ou rejetées par les individus. On pourrait penser que cet objectif est déjà atteint ou quasiment atteint puisque de très nombreuses expériences ont déjà été réalisées. Il n'en est rien car la presque totalité des travaux ont porté sur des

arguments déductifs. On peut noter d'ailleurs que ceux-ci ont porté presque exclusivement sur la microstructure du raisonnement, en entendant par là les inférences élémentaires qui en constituent le grain, alors que les études sur le raisonnement inductif ont porté presque exclusivement sur sa macrostructure et notamment sur les stratégies mises en œuvre. Or le raisonnement ordinaire fait appel à de très nombreuses inférences non déductives dont on sait peu de choses. Les développements récents survenus en IA et en logique invitent à les prendre en considération. Il est donc nécessaire et opportun d'élargir le champ de l'étude psychologique à des formes de raisonnement largement ignorées jusqu'à présent.

La première tâche doit s'appuyer sur la classification des types d'arguments établie par la logique. Rappelons que selon la définition technique usuelle, un *argument* est constitué par un couple prémisses-conclusion, qui n'est qu'un couple entrée-sortie (*NB.* — Dans les études expérimentales un argument est parfois défini seulement par les prémisses présentées puisqu'un sujet peut être invité à produire une conclusion à partir de celles-ci). Un argument indique seulement quelle conclusion est inférée de certaines prémisses, non comment l'inférence est effectuée. Ce résultat peut sembler limité, mais l'identification de différentes sortes de raisonnement à partir de critères externes présente non seulement un intérêt pratique évident, mais aussi un intérêt théorique, puisque cette carte impose des contraintes fortes sur la forme que peut prendre une théorie du raisonnement.

Cette première tâche comporte beaucoup plus de difficultés qu'il n'y paraît au premier abord. L'argument manifeste, celui qui est présenté par l'expérimentateur ou formulé par le participant d'une expérience, n'est pas toujours l'exact reflet de l'inférence qui est réalisée. Il faut se garder en pareil cas de projeter sur le profane les catégories du logicien, et de croire qu'un argument formellement déductif ou inductif corresponde à un raisonnement de même nature. Il y a deux raisons qui doivent inciter à cette prudence. La première est que les sujets d'une expérience n'accordent pas forcément le même statut à un couple prémisses-conclusion que l'expérimentateur ou le logicien. Une illustration simple de celle-ci concerne l'argument *a)* ci-dessous dénommé « affirmation du conséquent » (les prémisses sont au-dessus du trait horizontal, la conclusion au-dessous) :

a) Si Dionysos a bu alors il est jovial
Dionysos est jovial

Dionysos a bu

Cet argument doit être rejeté dans le cadre de la déduction, alors qu'il peut être accepté dans le cadre de l'induction où on s'intéresse à la conclusion la plus plausible. Certains « biais » présumés pourraient provenir d'un défaut de communication, l'expérimentateur ou le sujet se méprenant sur le statut accordé à un argument par l'autre (Margolis, 1987). La seconde raison, plus fâcheuse, est que les prémisses d'un argument manifeste peuvent ne pas correspondre aux prémisses du raisonnement effectif. Une illustration en est constituée par la transformation de l'argument *a*) en un autre argument, par exemple *b*) :

b) Si Dionysos a bu alors il est jovial
 (Si Dionysos est jovial alors il a bu)
 Dionysos est jovial

 Dionysos a bu

Ces décalages entre l'argument « objectif » ou « nominal » et l'argument « effectif » sont intéressants parce qu'ils permettent d'identifier des « modes opératoires » distincts. Le mode opératoire détermine les composantes de l'information qui sont prises en compte, ainsi que le mode d'expression de la réponse qui est adopté. Une question à examiner est de savoir si ces modes opératoires correspondent à des catégories établies par le logicien, par exemple les catégories de la déduction et de l'induction, ou encore les diverses formes de l'induction. Cela n'est pas assuré car certains d'entre eux résultent du non-respect des règles de la déduction. Le mode opératoire entraîne des différences interindividuelles nombreuses face à un même argument. Il peut entraîner aussi des différences intra-individuelles, un même individu pouvant adopter à l'égard d'un même type d'argument un mode d'approche variable selon le contenu et le contexte linguistique ou situationnel. Évidemment certains éléments de la situation peuvent favoriser l'homogénéité du groupe en sorte que tout le monde adopte un même parti pris, ou au contraire favoriser l'hétérogénéité du groupe et faire apparaître des partis pris différents. Toutes ces remarques invitent à la prudence pour parler de raisonnement déductif, inductif, ou autre. En effet le psychologue n'a pas toujours l'assurance qu'en produisant la conclusion d'un argument déductif ou inductif, le sujet observé ait procédé de façon déductive ou inductive. Ces appellations ne constituent souvent qu'une commodité de langage pour désigner le raisonnement avec des arguments déductifs, inductifs..., c'est-à-dire une métonymie.

Par ailleurs il ne faut pas rejeter dans quelque catégorie amorphe les arguments non déductivement valides. Ceux-ci peuvent être diffé-

renciés selon plusieurs critères. En particulier un argument déductivement invalide peut néanmoins être *inductivement fort*, c'est-à-dire que sa conclusion est jugée très plausible étant donné les prémisses quoique non certaine. Les arguments non déductifs sont jugés plus ou moins forts inductivement, et il est utile de savoir quels sont les facteurs qui interviennent dans ce jugement (Osherson *et al.*, 1990). La notion de degré de force d'un argument est à rapprocher de celle de *degré de confirmation* d'une hypothèse envisagée en épistémologie (Hempel, 1943) puis en IA, qui concerne la vraisemblance accordée à une conjecture par les autres propositions déjà établies. Les difficultés rencontrées par les chercheurs ayant tenté d'élaborer une logique de la confirmation ou une logique de l'induction n'enlèvent rien à l'intérêt de ces notions dans l'étude du raisonnement humain.

La diversité des formes d'inférence témoigne de la flexibilité du raisonnement et assure son adaptabilité face à la variabilité des facteurs externes et internes. Elle permet de dériver une conclusion même lorsque les informations utilisées comme prémisses n'ont pas la structure canonique des arguments déductifs de la logique classique, notamment lorsque les informations sont incomplètes ou incertaines. Elle permet également de moduler les inférences en fonction de la finalité de l'activité. On peut en effet poursuivre des objectifs fort différents lorsqu'on infère ce qui découle des propositions acceptées comme prémisses. Un objectif consiste à montrer que les prémisses entraînent nécessairement une conséquence particulière, comme dans la démonstration d'un théorème mathématique. En pareil cas l'intérêt ne porte pas sur le statut de vérité des prémisses ni sur celui de la conclusion. On entend montrer que celle-ci est une conséquence inéluctable de celles-là. Cet objectif est manifeste dans le raisonnement dit *contrefactuel* qui permet d'articuler deux propositions atomiques connues pour être fausses, comme : « Si Hitler avait eu la bombe atomique, il aurait gagné la guerre. » Un objectif tout différent consiste à déterminer ce qui est ou sera vrai dans une situation, ou à défaut ce qui est ou sera le plus plausible, compte tenu des informations disponibles sur celle-ci. Dans ce cas on recherche comme prémisses des informations vraies, ou à défaut on tient compte de leur degré de certitude afin de pouvoir moduler le degré de certitude de la conclusion. On est conduit à opérer de cette façon pour prendre les multiples décisions de la vie quotidienne ou professionnelle, car la qualité de la décision dépend de la vérité et de la complétude des informations sur la situation, non de la manière dont elles sont obtenues. Selon l'objectif poursuivi un raisonnement n'est pas évalué de la même façon.

Dans le premier cas l'évaluation concerne la rigueur du passage des prémisses à la conclusion, ce que les logiciens désignent par validité. Dans le second cas l'évaluation porte sur le statut de vérité de la conclusion, à tel point qu'on n'hésite pas à contester un raisonnement s'il aboutit à une conclusion non crédible.

Élargir l'étude à des formes de raisonnement qui ne relèvent pas de la déduction conduit à poser de nouvelles questions, et à formuler de nouvelles hypothèses. Une des questions concerne une possible filiation entre diverses formes de raisonnement, soit au cours du développement génétique, soit au cours du processus inférentiel. Les auteurs d'un manuel de référence écrivent que raisonner déductivement est d'une importance fondamentale pour l'intelligence humaine (Evans, Newstead et Byrne, 1993, p. 266). De façon plus radicale encore, Johnson-Laird (1993, p. 2) déclare : « Au cœur de la rationalité est la capacité à faire des déductions valides. » De nombreux autres auteurs partagent cette conviction. Cependant ils indiquent rarement en quoi la déduction serait fondamentale : est-ce par la fiabilité de ses conclusions, par sa fréquence, parce qu'elle servirait de base à d'autres formes de raisonnement, ou pour d'autres raisons ? La fiabilité des conclusions déductives est indéniable, mais il n'est pas assuré que les inférences déductives soient plus fréquentes ou plus basiques que d'autres. Une des idées défendues ici est que même dans l'hypothèse où le raisonnement avec des arguments déductifs serait basique, l'étude du raisonnement avec des arguments non déductifs pourrait être utile à la connaissance du premier. En effet rien ne garantit que les inférences avec des arguments déductifs, inductifs, abductifs... constituent des territoires indépendants. Il se pourrait par exemple que les inférences inductives soient assez fréquentes pour que certaines particularités de leur réalisation interviennent également dans les inférences déductives. Fodor (1983) considère que la confirmation, qui met en jeu des inférences inductives, est *isotrope*, c'est-à-dire qu'aucune information établie n'est *a priori* exclue lors de sa réalisation. Une telle caractéristique est de nature à expliquer pourquoi avec des arguments déductifs divers on a observé que les adultes avaient tendance à évaluer ou à enrichir les prémisses présentées par l'expérimentateur à partir de leurs connaissances et croyances personnelles. Un autre thème d'étude récent concernant l'effet de l'incertitude attribuée à un événement est de nature à expliquer la variabilité des raisonnements avec des énoncés comportant l'implication « si... alors... » en fonction de leur contenu ou contexte. En bref nous considérons que le raisonnement humain n'est pas cloisonné en secteurs, et

que ce que l'on apprend dans l'un pourrait éclairer ce qui se passe dans l'autre.

Après avoir collecté de nombreuses données empiriques relatives aux manifestations de la polymorphie, il convient d'en rendre compte. L'explication consiste à invoquer un ou plusieurs mécanismes inférentiels, en entendant par ce terme une suite spécialisée d'opérations élémentaires qui substitue aux propositions en entrée une proposition distincte en sortie. Les mécanismes inférentiels peuvent être généralement définis par un algorithme. Une véritable théorie du raisonnement doit non seulement indiquer le ou les mécanismes inférentiels invoqués, mais aussi spécifier dans quelles conditions chacun d'eux est déclenché, et comment son exécution est contrôlée. Par ailleurs une théorie peut éventuellement être complétée par un modèle qui en réalise l'implémentation dans un domaine particulier, généralement celui de la simulation informatique. Ce sont les mécanismes inférentiels invoqués qui caractérisent les théories du raisonnement. Or il est très difficile d'obtenir directement des informations sur ces mécanismes grâce à l'observation ou l'introspection, ce qui contribue à perpétuer les polémiques.

Si les principales théories se différencient par le mécanisme inférentiel invoqué, elles ont cependant un point commun qui est leur *monisme*, c'est-à-dire la prévalence accordée à un mécanisme sur les autres à l'intérieur du domaine de validité qu'elles se sont assigné. Ce monisme peut être relatif si on accorde à d'autres mécanismes un rôle subsidiaire, comme dans la théorie des règles formelles de Braine, ou absolu si on exclut tout autre mécanisme, comme dans la théorie des modèles mentaux de Johnson-Laird. La légitimité du monisme a rarement été discutée, en partie probablement parce que les théories du raisonnement ont été conçues au moins au départ pour une seule catégorie de raisonnement, le raisonnement déductif, voire une sous-catégorie de celui-ci. Or le monisme peut sembler d'autant moins plausible que le champ d'étude considéré est large. Il convient ainsi de se demander si des raisonnements aussi dissemblables par exemple que la déduction et l'induction peuvent relever d'un même mécanisme inférentiel. On ne saurait évidemment répondre à cette question en se référant au statut dissemblable de ces deux formes de raisonnement en logique, puisque précisément la part de celle-ci dans les inférences humaines fait actuellement l'objet d'un débat en psychologie comme indiqué précédemment.

Aujourd'hui le principal argument en faveur du monisme est métathéorique. Il réside dans le principe de parcimonie, qui consiste à

privilégier l'explication la plus simple, toutes choses égales par ailleurs. Il bénéficie d'une incontestable faveur dans le domaine de la psychologie du raisonnement. En conséquence chaque théorie cherche à étendre son explication initiale à de nouveaux domaines. Cette extension est caractéristique de la théorie des modèles mentaux, proposée d'abord pour le raisonnement déductif catégorique (Johnson-Laird, 1983), étendue au raisonnement déductif propositionnel (Johnson-Laird et Byrne, 1991), puis au raisonnement inductif (Johnson-Laird, 1993) et probabiliste (Johnson-Laird, 1994). La théorie des règles formelles d'inférence montre une semblable ambition même si son extension est encore limitée à la déduction (Rips, 1983, 1994). Le principe de parcimonie est légitime puisqu'il permet de ne pas accorder un grand crédit à une théorie qui se protégerait contre les faits incompatibles avec elle par la multiplication des hypothèses annexes. Cependant il ne doit pas être érigé en dogme. Il est clair que le critère de simplicité est à composer avec d'autres critères, et que la théorie la plus simple n'est pas forcément celle qu'il convient de privilégier si elle couvre moins de données qu'une autre. Par ailleurs on ne saurait garantir dans tous les domaines que la structure cachée de la nature soit forcément la plus simple possible. Ainsi on a parfois évoqué la redondance de la nature dans certains secteurs de la biologie, par exemple à propos du système vasculaire, redondance utile puisqu'elle permet des suppléances. Par ailleurs d'un point de vue formel le fait qu'un système possède plusieurs circuits ou algorithmes ayant des finalités comparables peut constituer un gain en fiabilité et en adaptabilité. Pour ces raisons préalables à un examen des données, l'hypothèse moniste d'une unicité des mécanismes inférentiels à travers la diversité des formes de raisonnement ne nous paraît pas bénéficier *a priori* d'une plus grande plausibilité que l'hypothèse adverse d'une pluralité de mécanismes. Cette dernière pourrait s'avérer plus plausible qu'il n'y paraît au premier abord si on considère l'extrême diversité des raisonnements humains.

Inversement la diversité des formes de raisonnement identifiées en se référant à la diversité des arguments ne permet pas de rejeter *a priori* le monisme théorique. Les principales théories privilégient la déduction, et on peut concevoir plusieurs façons de réduire le raisonnement inductif à un raisonnement déductif (Osherson, Smith et Shafir, 1986). La plus simple consiste à assimiler un argument inductif à un enthymème, c'est-à-dire à un argument déductif dont certaines prémisses demeurent implicites au lieu d'être mentionnées. Une autre façon de faire consiste à introduire une ou deux opérations supplé-

mentaires par rapport à celles qui caractériseraient la déduction. Ainsi pour l'argument suivant qui relève du raisonnement avec des prémisses incertaines, « s'il va à la pêche il attrape *souvent* du poisson ; il est allé à la pêche, donc il a *probablement* attrapé du poisson » : on pourrait invoquer un traitement déductif faisant abstraction de *souvent* dans la première prémisse, complété par une opération supplémentaire introduisant le modal *probablement* dans la conclusion afin de tenir compte de l'incertitude exprimée par *souvent*. Cependant l'hypothèse que toutes les formes de raisonnement se réduisent à la déduction conduit à décrire des procédures fort complexes compte tenu de la limitation de la capacité de calcul du système humain de traitement de l'information (cf. le § 6.7).

Ainsi la question du monisme théorique ne peut être tranchée par la spéculation en l'absence de données empiriques. Or les données disponibles ne permettent pas d'être convaincu de la capacité de l'une ou l'autre des théories existantes à rendre compte de la diversité observée. La thèse selon laquelle la polymorphie des formes de raisonnement correspond à une polymorphie des mécanismes inférentiels mérite l'examen. C'est cette question que cet ouvrage entend examiner. Les neuf chapitres qui en constituent le cœur seront consacrés à une présentation des données empiriques et des principales théories en compétition, avant de revenir dans la conclusion sur les questions soulevées ici. Avant de présenter le plan et les principes directeurs de l'ouvrage il peut être opportun de faire quelques commentaires sur quelques termes qui seront constamment employés.

Les termes « raisonnement » et « inférences » sont généralement utilisés de façon interchangeable en psychologie dans les travaux actuels sur le raisonnement, sans que les différents contextes d'utilisation fassent apparaître des différences manifestes d'acception (cf. le manuel de référence d'Evans *et al.*, 1993). Tout au plus peut-on noter dans des ouvrages plus généraux une certaine propension à utiliser le terme « inférence » lorsqu'on se réfère à des traitements automatiques ou très élémentaires, et le terme raisonnement lorsqu'on veut suggérer qu'il s'agit d'une activité complexe à un titre ou à un autre. À l'extérieur de la psychologie l'acception de raisonnement est souvent plus étroite que celle d'inférence (voir par ex. Blanché, 1973). Deux restrictions sont souvent utilisées. La première est que le raisonnement ferait appel à des inférences explicites s'exprimant dans un langage, quitte à invoquer un langage intérieur en l'absence de verbalisation. La seconde limite le raisonnement aux inférences justifiées ou justifiables par des principes communément admis, en particulier des prin-

cipes logiques (par ex. Kayser, 1993 ; Peirce, 1931). Cette dernière restriction s'accorde bien avec l'étymologie du terme raisonnement. Cependant il n'est pas aisé de savoir si l'auteur d'une inférence est en mesure de la justifier, surtout lorsqu'il s'agit d'un raisonnement court. On peut observer par exemple que les non-logiciens sont souvent très embarrassés si on leur demande de justifier les inférences élémentaires comme « si P alors Q ; P vrai ; donc Q ». On observe également que la contrainte d'avoir à justifier ses inférences n'est pas neutre et modifie l'activité inférentielle. Par contre la justification est plus aisée pour un raisonnement long car elle est constituée alors par l'énumération des raisonnements courts qui ont permis de poser des conclusions intermédiaires. Ainsi, bien que l'usage en psychologie soit de ne pas faire de distinction entre inférence et raisonnement, le fait qu'un raisonnement doive être justifié ou non constitue une dimension importante de celui-ci sur laquelle on reviendra dans la conclusion. En se référant à un usage élargi, on dira qu'une inférence (un raisonnement) est une activité mentale consistant à poser une nouvelle proposition à partir d'informations préalables déjà connues de l'individu ou qui viennent de lui être communiquées, ou consistant à réviser le degré de croyance (de certitude) antérieurement attribué à une proposition.

Deux autres termes souvent utilisés dans ce domaine, « connaissance » et « croyance », requièrent également quelques commentaires. Ces deux termes n'ont usuellement pas le même emploi, le premier se référant à des propositions présumées vraies, le second à des propositions dont la vérité n'est pas établie ou est douteuse. Pour parler de vérité il faut disposer d'un critère, et le critère d'un individu n'est pas toujours celui d'un autre. On pourrait éventuellement se référer à l'état de la science à un moment donné, mais la démarche de la science est si complexe que le prédicat « scientifiquement vrai » ressemble à l'étrange prédicat « v_{leu} » inventé par Goodman (1954) pour caractériser les émeraudes (une émeraude est « v_{leu} » si elle est verte lorsqu'elle est observée avant l'instant *t*, et bleue lorsqu'elle est observée après, *t* désignant par exemple l'an 2000). Surtout, le critère de l'expérimentateur pas toujours celui du sujet (curieusement dans quelques écrits l'auteur se réfère aux « croyances » des sujets, et mentionne ses propres « connaissances »). Ce qui est déterminant dans le processus inférentiel c'est le degré de confiance attribué aux propositions utilisées comme prémisses, les différents critères de vérité intervenant par leur effet sur le degré de certitude attribué à celles-ci. En général ce qui est considéré à tort ou à raison comme une vérité éta-

blie (par l'expérience, par la science, par la religion...) entraîne la certitude. On peut exceptionnellement ignorer certaines certitudes, et dire par exemple «je sais que suis mortel mais je ne peux pas le croire». Pour ces raisons on ne fait généralement pas de distinction entre croyance et connaissance dans divers secteurs de la philosophie de l'esprit ou de l'IA, et nous ferons de même dans cet ouvrage tout en reconnaissant qu'elle peut être pertinente en certaines occasions.

Notre examen de la diversité des formes de raisonnement sera délibérément limité à ceux qui ne sont pas automatiques, ce terme étant défini par rapport à l'opposition entre des processus de traitement *automatiques* et des processus dits *contrôlés*. Les critères les plus fréquemment invoqués sont que les premiers sont très rapides, inconscients, incoercibles, et ne requièrent aucune attention, ce qui permet de les produire parallèlement à d'autres processus. Les seconds par contre ont des caractéristiques inverses qui les rendent vulnérables mais flexibles. Il n'est pas si aisé qu'on pourrait le croire d'identifier des processus automatiques lorsqu'on restreint ce terme aux processus qui satisfont simultanément à tous les critères ci-dessus, et en particulier au dernier. Ils interviennent plus spécialement dans ce que Fodor (1983) a appelé des systèmes ou modules «périphériques» par opposition au système «central». Les premiers mettent en jeu des algorithmes de traitement spécialisés en fonction notamment des diverses entrées sensorielles, et le second confronte les sorties de ces différents systèmes entre elles et avec les croyances antérieures. On peut s'attendre donc à ce que les inférences automatiques soient circonscrites à des activités très élémentaires. De fait si la compréhension du discours oral ou écrit met en jeu des inférences nombreuses et diverses, seules quelques-unes sont véritablement automatiques. Un exemple est constitué par les inférences qui désignent un antécédent comme étant le référent de termes anaphoriques (voir Graesser et Bower, 1990; Kekenbosch, 1994) (par ex. dans l'énoncé «Actéon surprit Artémis au bain; *la déesse irritée le transforma en cerf*», «la déesse» désigne Artémis, et «le» désigne Actéon). Ces inférences sont trop limitées à des opérations très spécialisées pour présenter un intérêt pour notre propos. Celui-ci sera donc restreint aux inférences réalisées dans le cadre d'un traitement contrôlé, tout en considérant qu'il existe différents niveaux de contrôle et qu'on ne met peut-être pas en jeu les mêmes formes de raisonnement à ces différents niveaux.

L'ouvrage est découpé en trois parties. La première concerne les raisonnements élémentaires avec des arguments déductifs, la seconde

des raisonnements élémentaires avec des arguments non déductifs, et la troisième partie le maintien de la cohérence au sein des connaissances ou croyances compte tenu de la multiplicité des inférences réalisées. Dans la première et la seconde partie les chapitres correspondent à des formes de raisonnement définies à partir des catégories des logiciens pour classer les d'arguments. Cette façon de faire permet d'avoir des repères clairs et stables pour déterminer, au sein d'une catégorie d'arguments homogènes du point de vue de leur statut logique, quels sont ceux qui sont acceptés ou rejetés en fonction du contenu ou du contexte. Ainsi dans la première partie le premier chapitre concerne le raisonnement propositionnel avec les connecteurs *et*, *ou*, *ni... ni*, *si... alors...*, un second chapitre réserve un traitement particulier au connecteur *si... alors*, et le troisième chapitre concerne le raisonnement avec les quantificateurs *tous*, *aucun*, *quelques...* Les chapitres 4, 5 et 6 de la deuxième partie examinent respectivement l'induction, l'abduction et le raisonnement avec des prémisses incertaines. Dans la troisième partie le chapitre 7 est consacré à la révocabilité des conclusions avec le raisonnement non monotone, le chapitre 8 à la coordination entre les inférences, et le dernier chapitre aux formes de la rationalité.

L'exposé des résultats empiriques et des principales positions théoriques sera effectué dans ces neuf chapitres en évitant d'adopter des positions trop personnelles dans les principaux débats en cours. Par contre plusieurs partis pris ont été adoptés pour la sélection des informations présentées puisqu'il ne saurait être question d'être exhaustif. Le premier a été d'accorder aux raisonnements ou problèmes examinés, un espace qui soit davantage fonction de leur intérêt intrinsèque que du nombre de travaux qui leur sont consacrés. Ainsi il n'est pas apparu opportun de consacrer plus qu'un paragraphe au fameux paradigme expérimental des quatre cartes de Wason malgré les centaines de publications qu'il a suscitées. Divers auteurs conviennent aujourd'hui qu'il n'est pas le plus adapté à l'étude du raisonnement, ou plus radicalement contestent sa pertinence, car d'une part il est sensible à de nombreux facteurs, et d'autre part une réponse valide requiert une démarche plus complexe que ne le laisse supposer la simplicité de la tâche (voir Andler, 1995). Il n'est pas apparu opportun non plus de relater par le menu les nombreuses expériences conçues pour trancher entre la théorie des modèles mentaux et celle des règles formelles d'inférence bien qu'il s'agisse d'un secteur très animé. Les partisans de chacune de ces théories avancent des résultats empiriques favorables à leur thèse, mais généralement

que de nombreux adultes sont capables de faire des inférences avec des énoncés abstraits constituant des arguments déductifs, ce qui montre qu'ils sont tout à fait capables d'utiliser des indices syntaxiques lorsque l'information sémantique fait défaut.

Il y a deux hypothèses en réponse à la question ci-dessus. La première c'est que l'utilisation exclusive de l'information syntaxique résulte d'un entraînement, et que les individus sont inégalement entraînés à cela par leur scolarité ou leurs interactions avec l'environnement. La seconde, c'est que les individus privilégient l'information syntaxique au détriment de l'information sémantique dans certaines circonstances seulement. La première hypothèse fait appel aux différences interindividuelles, la seconde aux différences entre situations. Le rôle de l'entraînement a parfois été contesté à partir des résultats d'expériences comportant un entraînement limité instauré par l'expérimentateur pour les besoins de l'étude, ou une comparaison de groupes de sujets ayant des niveaux de scolarisation différents. Cependant ce qui s'avère crucial c'est moins le niveau des études que leur nature. Cela a été notamment mis en évidence pour le raisonnement conditionnel en comparant des étudiants en sciences à des étudiants en lettres (Politzer, 1981), et pour la tâche de sélection des quatre cartes de Wason en comparant des étudiants en mathématiques à des étudiants d'autres disciplines scientifiques (Jackson et Griggs, 1988). En ce qui concerne la deuxième hypothèse relative à l'influence des contextes sur le primat accordé à la morphologie ou au contenu sémantique, les données pertinentes ne sont pas nombreuses. Certains résultats, notamment dans la tâche des quatre cartes, suggèrent que le contexte situationnel pourrait avoir un effet via l'imputation d'utilités aux différentes issues (§ 2.5). Une autre composante du contexte résiderait dans la similitude entre la tâche de raisonnement d'une part et des exercices scolaires ou des tâches réelles d'autre part, les premières favorisant la prise en compte des aspects plus formels. On peut interpréter en ce sens l'expérience montrant que 40% des étudiants sont capables de passer d'un raisonnement « fondé sur des suppositions » à un raisonnement « fondé sur des croyances » et *vice versa* en conformité avec la demande de l'expérimentateur (§ 6.9). Une autre composante du contexte réside dans les contraintes ou les suggestions imposées par le mode de présentation des informations à l'activité inférentielle, qui permet de rendre manifestes ou d'occulter les relations entre les éléments des énoncés (Da Silva Neves, 1995 ; § 4.4). Cependant le rôle du contexte a été largement établi ailleurs, soit

dans la résolution de problème en laboratoire (Clément, 1994), soit avec des exercices plus scolaires (Bastien, 1987).

Ces remarques suggèrent qu'une déduction « pure », c'est-à-dire fondée uniquement sur les aspects formels des prémisses de l'argument et non sur leur contenu, n'existe que dans des conditions particulières. Elle n'interviendrait que dans deux cas, soit comme inférence par défaut en l'absence de tout contenu sémantique, soit comme inférence « corrigée » par inhibition de celui-ci s'il est présent. Adopter cette façon de voir conduit à nouveau à mettre en doute l'affirmation que la déduction serait fondamentale en ce sens supplémentaire qu'elle serait « basique ». Par ce terme on se réfère à l'idée qu'elle constituerait le fondement sur lequel les autres raisonnements pourraient s'édifier grâce à l'introduction d'informations et d'opérations supplémentaires.

Examinons maintenant les données relatives au deuxième objectif de l'étude du raisonnement, à savoir l'identification du ou des mécanismes(s) inférentiel(s) responsable(s) de sa production. Rappelons que deux mécanismes sont fréquemment invoqués actuellement à propos de la déduction, les règles formelles d'inférence (la « logique mentale ») et les modèles mentaux, et accessoirement les schémas pragmatiques et un mécanisme de propagation dans un réseau connexionniste. Les données empiriques disponibles ont été présentées succinctement dans les chapitres 1 à 9. Disons rapidement que leur examen aboutit à une conclusion très différente de celle qui est habituellement attendue. En effet le problème est présenté très souvent sous la forme d'une alternative entre deux options mutuellement exclusives, en sorte qu'opter pour l'une équivaut à rejeter l'autre, ou *vice versa*. Notre conclusion conduit à rejeter ce genre d'alternative. Elle conduit par surcroît à remettre en cause les termes du choix en considérant que la théorie des modèles mentaux et la théorie des règles formelles d'inférence doivent toutes les deux être révisées, parce que sous leur forme actuelle aucune d'elles ne correspond exactement aux données recueillies.

Formuler le problème du choix entre des mécanismes inférentiels sous forme d'une alternative correspond au monisme théorique mentionné dans l'introduction. Cette option stipule qu'un seul mécanisme est responsable de toutes les formes de raisonnement. Elle s'oppose à une option pluraliste qui comporte deux versions. Dans une version les différents mécanismes inférentiels seraient en quelque sorte spécialisés pour un raisonnement particulier, et il y aurait une correspondance bijective entre mécanismes et formes de raisonnement. Dans la

seconde version les mécanismes seraient plus polyvalents, en sorte que plusieurs mécanismes distincts pourraient intervenir dans une même forme de raisonnement, ce qui rendrait possible des suppléances et une relative flexibilité.

En adoptant provisoirement le point de vue du monisme théorique, il existe actuellement non pas deux candidats sérieux mais trois, la théorie des règles formelles *révisée*, la théorie des modèles mentaux *révisée* et le connexionnisme. La première se prête mal à la modélisation du raisonnement inductif ou sous incertitude. Certes, elle peut être étendue par l'incorporation de règles d'inférence non déductives (cf. Collins et Michalski, 1989 ; Collins *et al.*, 1975), et par l'incorporation de règles de type probabiliste, mais cette extension ne permet pas d'expliquer comment on attribue des forces différentes à des arguments de même structure syntaxique (cf. le § 4.3). La théorie des modèles mentaux s'affirme théorie universelle quoique incomplète (Johnson-Laird et Byrne, 1991, 205-206, 215), et elle a proposé des extensions pour le raisonnement inductif et probabiliste (Johnson-Laird, 1993, 1994).

Cette ambition se heurte à au moins deux objections. La première concerne le coût computationnel de cette extension. Ce type de critique a été formulé par deux partisans du connexionnisme, Oaksford et Chater (1993). Ils considèrent que les deux principales théories de la déduction nécessitent des ressources qui excèdent la capacité du système humain de traitement, et ils les rejettent toutes deux comme étant invalides pour cette raison. La critique est excessive si on considère que les problèmes de déduction que traitent les profanes sont d'un ordre de complexité très inférieur à celui des problèmes traités par le logicien. Même en prenant en considération la multiplicité des connaissances qui peuvent être activées par le contenu sémantique des prémisses, leur nombre est limité par des considérations de seuil d'activation ou de pertinence selon le niveau d'analyse auquel on se place. En outre la complexité computationnelle serait excessive pour un système qui ne commettrait pas d'erreur, non pour un système qui serait faillible comme c'est le cas pour le système humain. Par contre l'objection d'un coût prohibitif demeure si on considère la proposition de Johnson-Laird (1994) pour représenter les différents degrés de plausibilité attribués aux prémisses et à la conclusion. Rappelons que celle-ci consiste à représenter la plausibilité d'une proposition en créant des clones du modèle de celle-ci, plus des clones du modèle de la proposition alternative. Ces clones doivent être en nombre suffisant pour que leur rapport soit conforme à la plausibilité attribuée à

chaque proposition. Il en résulte une multiplication du nombre de modèles requis pour la moindre inférence avec des prémisses incertaines, ce qui rend cette démarche très lourde et partant fort peu vraisemblable (cf. le § 6.8).

La deuxième objection résulte des différences interindividuelles, et plus spécialement des différences qui concernent non pas le niveau de performance, mais le mode opératoire. Il est très difficile d'obtenir des indications directes de la démarche adaptée par les sujets. On a obtenu des indications favorables à l'existence de deux démarches différentes par des procédures indirectes en mesurant le temps de réaction pour diverses catégories de réponse ou en imposant des contraintes de temps (Galotti, Baron et Sabini, 1986; voir aussi Robert, 1993). Cependant l'observation fine du comportement dans des situations très ouvertes où les sujets peuvent utiliser un brouillon ou faire des manipulations pour résoudre des syllogismes catégoriques est très défavorable à la théorie des modèles mentaux sous la forme définie par Johnson-Laird. En effet l'une des deux démarches observées chez les sujets utilise des « modèles », mais ceux-ci reposent sur une représentation des classes elles-mêmes du type des cercles d'Euler, et non sur une représentation des éléments des classes comme il est affirmé pour ce type d'inférences (Ford, 1995).

Le connexionnisme constitue un autre candidat potentiel au monisme théorique. Certes il n'a pas fait preuve jusqu'à présent d'une aptitude comparable à celle de la théorie des modèles mentaux à rendre compte des principaux phénomènes observés avec les divers types de syllogismes déductifs. Par contre il semble mieux adapté à modéliser les raisonnements non déductifs, comme le raisonnement inductif (par ex. Sloman, 1993; voir le § 4.8) ou le raisonnement non monotone (par ex. Balkenius et Gärdenfors, 1991). Cependant l'aptitude à modéliser des états représentationnels complexes n'entraîne pas l'aptitude à confronter simultanément plusieurs propositions distinctes pour déterminer leur statut de vérité respectif comme dans le dilemme du prisonnier ou les célèbres problèmes de Smullyan (1984). Cela ne permet pas non plus d'expliquer les phénomènes de représentation croisée dans lesquels un individu considère les anticipations ou raisonnements d'autrui, qui doit lui-même intégrer les anticipations et raisonnements du premier, qui doit à son tour tenir compte de cette représentation modifiée, etc. (Walliser, 1985). Or ces différents problèmes relèvent de la compétence des individus. Notons également que le connexionnisme comme monisme se heurte lui aussi à l'objection des différences interindividuelles dans le mode opératoire. Cette

dernière objection pourrait cependant être levée si on considère que le connexionnisme décrit une étape préliminaire du raisonnement dans laquelle des processus subsymboliques permettraient d'élaborer des représentations qui feraient ensuite l'objet d'un traitement d'une nature différente.

L'option pluraliste mentionnée précédemment, au lieu de privilégier un mécanisme inférentiel en excluant tous les autres, affecte chacun d'eux à une forme particulière de raisonnement. Elle a déjà été évoquée par Rips (1990) lorsqu'il envisage l'hypothèse selon laquelle les théories relevant du point de vue « strict » ou discret comme les théories des règles formelles d'inférence concerneraient les arguments déductifs, tandis que les théories relevant du point de vue « lâche » ou continu comme le connexionnisme concerneraient la révision des croyances (cf. l'introduction). Partisan des règles formelles d'inférence, Rips rejette cette option car il rejette le deuxième point de vue, considérant qu'il ne permet pas d'expliquer d'où viennent les croyances ni de traiter les relations causales. Cependant l'option pluraliste mérite un examen plus attentif car d'une part la correspondance entre mécanismes et raisonnements peut prendre diverses formes et pas seulement celle qui est envisagée par Rips, et d'autre part les deux reproches adressés au point de vue « lâche » ne sont pas rédhibitoires et concernent également le point de vue « strict ».

Sous sa forme la plus générale, l'option pluraliste présente l'attrait d'être compatible avec la grande diversité des procédures de raisonnement qui ont été soit observées directement, soit inférées et testées à partir des propositions d'une théorie. Elle pourrait comporter une difficulté préalable si on considère qu'elle présuppose la capacité de différencier différentes formes d'arguments ou de raisonnements, et que cette capacité n'est pas établie. Cependant, toute distinction entre des formes de raisonnement ne repose pas nécessairement sur des critères liés aux caractéristiques des arguments, mais aussi sur des critères liés aux modalités de traitement. Ainsi il semble opportun de faire une distinction entre un raisonnement *implicite* qui interviendrait constamment lors de tout apport d'informations nouvelles et permettrait de se faire une opinion personnelle et un raisonnement *explicite* qui interviendrait lorsqu'on veut justifier son opinion pour autrui ou pour soi-même. Les termes *explicite* et *implicite* sont préférés ici aux termes *démonstratif* et *non démonstratif* pour bien marquer que la recherche de justification peut concerner également des arguments inductifs ou probabilistes, le terme *démonstratif* étant réservé par de nombreux auteurs au raisonnement déductif.

Le raisonnement explicite, contrairement au raisonnement implicite, nécessite qu'on identifie les propositions utilisées comme prémisses de l'inférence ainsi que le type d'argument concerné. Cette seconde condition concerne la perception d'une relation particulière entre prémisses et conclusion qui soit considérée comme légitime et communément admise, et non pas la référence à quelque principe logique. Le raisonnement explicite intervient dans la démonstration, l'argumentation et dans les situations où on juge utile de pouvoir garantir ses conclusions, par exemple parce qu'on anticipe une demande de justification par des tiers. Le raisonnement implicite concerne le réajustement permanent du stock de croyances de l'individu sous l'effet des apports d'informations nouvelles, et il fait appel à des critères ou opérations peu accessibles à la conscience, ce qu'on appelle communément l'intuition. On peut considérer que les données expérimentales recueillies par les psychologues concernent habituellement le raisonnement explicite, mais que les sujets pourraient avoir recours à un raisonnement implicite soit parce que la situation est trop contraignante, soit faute de découvrir une justification. Compte tenu de la distinction entre raisonnement explicite et implicite, la capacité de différencier différentes sortes d'arguments ne serait requise que pour le premier. Or cette capacité appartiendrait bien à la compétence humaine comme le montre une expérience exploratoire dans laquelle les sujets devaient classer 24 arguments déductifs, inductifs et abductifs en autant de catégories qu'ils le jugeaient souhaitable. Les arguments ont été répartis entre une, deux, trois et plus de trois catégories par respectivement 13 %, 20 %, 44 % et 23 % des sujets. L'analyse mathématique des clusters fait apparaître trois regroupements dominants, un pour les arguments déductifs, un pour les arguments avec induction classique (généralisation), et le dernier pour l'abduction et l'extrapolation (George, 1996b).

Spécifier la forme précise que devrait prendre l'option pluraliste est hasardeux en l'état actuel de nos connaissances. Tout au plus peut-on en esquisser certains contours en formulant quelques hypothèses. On peut supposer tout d'abord que le raisonnement implicite met en jeu des mécanismes inférentiels ayant un coût computationnel faible ou nul, et le raisonnement explicite des mécanismes ayant un coût élevé. La raison en est que le premier survient souvent parallèlement à d'autres activités et relève donc de processus automatiques, alors que le second comporte un coût attentionnel et relève donc de processus contrôlés. La seconde hypothèse est que le raisonnement implicite devrait faire appel à des mécanismes inférentiels très polyvalents.

lents et peu nombreux, voire à un seul, contrairement au raisonnement explicite. En effet, certains mécanismes polyvalents permettent d'éviter d'avoir à analyser les informations pour sélectionner celles qui seront constitutives d'un argument, et d'avoir à catégoriser celui-ci pour déterminer la séquence d'opérations appropriées. Par contre pour être en mesure de justifier une inférence, il faut isoler les informations et les indices sémantiques ou morphologiques permettant de montrer que certaines prémisses entraînent ou favorisent la conclusion proposée, ce qui suggère une spécialisation et une diversification des mécanismes inférentiels. La troisième hypothèse est que le raisonnement implicite peut faire intervenir des critères non verbalisables contrairement au raisonnement explicite. En effet un raisonnement « privé » peut être convaincant pour son auteur même si celui-ci n'est pas en mesure d'expliquer sur quoi repose sa conviction intime, alors que dans un raisonnement « public » ou susceptible de l'être il faut transmettre toutes les informations utilisées afin que le destinataire reconstitue l'inférence pour son propre compte.

Ces hypothèses suggèrent que le raisonnement implicite pourrait faire appel à des mécanismes inférentiels semblables aux « heuristiques » décrites par Kahneman et Tversky (Kahneman, Slovic et Tversky, 1982). Le processus inférentiel consisterait en une évaluation du degré de confirmation d'une proposition par d'autres propositions mieux établies en se référant à la « force » de certaines variables permettant d'établir une relation entre la première et les secondes. Tversky et Kahneman mentionnent la représentativité et la disponibilité (§ 6.7) mais les données suggèrent que plusieurs autres variables affectent cette évaluation comme la similitude, la typicalité (cf. les § 4.3 et 4.8), la force de la relation causale (cf. le § 4.7). Les différentes variables qui concernent diverses sortes de relations entre deux propositions peuvent éventuellement se combiner en une variable unique de type associatif si ce qui les différencie est ignoré ou devient indiscernable. Dans ce cas limite les inférences réalisables seraient prédictibles par la structure de ces associations comme le stipulent les modèles connexionnistes. Un tel mécanisme connexionniste serait totalement « aveugle » en ce sens qu'il est dépourvu de toute possibilité de sélectivité. Par contre on peut présumer que, lorsqu'il est possible de différencier plusieurs sortes de relations entre deux propositions, celles-ci n'ont vraisemblablement pas toutes le même poids, les relations causales étant vraisemblablement jugées plus dignes de confiance que des relations contingentes de coprésence par exemple.

Le raisonnement explicite ferait appel notamment aux deux mécanismes les plus souvent invoqués à propos de la déduction, les règles formelles d'inférence et les modèles mentaux. Les premières devraient être révisées et élargies. Elles devraient être révisées en renonçant à stipuler qu'elles sont automatiques et universelles, puisque certains résultats suggèrent plutôt qu'elles fonctionnent comme des règles par défaut (§ 7.7) et qu'elles sont très dépendantes du niveau et du type de scolarité (cf. *supra*). Elles devraient être élargies de façon à englober des règles ou schémas d'inférence relatifs à des aspects spécifiques du réel comme l'espace et le temps, par exemple la transitivité de la relation *Plus Loin* (x, y) ou *Avant* (x, y). Les modèles mentaux devraient également être révisés et élargis, ne serait-ce que parce que le support représentationnel invoqué pour les syllogismes catégoriques est inadéquat (cf. *supra*). Il n'y a aucune raison de penser que les deux mécanismes précédents sont les seuls mécanismes disponibles pour le raisonnement explicite. On peut par exemple avancer l'hypothèse qu'une partie des inférences déductives repose sur une ou plusieurs paraphrases linguistiques permettant de transformer une proposition donnée en une autre jugée à tort ou à raison équivalente, ce qui pourrait éviter d'avoir recours à des règles spécifiques ou à une construction de modèles mentaux. Ainsi l'inférence *si p alors q, or non-q, donc non-p* pourrait reposer sur une transformation de la première proposition en *p seulement si q*, puis une nouvelle transformation en *si non-q alors non-p*, et enfin sur le recours à l'inférence commune *Modus Ponens*. D'autres formes de substitutions verbales ont été relevées pour les syllogismes catégoriques (Ford, 1995), et d'autres peuvent reposer sur les inférences immédiates fondées sur des relations d'implication ou d'incompatibilité entre deux propositions, option déjà envisagée par le philosophe Cherniak (1986). Dans l'hypothèse où il existerait plusieurs mécanismes inférentiels disponibles, la sélection de l'un d'eux serait déterminée par son adéquation à la structure des propositions constituant les prémisses, et si plusieurs sont adéquats, par un ordre de préférence fondé par exemple sur la fréquence ou sur la simplicité d'utilisation.

Comme on a stipulé par ailleurs que le raisonnement explicite n'est pas limité aux arguments déductifs, les mécanismes mentionnés devraient permettre d'exprimer qu'une conclusion est plausible mais non certaine, ou ne concerne qu'une fraction des membres d'une classe. Ces mécanismes le permettent de diverses façons, par exemple en relevant que l'une des prémisses d'un argument déductif est incertaine et en transférant cette incertitude à la conclusion (§ 6.3), en uti-

lisant des quantificateurs plus fins que ceux qui sont usuellement étudiés (§ 3.4) comme *la plupart des X* ou *peu de X*, en soulignant que l'adjonction d'une prémisse supplémentaire transforme l'argument en un argument déductivement valide (§ 4.3), etc. Il est même possible d'utiliser des règles simples de maximum et de minimum pour exprimer différents degrés de plausibilité de la conclusion à partir de celle des prémisses (§ 6.6).

On aurait ainsi deux groupes de mécanismes inférentiels, les uns pour le raisonnement implicite, les autres pour le raisonnement explicite, les premiers dépendant principalement du contenu sémantique et de la force des croyances, les seconds s'appuyant principalement sur la morphologie des arguments et pouvant se libérer de la force des croyances. Ces deux groupes seraient distincts mais non mutuellement exclusifs. Plus précisément notre hypothèse est que le raisonnement implicite serait difficile à inhiber à moins que le caractère abstrait des propositions considérées ne permette pas son intervention. Le raisonnement explicite serait précédé ou accompagné par des processus relevant du raisonnement implicite. Cette hypothèse résulte des observations précédentes sur la difficulté de faire abstraction du contenu sémantique des prémisses et de la supposition que le raisonnement implicite comporterait un coût computationnel faible ou nul. Ainsi le raisonnement implicite serait beaucoup plus fréquent parce qu'il interviendrait soit seul soit suivi d'un raisonnement explicite, alors que le raisonnement explicite serait habituellement accompagné d'un raisonnement implicite préliminaire. Cela expliquerait pourquoi le mode de raisonnement « basé sur des croyances » est plus fréquent et plus aisé à adopter que le mode « fondé sur des suppositions » (§ 6.9). Par ailleurs l'hypothèse d'une intervention successive de deux modes de raisonnement différents, implicite puis explicite, conduit à envisager une possibilité de conflit entre leurs sorties respectives. Ce conflit devrait se manifester par des rectifications de la réponse initiale à un problème, et par une instabilité de la performance à des problèmes de même type. Ces deux phénomènes sont observés mais ils ont rarement fait l'objet de discussions. Le conflit peut être évité si le second processus est conduit de façon à ce que sa conclusion s'aligne sur celle du premier, c'est-à-dire s'il consiste en une justification ou rationalisation de la réponse initiale. L'intervention de rationalisation dans le raisonnement est également attestée.

Nous pensons que les travaux ultérieurs confirmeront le bien-fondé d'une approche élargie à d'autres formes de raisonnement que la déduction et débarrassée du dogme métathéorique de l'existence

d'un mécanisme inférentiel unique et omniprésent. Nous avons avancé diverses hypothèses au cours de cette conclusion, hypothèses qui devront être spécifiées ou remaniées ultérieurement à la suite de l'obtention de données nouvelles. Ces hypothèses ne font qu'esquisser l'une des formes possibles d'une approche centrée sur la polymorphie du raisonnement. D'autres formes sont concevables. Les travaux ultérieurs devraient poursuivre cette exploration.



Bibliographie

- Adams E., *The logic of conditionals*, Dordrecht, Reidel, 1975.
- Ahn W., Kalish C. W., Medin D. L. et Gelman S. A., The role of covariation versus mechanism information in causal attribution, *Cognition*, 54, 1995, 299-352.
- Alloy L. B. et Tabachnik N., Assessment of covariation by humans and animals, the joint influence of prior expectations and current situational information, *Psychological Review*, 91, 1984, 112-149.
- Anderson J. R., *The architecture of cognition*, Hillsdale, Lawrence Erlbaum Associates, 1983.
- Anderson J. R. et Sheu C. F., Causal inferences as perceptual judgements, *Memory and Cognition*, 23, 1995, 510-524.
- Andler D., Logique, raisonnement et psychologie, in J. Dubucs et F. Lepage, *op. cit.*, 1995.
- Balkenius C. et Gärdenfors P., Nonmonotonic inferences in neural networks, in J. Allen, R. Fikes et E. Sandewall (Eds), *Principles of knowledge representation and reasoning*, San Mateo, Morgan Kaufman, 1991.
- Bastien C., *Schémes et stratégies dans l'activité cognitive de l'enfant*, Paris, PUF, 1987.
- Beauvois J.-L. et Joule R. V., *Soumission et idéologies*, Paris, PUF, 1987.
- Begg I. et Harris G., On the interpretation of syllogisms, *Journal of Verbal Learning and Verbal Behavior*, 21, 1982, 595-620.
- Bereiter C. et Scardamalia M., When weak explanations prevail, *Behavioral and Brain Sciences*, 12, 1989, 468-469.
- Bisseret A., Figeac-Letang C. et Falzon P., Modélisation de raisonnements opportunistes : l'activité des spécialistes de régulation des carrefours à feu, *Psychologie française*, 33, 1988, 161-168.
- Blanché R., *Le raisonnement*, Paris, Armand Colin, 1973.
- Borel M.-J., Grize J.-B., Miéville D., *Essai de logique naturelle*, Berne, Peter Lang, 1983.
- Boutilier C., Conditional logic of normality : A modal approach, *Artificial Intelligence*, 68, 1983, 87-154.

- Braine M. D. S., On the relation between the natural logic of reasoning and standard logic, *Psychological Review*, 85, 1978, 1-21.
- Braine M. D. S., The « natural logic » approach to reasoning, in W. F. Overton (Ed.), *Reasoning, necessity, and logic. Developmental perspectives*, Hillsdale, NJ, Erlbaum, 1990.
- Braine M. D. S. et O'Brien D. P., A theory of *If*: A lexical entry, reasoning program, and pragmatic principles, *Psychological Review*, 98, 1991, 182-203.
- Braine M. D. S., Reiser B. J. et Rumin B., Some empirical justification for a theory of natural propositional logic, in G. H. Bower (Ed.), *The psychology of learning and motivation* (vol. 18, p. 313-371), New York, Academic Press, 1984.
- Byrne R. M. Suppressing valid inferences with conditionals, *Cognition*, 31, 1989, 61-83.
- Carnap R. et Jeffrey R. (Eds), *Logical foundations of probability*, Berkeley, University of California Press, 1971.
- Chan D. et Chua F., Suppression of valid inferences : Syntactic views, mental models, and relative salience, *Cognition*, 53, 1994, 217-238.
- Charniak E. et McDermott D., *Introduction to Artificial Intelligence*, New York, Addison-Wesley, 1985.
- Cheng P. W. et Holyoak R. J., Pragmatic reasoning schemas, *Cognitive Psychology*, 17, 1985, 391-416.
- Cheng P. W. et Novick L. R., A probabilistic contrast model of causal induction, *Journal of Personality and Social Psychology*, 58, 1990, 545-567.
- Cherniak C., *Minimal rationality*, Cambridge, MIT Press, 1986.
- Chomsky N., *Aspects of the theory of syntax*, Cambridge, MIT Press, 1965.
- Church A., *Introduction to mathematical logic*, Princeton, Princeton University Press, 1956.
- Clément E., *La représentation de l'action : l'interprétation des consignes dans les problèmes isomorphes*, thèse de doctorat de l'Université de Paris VIII, 1994.
- Cohen L. J., Can human irrationality be experimentally demonstrated?, *Behavioral and Brain Sciences*, 4, 1981, 317-370.
- Collins A. et Michalski R., The logic of plausible reasoning : A core theory, *Cognitive Science*, 13, 1989, 1-49.
- Collins A., Warnock E. H., Aiello N. et Miller M. L., Reasoning from incomplete knowledge, in D. G. Bobrow et A. Collins (Eds.), *Representation and Understanding*, New York, Academic Press, 1975.
- Corson Y., Recherche, test et chronologie d'hypothèses en situations ouvertes, *Cahiers de psychologie cognitive*, 4, 1984, 421-447.
- Cosmides L., The logic of social exchange : Has natural selection shaped how humans reason ? Studies with the Wason selection task, *Cognition*, 31, 1989, 187-276.
- Cummins D. D., Lubart T., Alkins O. et Rist R., Conditional reasoning and causation, *Memory and Cognition*, 19, 1991, 274-282.
- Da Silva Neves R. M., *Interaction entre induction et déduction dans le raisonnement humain*, thèse de doctorat non publiée, Université de Paris VIII, 1995.
- Doherty M. E., Mynatt C. R., Tweney R. D. et Schiavo M. D., Pseudodiagnosticity, *Acta Psychologica*, 43, 1979, 162-167.

- Dominovski R. L., How people discover concepts ?, in R. L. Solso (Ed.), *Theories in cognitive psychology : The Loyola symposium*, Potomac, Lawrence Erlbaum, 1974.
- Ducrot O., *Dire et ne pas dire*, Paris, Hermann, 1972.
- Dubucs J. et Lepage F. (Eds), *Méthodes logiques pour les sciences cognitives*, Paris, Hermès, 1995.
- Einhorn H. J. et Hogarth R. M., Judging probable cause, *Psychological Bulletin*, 99, 1986, 3-19.
- Elio R. et Pelletier F. J., The effect of syntactic form on simple belief revisions and updates, in A. Ram et K. Eiselt (Eds), *Proceedings of the Sixteenth Annual Conference of the Cognitive Society*, 1994, 260-265.
- Engel P., *Introduction à la philosophie de l'esprit*, Paris, La Découverte, 1994.
- Evans J. St. B. T., *The psychology of deductive reasoning*, Londres, Routledge & Kegan Paul, 1982.
- Evans J. St. B. T., *Bias in human reasoning : Causes and consequences*, Hove, Lawrence Erlbaum Associates, 1989.
- Evans J. St. B. T., Bias and rationality, in Manktelow et Over, *op. cit.*, 1993.
- Evans J. St. B. T., Barston J. L. et Pollard P., On the conflict between logic and belief in syllogistic reasoning, *Memory and Cognition*, 11, 1983, 295-306.
- Evans J. St. B. T. et Newstead S. E., A study of disjunctive reasoning, *Psychological Research*, 41, 1980, 373-388.
- Evans J. St. B. T., Newstead S. E. et Byrne R. M. J., *Human reasoning : The psychology of deduction*, Londres, Lawrence Erlbaum Associates, 1993.
- Fillenbaum S., OR : Some uses, *Journal of Experimental Psychology*, 102, 1974, 913-921.
- Fillenbaum S., Inducements : On the phrasing and logic of conditional promises, threats, and warnings, *Psychological Research*, 38, 1976, 231-250.
- Fillenbaum S., Mind your p 's and q 's : The role of content and context in some uses of *and*, *or* and *if*, in G. H. Bower (Ed.), *The psychology of learning and motivation* (vol. 11), New York, Academic Press, 1977.
- Fodor J. A., *The langage of thought*, Cambridge, Harvard University Press, 1975.
- Fodor J. A., *The modularity of mind*, Cambridge, MIT Press, 1983 (trad. franç. : *La modularité de l'esprit*, Paris, Éditions de Minuit).
- Ford M., Two modes of mental representation and problem solution in syllogistic reasoning, *Cognition*, 54, 1995, 1-71.
- Frege G., *Begriffsschrift*, Halle, Nebert, 1879 (trad. angl. : J. van Heijenoort, From Frege to Gödel, Harvard, Harvard University Press, 1967).
- Gärdenfors P. et Makinson D., Nonmonotonic inference based on expectations, *Artificial Intelligence*, 65, 1994, 197-245.
- Galotti K. M., Baron J. et Sabini J. P., Individual differences in syllogistic reasoning : Deduction rules or mental models ?, *Journal of Experimental Psychology : General*, 115, 1986, 16-25.
- Gentzen, G., Untersuchungen über das logische Schliessen, *Mathematische Zeitschrift*, 39, 1935, 176-210, 405-431 (trad. franç. et présentation par R. Feys et J. Ladrière, *Recherche sur la déduction logique*, Paris, PUF, 1955 ; trad. angl. : *The collected papers of Gerhard Gentzen*, édité by M. E. Szabo, Amsterdam, North Holland, 1969).

- George C., Attention et traitement de l'information dans l'apprentissage instrumental humain, *L'Année psychologique*, 80, 1980, 481-500.
- George C., Généralisation déductive et inductive dans un apprentissage par l'action, *L'Année psychologique*, 86, 1986, 183-200.
- George C., Dissociation des difficultés dans la tâche de sélection de Wason, *L'Année psychologique*, 90, 1990, 169-193.
- George C., Facilitation in the Wason's selection task with a consequent referring to an unsatisfactory outcome, *British Journal of Psychology*, 82, 1991, 463-472.
- George C., Rules of inference in the interpretation of the conditional connective, *Cahiers de psychologie cognitive / European Bulletin of Cognitive Psychology*, 12, 1992, 115-140.
- George C., Discrepancies between response patterns on conditional reasoning tasks : computational errors or changes in interpretation ?, *Cahiers de psychologie cognitive / Current Psychology of Cognition*, 13, 1994, 821-832.
- George C., The endorsement of the premises : Assumption-based or belief-based reasoning, *British Journal of Psychology*, 86, 1995, 93-111.
- George C., *Reasoning with uncertain premises, Thinking and Reasoning*, sous presse, 1996a.
- George C., *La classification des arguments par des non-logiciens*, manuscrit soumis pour publication, 1996b.
- George C. et Politzer G., *Propositional reasoning as constraint satisfaction*, manuscrit soumis pour publication, 1996.
- Gigerenzer G., How to make cognitive illusions disappear : Beyond « Heuristics and Biases », *European Journal of Social Psychology*, 2, 1991, 83-115.
- Gineste M.-D. (sous presse), *Analogie et cognition*, Paris, PUF.
- Gochet P. et Gribomont P., *Logique : méthodes pour l'informatique fondamentale*, vol. 1, Paris, Hermès, 1990.
- Godu B., *Stratégies de raisonnement et supposition*, mémoire de maîtrise non publié, Université de Paris VIII, 1995.
- Goodman N., *Fact, fiction and forecast*, Cambridge, Harvard University Press, 1954 ; trad. franç. : *Faits, fictions et prédictions*, Paris, Éditions de Minuit, 1984.
- Graesser A. C. et Bower G. H. (Eds), *Inferences and text comprehension*, New York, Academic Press, 1990.
- Grice H. P., Logic and conversation, in P. Cole et J. L. Morgan (Eds), *Syntax and semantics*, 1975, vol. 3 : *Speech acts*, New York, Academic Press (trad. franç., in *Communications*, 1979, 30, 57-71).
- Griggs R. A. et Cox J. R., The elusive thematic materials effect in the Wason selection task, *British Journal of Psychology*, 73, 1982, 407-420.
- Halpern J. Y. et Moses, Y., A guide to the modal logics of knowledge and belief : Preliminary draft, *Proceedings of the 9th International Joint Conference on Artificial Intelligence*, 1985, 480-490.
- Harman G., *Change in view*, Cambridge, MIT Press, 1986.
- Haton J.-P., Bouzid N., Charpillat F., Haton M.-C., Lâasri B., Lâasri H., Marqui P., Mondot T. et Napoli A., *Le raisonnement en intelligence artificielle*, Paris, InterÉditions, 1991.

- Heller R. F., Saltzstein H. D. et Caspe W. B., Heuristics in medical et non-medical decision-making, *Quarterly Journal of Experimental Psychology*, 44A, 1992, 211-235.
- Hempel C. G., A purely syntactical definition of confirmation, *Journal of Symbolic Logic*, 8, 1943, 122-143.
- Henle M., On the relation between logic and thinking, *Psychological Review*, 69, 1962, 366-378.
- Hewson C. et Vogel C., Psychological evidence for assumptions of path-based inheritance reasoning, in A. Ram et K. Eiselt (Eds), *Proceedings of the 16th Annual Conference of the Cognitive Society*, 1994, 409-414.
- Hilton D. J., Conversational processes and causal explanation, *Psychological Bulletin*, 107, 1990, 65-81.
- Hilton D. J. et Slugoski B. R., Knowledge-based causal attribution : The abnormal condition focus model, *Psychological Review*, 93, 1986, 75-88.
- Hintikka J., *Knowledge and belief : An introduction to the logic of the two notions*, Ithaca, Cornell University Press, 1962.
- Holland J. H., Holyoak K. J., Nisbett R. E. et Thagard P. R., *Induction : Processes of inference, learning and discovery*, Cambridge, MA, Bradford Books et MIT Press, 1986.
- Houdé O. et Miéville D., *Pensée logico-mathématique*, Paris, PUF, 1993.
- Hume D., *A treatise of human nature*, 1739-1740 (trad. franç. : *Traité de la nature humaine*, Paris, Aubier, 1966).
- Inhelder B. et Piaget J., *De la logique de l'enfant à la logique de l'adolescent*, Paris, PUF, 1955.
- Jackson F. (Ed.), *Conditionals*, Oxford, Oxford University Press, 1991.
- Jackson S. L. et Griggs R. A., Education and the selection task, *Bulletin of the Psychonomic Society*, 26, 1988, 337-330.
- Jacob P., *L'empirisme logique*, Paris, Éd. de Minuit, 1980.
- Johnson-Laird P. N., *Mental models*, Cambridge, Cambridge University Press, 1983.
- Johnson-Laird P. N., *Human and machine thinking*, Hillsdale, Erlbaum, 1993.
- Johnson-Laird P. N., Mental models and probabilistic thinking, *Cognition*, 50, 1994, 189-209.
- Johnson-Laird P. N. et Bara B. G., Syllogistic inference, *Cognition*, 16, 1984, 1-61.
- Johnson-Laird P. N. et Byrne R., *Deduction*, Hillsdale, Erlbaum, 1991.
- Johnson-Laird P. N. et Tagard J., How implication is understood, *American Journal of Psychology*, 10, 1969, 367-373.
- Johnson-Laird P. N., Byrne R. et Schaeken W., Propositional reasoning by model, *Psychological Review*, 99, 1992, 418-439.
- Johnson T. R., Krems J. et Amra N. K., A computational model of human abductive skill and its acquisition, *Proceedings of the 16th Annual Conference of the Cognitive Science Society*, 1994.
- Kahneman D., Slovic P. et Tversky A. (Eds), *Judgement under uncertainty : Heuristics and biases*, Cambridge, Cambridge University Press, 1982.
- Kant E., *Kritik der reinen Vernunft*, 1787 (trad. franç. : *Critique de la raison pure*, Paris, PUF, 1971).

- Kayser D., Un point de vue artificialiste sur le raisonnement, in J.-F. Le Ny (Ed.), *Intelligence naturelle et intelligence artificielle*, Paris, PUF, 1993.
- Kelley H. H., The process of causal attribution, *American Psychologist*, 28, 1973, 103-128.
- Kekenbosch C., *La mémoire et le langage*, Paris, Nathan, 1994.
- Klahr D. et Dunbar K., Dual space search during scientific reasoning, *Cognitive Psychology*, 12, 1988, 1-48.
- Klayman J. et Brown K., Debias the environment instead of the judge : An alternative approach to reducing error in diagnostic (and other) judgement, *Cognition*, 49, 1993, 97-122.
- Klayman J. et Ha Y.-W., Confirmation, disconfirmation and information in hypothesis testing, *Psychological Review*, 94, 1987, 211-228.
- Kleene S. C., *Mathematical logic*, New York, Wiley, 1967 (trad. franç. : *Logique mathématique*, Paris, Armand Colin, 1971).
- Lea Sombé, *Raisonnements sur des informations incomplètes en Intelligence artificielle*, Toulouse, Techna, 1989.
- Léonard F., Un modèle du sujet : l'équilibre de Heider, in S. Moscovici, *Introduction à la psychologie sociale*, vol. I, Paris, Larousse, 1972.
- Lucas, F., *La flexibilité du raisonnement*, mémoire de maîtrise non publié, Université de Paris VIII, 1994.
- Mackie J. L., Causes and causation, *American Philosophical Quarterly*, 1965, 245-255 et 261-264 ; reproduit in E. Sosa et M. Tooley, *Causation*, Oxford, Oxford University Press.
- Malherbe J.-F., *Épistémologies anglo-saxonnes*, Paris, PUF, 1981.
- Manktelow K. I. et Over D. E., Social rules and utilities in reasoning with deontic conditionals, *Cognition*, 43, 1991, 85-105.
- Manktelow K. I. et Over D. E. (Eds), *Rationality*, Londres, Routledge, 1993.
- Marcus S. L. et Rips L. J., Conditional reasoning, *Journal of Verbal Learning and Verbal Behavior*, 21, 1979, 199-223.
- Margolis L., *Patterns, thinking and cognition : A theory of judgement*, Chicago, Chicago University Press, 1987.
- Markovits H., Awareness of the « possible » as a mediator of formal thinking in conditional reasoning problems, *British Journal of Psychology*, 75, 1984, 367-376.
- Matalon B., Étude génétique de l'implication, in W. Beth et al. (Eds), *Implication, formalisation et logique naturelle*, *Études d'épistémologie génétique*, vol. 16, Paris, PUF, 1962.
- Matthews L. J. et Patton J. H., Failures to shift following disconfirmation in concept identification, *Journal of Experimental Psychology : Human Learning and Memory*, 104, 1975, 91-94.
- McCarthy J., Circonscription : A form of non-monotonic reasoning, *Artificial Intelligence*, 13, 1980, 27-39.
- McDermott D. et Doyle J., Non-monotonic logic L, *Artificial Intelligence*, 13, 1980, 41-72.
- Medin D. L., Wattenmaker W. D. et Michalski R. S., Constraints and preference in inductive learning : an experimental study of human performance, *Cognitive Science*, 11, 1987, 229-339.

- Mill J. S., *A system of logic*, Londres, Longmans, 1843 (trad. franç. : *Système de logique*, Bruxelles, Mardaga, 1988).
- Mynatt C. R., Doherty M. E., et Dragan W., Information relevance, working memory, and the consideration of alternatives, *The Quarterly Journal of Experimental Psychology*, 46A, 1993, 759-778.
- Newell A., Reasoning, problem-solving and decision processes : The problem space as a fundamental category, in R. Nickerson (Ed.), *Attention and performance*, vol. 8, Hillsdale, NJ, Lawrence Erlbaum, 1981.
- Newell A., *Unified theory of cognition*, Cambridge, Harvard University Press, 1990.
- Newstead S. E., Interpretational errors in syllogistic reasoning, *Journal of Memory and Language*, 28, 1989, 78-91.
- Newstead S. E. et Griggs R. A., Drawing inferences from quantified statements : A study of the square of opposition, *Journal of Verbal Learning and Verbal Behavior*, 22, 1983, 535-546.
- Nisbett R. E., Krantz D. H., Jepson C. et Kunda Z., The use of statistical heuristics in everyday inductive reasoning, *Psychological Review*, 90, 1983, 339-363.
- Oakhill J., Johnson-Laird P. N. et Garnham A., Believability and syllogistic reasoning, *Cognition*, 31, 1989, 117-140.
- Oaksford M. et Chater N., Reasoning theory and bounded rationality, in Manktelow et Over, *op. cit.*, 1993.
- O'Brien D. P., Mental logic and human irrationality : We can put a man on the moon, so why can't we solve those logical-reasoning problems ?, in Manktelow et Over, *op. cit.*, 1993.
- Osherson D. N., Smith E. E. et Shafir E. B., Some origins of belief, *Cognition*, 24, 1986, 197-224.
- Osherson D. N., Smith E. E., Wilkie O., Lopez A. et Shafir E. B., Category-based induction, *Psychological Review*, 97, 1990, 185-200.
- Payne J. W., Bettman J. R. et Johnson E. J., Behavioral decision research : A constructive process perspective, *Annual Review of Psychology*, 43, 1992, 87-131.
- Peirce C. S., *Collected papers*, Cambridge, MA, Harvard University Press, 1931 (3^e éd., 1974).
- Pennington N. et Hastie R., Reasoning in explanation-based decision making, *Cognition*, 49, 1993, 123-163.
- Piaget J., *La psychologie de l'intelligence*, Paris, Armand Colin, 1947.
- Pohlmann M., *Stratégies de raisonnement et structuration des prémisses*, mémoire de maîtrise non publié, Université de Paris VIII, 1995.
- Poitou J.-P., *La dissonance cognitive*, Paris, Armand Colin, 1974.
- Politzer G., Differences in interpretation of implication, *American Journal of Psychology*, 94, 1981, 461-477.
- Politzer G., Laws of language use and formal logic, *Journal of Psycholinguistic Research*, 15, 1986, 47-92.
- Politzer G., Résolution des syllogismes catégoriques, *L'Année psychologique*, 88, 1988, 461-477.
- Politzer G., Immediate deduction between quantified sentences, in K. J. Gilhooly, M. T. G. Keane, R. H. Logie et G. Erdos (Eds), *Lines of thinking. Reflections on the psychology of thought*, Londres, Wiley, 1990.

- Politzer G., *La psychologie du raisonnement : lois de la pragmatique et logique formelle*, thèse de doctorat d'État, Université de Paris VIII, 1993.
- Politzer G. et Braine M. D. S., Response to inconsistent premises cannot count as suppression of valid inferences, *Cognition*, 38, 1991, 103-108.
- Politzer G. et George C., *Reasoning under uncertainty and nonmonotonicity : A pragmatic account*, Communication au XXV^e Congrès international de psychologie, Bruxelles, 19-24 juillet 1992.
- Politzer G. et Nguyen-Xuan A., Reasoning about conditional promises and warnings : Darwinian algorithms, mental models, relevance judgements or pragmatic schemas ?, *Quarterly Journal of Experimental Psychology*, 44, 1992, 401-412.
- Popper K. R., *Logik der Forschung*, Vienne, Julius Springer Verlag, 1935 (trad. angl. : *The logic of scientific discovery*, Londres, Hutchinson, 1959 ; trad. franç. : *La logique de la découverte scientifique*, Paris, Payot, 1973).
- Popper K. R., Degree of confirmation, *The British Journal for the Philosophy of Science*, 5, 1954, 143-149.
- Quine W. V. O., *Methods of logic*, New York, Holt, Rinehart Winston, 1950 (trad. franç. : *Méthodes de logique*, Paris, A. Colin, 1972).
- Reiter R., A logic for default reasoning, *Artificial Intelligence*, 13, 1980, 81-132.
- Richard J.-F., *Les activités mentales*, Paris, Armand Colin, 1990.
- Richard J.-F., Poitrenaud S. et Tijus C., Problem-solving restructuration : Elimination of implicit constraints, *Cognitive Science*, 17, 1993, 497-529.
- Rippol T., La recherche sur le raisonnement par analogie : objectifs, difficultés et solutions, *L'Année psychologique*, 92, 1992, 263-288.
- Rips L. J., Cognitive processes in propositional reasoning, *Psychological Review*, 90, 1983, 38-71.
- Rips L. J., Reasoning, *Annual Review of Psychology*, 41, 1990, 321-353.
- Rips L. J., *The psychology of proof*, Cambridge et Londres, Bradford Book & MIT Press, 1994.
- Rivenc F., *Introduction à la logique*, Paris, Payot, 1989.
- Roberts M. J., Human reasoning : Deduction rules or mental models, or both ?, *The Quarterly Journal of Experimental Psychology* 46A, 1993, 569-589.
- Rumelhart D. E., Towards a microstructural account of human reasoning, in S. Vosnadiou et A. Ortony (Eds), *Similarity and analogical reasoning*, Cambridge, Cambridge University Press, 1989.
- Shao J. et Nguyen-Xuan A., Inférences et résolution de problèmes ; la sériation de poids avec une balance insensible, *Archives de psychologie*, 62, 1994, 107-136.
- Shackle H. et Fischhoff B., Strategies for information search in causal analysis, *Memory and Cognition*, 10, 1982, 520-530.
- Shortliffe E., *Computer-based medical consultation : MYCIN*, New York, Elsevier, 1976.
- Simon H. A., A behavioral model of rational choice, *Quarterly Journal of Economics*, 69, 1955, 99-118.
- Simon H. A., *Models of bounded rationality*, Cambridge, MIT Press, 1982.
- Smith E. E., Shafir E. et Osherson D., Similarity, plausibility, and judgements of probability, *Cognition*, 49, 1993, 67-96.
- Sloman S. A., Feature-based induction, *Cognitive Psychology*, 25, 1993, 231-280.

- Sloman S. A., When explanations compete : The role of explanatory coherence on judgements of likelihood, *Cognition*, 52, 1994, 1-21.
- Smullyan R., *Le livre qui rend fou*, Paris, Dunod, 1984.
- Sperber D. et Wilson D., *Relevance : Communication and cognition*, Oxford, Blackwell, 1986 (trad. franç. : *La pertinence : communication et cognition*, Paris, Éditions de Minuit, 1989).
- Stalnaker R., A theory of conditionals, *Studies in Logical Theory, American Philosophical Quarterly*, monograph n° 2, 1968, 98-112 ; réimprimé in F. Jackson (Ed.), 1991.
- Sternberg R. J., Developmental patterns in the encoding and combination of logical connectives, *Journal of Experimental Child Psychology*, 28, 1979, 469-498.
- Stevenson R. M. et Over D. E., Deduction from uncertain premises, *The Quarterly Journal of Experimental Psychology*, 48A, 1995, 613-643.
- Stich S. P., Could man be an irrational animal ?, *Synthèse*, 64, 1985, 115-135.
- Thagard P., Explanatory coherence, *Behavioral and Brain Sciences*, 12, 1989, 435-467.
- Toulmin S. E., *The uses of argument*, Cambridge, Cambridge University Press, 1958 (trad. franç. : *Les usages de l'argumentation*, Paris, PUF, 1993).
- Tversky A. et Kahneman D., Judgement under uncertainty : Heuristics and biases, *Science*, 185, 1974, 1124-1131.
- Voss J. F., Blais J., Means M. L., Greene T. R. et Ahwesh E., Informal reasoning and subject matter knowledge in the solving of economics problem by naive and by novice individuals, in L. B. Resnich (Ed.), *Knowing, learning and instruction : Essays in honor of Robert Glaser*, Hillsdale, Lawrence Erlbaum, 1989.
- Walliser B., *Anticipations, équilibres et rationalité économique*, Paris, Calmann-Lévy, 1985.
- Walliser B., Instrumental rationality and cognitive rationality, *Theory and decision*, 27, 1989, 7-36.
- Walliser B., Logique épistémique et théorie de jeux, *Revue économique*, 42, 1991, 801-832.
- Weber E. U., Böckenholt U., Hilton D. et Wallace B., Determinants of diagnostic hypotheses generation : Effects of information, base rates, and experience, *Journal of Experimental Psychology : Learning, Memory and Cognition*, 19, 1993, 1151-1164.
- Weil-Barais A., Les apprentissages en sciences physiques, in G. Vergnaud (coord.), *Apprentissages et didactique, où en est-on ?*, Paris, Hachette, 1994.
- White P. A., Causal processing : Origins and development, *Psychological Bulletin*, 104, 1988, 36-52.
- Whitehead A. N., et Russell B., *Principia mathematica*, Cambridge, Cambridge University Press, 1913.
- Wright P. et Hull A. J., Reading to do : Creating contingent action plans, *British Journal of Psychology*, 79, 1988, 187-211.
- Zwirn D. et Zwirn H., Confirmation non probabiliste, in J. Dubucs et F. Lepage, *op. cit.*, 1995.



Imprimé en France
Imprimerie des Presses Universitaires de France
73, avenue Ronsard, 41100 Vendôme
Février 1997 — N° 43 315

COLLECTION
PSYCHOLOGIE ET SCIENCES DE LA PENSÉE

- Dehaene S., *Le cerveau en action. Imagerie cérébrale fonctionnelle en psychologie cognitive.*
- George C., *Polymorphisme du raisonnement humain. Une approche de la flexibilité de l'activité inférentielle.*
- Gineste M.-D., *Analogie et cognition. Étude expérimentale et simulation informatique.*
- Houdé O. et Miéville D., *Pensée logico-mathématique. Nouveaux objets interdisciplinaires.*
- Houdé O., *Rationalité, développement et inhibition. Un nouveau cadre d'analyse.*
- McAdams S. et Bigand E., *Penser les sons. Psychologie cognitive de l'audition.*
- McCarthy R. A. et Warrington E. K., *Neuropsychologie cognitive. Une introduction clinique.*
- Pacherie É., *Naturaliser l'intentionnalité. Essai de philosophie de la psychologie.*
- Pélissier A. et Tête A., *Sciences cognitives. Textes fondateurs (1943-1950).*
- Proust J., *Perception et intermodalité. Approches actuelles de la question de Molyneux.*
- Sprenger-Charolles L. et Casalis S., *Lire. Lecture et écriture: acquisition et troubles du développement.*
- Tisseau G., *Intelligence artificielle. Problèmes et méthodes.*
- Zesiger P., *Écrire. Approches cognitive, neuropsychologique et développementale.*

PRESSES
UNIVERSITAIRES
DE FRANCE

Christian George

Polymorphisme du raisonnement humain

*Une approche de la flexibilité
de l'activité inférentielle*