
PEUT-ON NATURALISER LA CONSCIENCE?

Tiziana Zalla

C. R. E. A. - Ecole Polytechnique
1, Rue Descartes
75005 Paris
zalla@poly.polytechnique.fr

Résumé

On assiste ces derniers temps à des tentatives de plus en plus nombreuses d'élaboration de théories de la conscience dans le cadre des sciences cognitives. Le problème du rôle causal des propriétés phénoménales de la conscience apparaît comme complexe et controversé. Il s'agit de s'interroger sur la nature des relations entre les propriétés computationnelles ou fonctionnelles de l'esprit et les contenus qualitatifs de l'expérience. L'apport d'une approche pluridisciplinaire se révèle utile et nécessaire pour développer une conception naturaliste de l'esprit. Dans cet article, j'entends montrer comment un problème philosophique comme celui des caractéristiques phénoménaux de la conscience est susceptible d'être abordé dans les cadres explicatifs d'autres disciplines, telles que la psychologie expérimentale et la neuropsychologie.

Mots clés : conscience, propriétés phénoménales, propriétés fonctionnelles, spécificité modale, modularité.

1. Les propriétés phénoménales de la conscience : présentation du problème

Le problème de la conscience a émergé au cours de ces dernières années avec l'avancement de la recherche dans les sciences cognitives. Les théories concernant les processus cognitifs - perception, mémoire, apprentissage, langage - formulées dans les dernières années n'ont pas envisagé les problèmes de la conscience alors que sa compréhension est devenue un problème légitime et central dans l'étude de la cognition et du comportement chez l'homme. De toute évidence, le problème de la conscience est né et s'est posé au sein de la philosophie de l'esprit dans que certains pré-supposés de la doctrine béhavioriste ont été remis en question. Le béhaviorisme a en fait traité la conscience comme un sujet sans aucune légitimité scientifique. Le débat actuel est animé par quatre doctrines : le fonctionnalisme, le phénoménisme, le physicalisme réductionniste et l'éliminativisme.

Malgré de nombreuses divergences, les différentes théories fonctionnalistes partagent certains principes philosophiques généraux. Elles considèrent l'esprit comme un système informationnel, non "énergétique", et elles considèrent la psychologie comme l'étude de la vie mentale. En s'accordant avec la doctrine béhavioriste, le fonctionnalisme admet l'existence d'états mentaux internes qui sont des causes des comportements. Selon le fonctionnalisme, un état mental est un état fonctionnel qui est défini par ses relations avec les stimuli, les réponses comportementales et les autres états mentaux. J'appellerai phénoménisme la doctrine selon laquelle il existe des propriétés subjectives et qualitatives de l'expérience consciente qui échappent à toute explication physicaliste et fonctionnaliste. La critique principale adressée au fonctionnalisme porte

sur le fait que les états psychologiques n'y sont pas caractérisés par leur contenu, mais seulement par le rôle causal qu'ils jouent en tant que symboles formels. Ainsi comme, le fonctionnalisme computationnel ignore deux types de propriétés mentales : les propriétés intentionnelles et les propriétés qualitatives de l'expérience. Pour que la conscience puisse être prise en compte par une explication fonctionnaliste, elle doit obéir à un principe de pertinence causale, c'est-à-dire qu'elle doit avoir une efficacité causale sur les autres états mentaux. Les défenseurs des qualia s'interrogent donc sur la validité de la doctrine fonctionnaliste pour l'étude des caractéristiques subjectives et phénoménaux de la conscience. Dans la taxinomie psychologique des états mentaux, il y a en effet des états qui semblent plus difficiles à définir uniquement d'après leur contenu représentationnel ; les expériences perceptives des couleurs, des sons, les sensations, les sentiments semblent plutôt se caractériser par un contenu qualitatif. Ces états purement qualitatifs seraient alors des contre-exemples de la thèse de Brentano (1924-1928) selon laquelle l'intentionnalité est une marque du mental¹. Un des arguments apportés contre le computationnalisme consiste à montrer que, du point de vue du fonctionnalisme, il est concevable qu'une créature ait exactement les mêmes caractérisations fonctionnelles qu'un être humain conscient alors que rien dans cette caractérisation ne révèle si son expérience des couleurs est similaire à celle de l'être humain. L'aspect qualitatif de ses expériences conscientes peut même être inversé : sa vision d'un objet rouge correspondrait à notre expérience qualitative du vert et, à l'inverse, ce qu'elle appelle vert correspondrait chez nous à l'expérience phénoménale du rouge. L'aspect qualitatif de l'expérience peut être

¹ Selon Brentano (1924-1928), l'intentionnalité est le caractère qui différencie les phénomènes psychiques des phénomènes physiques. Elle est définie comme une relation à un objet, un "être dirigé vers" un objet, à un contenu de la représentation.

aussi complètement absent chez cette créature sans que rien dans l'organisation fonctionnelle de son esprit ne révèle les quels systèmes sont intrinsèquement liés à l'expérience. Il semble donc que deux états mentaux sont susceptibles de jouer le même rôle fonctionnel indépendamment du fait d'être ou non accompagnés par une expérience qualitative consciente. Par conséquent, il y a un aspect important du mental, notamment l'expérience consciente, qu'il est impossible de capturer avec les approches computationnelles et fonctionnalistes du mental.

L'avancement des recherches en neurosciences nous amène à aborder l'étude de la conscience d'une manière empirique et pluridisciplinaire. Cependant, une telle recherche empirique sur la conscience doit être précédée, à mon avis, par un travail d'analyse conceptuel. Les qualités phénoménales ne doivent pas être confondues avec les propriétés fonctionnelles, ni avec le contenu sémantique ou intentionnel des représentations. Ned Block (1992; 1995) a remarqué que le terme "conscience" correspond en fait à un faisceau de concepts et que la confusion de ces concepts est à l'origine de la méconnaissance des caractéristiques phénoménales de la conscience. Par "conscience phénoménale" ou "qualia", on entend les propriétés subjectives et qualitatives des sensations, des émotions, des perceptions, et selon certains auteurs, de tous les états mentaux (cf. Block, 1994). Une qualité phénoménale désigne un caractère de l'état conscient qui est inaccessible introspectivement, à savoir comment les expériences sont ressenties et les choses nous apparaissent. Depuis les années soixante, en adoptant le paradigme fonctionnaliste, la psychologie cognitive a surtout donné des explications des processus cognitifs uniquement en termes de mécanismes computationnels inconscients. L'esprit computationnel, tel qu'il est conçu par le fonctionnalisme, et l'esprit phénoménal postulent deux domaines de phénomènes et de propriétés différents du niveau physique. Les deux appartiennent au domaine des phénomènes mentaux même s'ils désignent des propriétés différentes des états mentaux. Selon le fonctionnalisme les états mentaux ont des propriétés représentationnelles ou intentionnelles, alors que les défenseurs de l'esprit phénoménal font référence aux propriétés qualitatives et subjectives de l'expérience. En philosophie de l'esprit on a aussi assisté à la proposition d'une approche éliminativiste ou réductionniste des processus mentaux qui s'accompagne de la négation de la dimension subjective et irréductible de la conscience. Thomas Nagel (1974) en formulant le problème de la subjectivité et du dualisme par la question "Qu'est-ce que cela fait d'être une chauve-souris?" ("What is it like to be a bat?") entame de façon intuitive une orientation de pensée anti-physicaliste. Par physicalisme, Nagel entend la conception générale selon laquelle le monde n'est rien d'autre qu'un système physique énorme et compliqué, au sens que les faits fondamentaux sont physiques et que toute description complètement existante est de nature physique. De la même manière, toutes les choses existantes dépendent ontologiquement et non causalement de ces faits physiques. Or, Nagel fait remarquer que la modalité d'existence des expériences mentales ne semble pas être la même que celle des objets et de ces événements physiques. Les états mentaux conscients diffèrent des autres faits objectifs parce qu'ils sont l'objet non seulement d'une description à la

troisième personne, mais aussi d'une conscience immédiate qui ne se manifeste qu'à la première personne. Nagel refuse les tentatives d'appliquer les idées objectivistes des sciences physiques à l'analyse des phénomènes conscients connus de manière introspective et à la première personne. L'argument de Nagel est que par une description objectiviste on prétend employer les concepts expérimentiels tout en niant les propriétés intrinsèques de l'expérience. Selon sa définition, la conscience est un phénomène de nature subjective et inaccessible à un observateur externe. Considérons le cas d'un état conscient tel que la douleur. Quel fait du monde correspondrait à l'énoncé, "Je ressens une douleur"? Il semble y avoir deux genres de faits. Tout d'abord, le fait que j'ai prouvé quelques sensations conscientes désagréables, et que je fais une expérience de ces sensations d'un point de vue subjectif, à la première personne. Ces sensations sont constitutives de ma douleur présente. Cependant, la douleur est aussi provoquée par certains processus neurophysiologiques sous-jacents constitués en grande partie par des configurations d'activité neuronale siégeant dans le thalamus et d'autres régions du cerveau. Or, si nous essayons de réduire la sensation subjective consciente de la douleur à l'activité neuronale objective en affirmant que la douleur n'est que ce processus neurophysiologique, nous effectuons une réduction ontologique. La conséquence d'une telle réduction est que les caractéristiques essentielles de la douleur sont négligées car aucune description objectiviste à la troisième personne en termes de faits physiologiques ne peut sauvegarder le caractère qualitatif de la douleur et des autres expériences subjectives. La même idée était déjà avancée par Saul Kripke (1972) lorsqu'il affirmait que par le terme "douleur", nous désignons quelque chose de différent d'un état du corps ou du cerveau. Donc, la douleur ne peut pas être identique aux états neurophysiologiques induits par l'activité des neurones du thalamus. Nagel affirme ainsi qu'aucune description physique et fonctionnelle des autres organismes ne sera suffisante pour connaître l'expérience que ces organismes éprouvent. De là on conclut que la nature des propriétés phénoménales ne peut être ni physique, ni simplement fonctionnelle.

Cette manière de concevoir l'existence des qualia, ainsi que leurs relations avec les états physiques et les états fonctionnels du cerveau, est cependant controversée. En particulier, les théories qui associent la conscience aux processus de haut niveau de traitement de l'information soutiennent généralement une conception épiphénoménaliste des qualia. Les qualia sont, plus précisément, des phénomènes de projection des processus ou des états mentaux computationnels et, en conséquence, ils sont causalement inefficaces. Ils sont causés par les états fonctionnels, mais à l'inverse, ils n'ont aucune efficacité causale sur les états fonctionnels. Selon Ray Jackendoff (1987, p. 276) les théories computationnelles du fonctionnement cognitif reposent principalement sur deux hypothèses : 1. le présupposé de la "suffisance computationnelle" selon lequel toute distinction phénoménale est déterminée par ou dérive d'une distinction computationnelle correspondante; 2. le présupposé de la "non efficacité" de la conscience selon lequel la conscience d'une entité E n'a en elle-même aucun effet sur l'esprit computationnel. La nature de la conscience est alors déterminée par la structure computationnelle des processus sous-jacents

de telle manière que toutes les distinctions phénoménales sont produites, déterminées, ou projetées à partir de structures computationnelles distinctes.

La difficulté de rendre compte de l'expérience subjective et des aspects qualitatifs des états mentaux à partir de la connaissance de leurs bases physiques concerne directement les neurosciences. La non-conformité à l'exigence scientifique d'observabilité publique en fait un problème épistémologique. À cet égard, toutes les théories de l'esprit sont concernées, notamment le fonctionnalisme, l'éliminativisme, le physicalisme et les théories réductionnistes. Le sujet de l'expérience occupe une position d'autorité dans l'expression des jugements concernant ses expériences. L'introspection reste donc l'accès privilégié à nos états mentaux conscients - pensées, croyances, intentions - et aux expériences sensorielles et perceptives. L'enjeu pour une théorie de la conscience est alors de tenir compte à la fois de deux perspectives : l'expérience introspective et les descriptions objectives à la troisième personne. La fiabilité de l'introspection comme méthode de recueil de données fait l'objet de controverses (Cf. Nisbett et Wilson, 1977). Pourtant, je crois qu'une approche exhaustive doit tenir compte des comptes rendus verbaux comme source indirecte et comme indication fiable de l'expérience subjective dans des conditions explicitées par Ericsson et Simon (1984)².

Nagel (1974, p.164) reconnaît que le point de vue subjectif n'est pas seulement une perspective individuelle : les membres d'une même espèce animale et doués des mêmes systèmes perceptifs le partagent. Des organismes d'espèces radicalement différentes pourraient comprendre les mêmes événements physiques en termes objectifs sans partager les mêmes formes phénoménales par lesquelles les événements se manifestent par les systèmes perceptifs de chaque espèce. La réduction aux points de vue particuliers des connaissances physiques n'est envisageable qu'à condition que le point de vue spécifique à l'espèce soit omis de la réduction. Nagel envisage ainsi une forme de "phénoménologie objective" qui ne dépende guère de l'empathie ou de l'imagination et qui soit capable de décrire le caractère subjectif des expériences en une forme compréhensible par des organismes qui sont dans l'impossibilité d'avoir ces mêmes expériences (par exemple, un être aveugle à qui manque l'expérience des couleurs). Cela pourrait se présenter sous la forme d'une découverte ou de la création d'un point de vue

commun avantageux à partir duquel on peut voir quels aspects de l'expérience subjective sont accessibles à ce type de description objective et comment l'expérience subjective et la neurophysiologie sont connectées. Mais, conclut Nagel, en l'état présent, aucune théorie physique de l'esprit ne peut rendre compte des aspects subjectifs de l'expérience. Ces aspects qui restent à expliquer constituent encore le mystère de la conscience.

Le défi majeur d'une telle approche véritablement naturaliste consiste à faire face à l'épiphiénoménalisme. Une théorie qui serait à la fois naturaliste et non-réductrice des états de conscience doit démontrer que les états phénoménaux ont un rôle causal, non en vertu de leurs états physiques ou computationnels, mais en raison de propriétés phénoménales spécifiques aux états mentaux. Selon l'hypothèse que je veux défendre, les propriétés qualitatives des expériences jouent un rôle important dans les processus inférentiels permettant de déterminer l'origine des représentations mentales, et dans l'identification des états internes. Plus précisément, les propriétés qualitatives sont des indices à partir desquels des processus d'ordre supérieur peuvent distinguer les souvenirs des perceptions et des connaissances générales, les événements perçus des événements imaginés.

2. La spécificité modale de la conscience phénoménale

Dans le cadre d'une approche pluridisciplinaire, je propose d'aborder le problème des propriétés phénoménales en essayant de démontrer qu'il est possible d'entamer une approche expérimentale de la conscience, et de nier ainsi sa nature en principe irréductible. Contrairement à la notion de conscience computationnelle, la conscience phénoménale n'est pas une notion opérationnelle. Je crois pourtant qu'une solution empirique pourrait s'imposer face aux difficultés de détermination et de naturalisation de ces propriétés phénoménales. Par définition, les propriétés phénoménales sont intrinsèques, c'est-à-dire qu'elles ne dépendent pas de leur relation avec le monde extérieur. Donc, elles ne sont identifiables ni aux propriétés fonctionnelles et représentationnelles qui sont relationnelles, ni aux propriétés neuronales car ces dernières ne respectent pas les conditions d'accessibilité cognitive.

Block (1991; 1994, 1995) a remarqué que la confusion entre le concept de conscience phénoménale et le concept de conscience comme accès à un système de traitement de l'information (ou computationnelle) - qu'il s'agisse d'un système exécutif ou de supervision, attentionnel ou de contrôle - est à l'origine de maintes argumentations fallacieuses. Il soutient qu'il s'agit de deux phénomènes distincts, bien que ces deux types de conscience se manifestent habituellement ensemble. Une théorie computationnelle de la conscience pourrait démontrer que certaines structures constituantes du système cognitif possèdent les conditions fonctionnelles nécessaires à l'expérience consciente. Une telle théorie computationnelle pourrait indiquer les procédures d'accès et de traitement d'information à travers lesquelles certaines représentations atteignent le niveau de conscience; elle

² Ericsson et Simon remarquent que Nisbett et Wilson n'ont pas spécifié les conditions sous lesquelles le sujet serait incapable d'effectuer un compte rendu fiable de ses expériences privées. Les auteurs développent un modèle précisant les conditions nécessaires pour un compte rendu adéquat : l'information reportée est plus fiable si elle est stockée en mémoire de travail que si elle est récupérée de façon rétrospective en mémoire à long terme. Ils soutiennent donc qu'une théorie de l'introspection est viable si elle est limitée aux états mentaux présents et non aux expériences mentales passées, puisque on a un accès direct aux souvenirs et non aux expériences mentales originaires de ces souvenirs. Quand l'accès est limité, le système verbal effectuée des inférences sur l'origine des processus et des états mentaux, ce qui sous des conditions particulières, donne lieu à des confabulations.

pourrait ainsi d crire quels processus fonctionnels sont conscients. En revanche, une telle approche ne cherche pas   expliquer pourquoi une partie seulement des processus cognitifs humains est consciente. Elle ne cherche pas non plus   indiquer les avantages  volutifs de la conscience et comment celle-ci  mergea au cours de l' volution phylog n tique. Toutefois, je consid re ici que le niveau computationnel est suffisamment abstrait pour d crire le degr  de complexit  de l'esprit et du comportement humain, si bien qu'il appara t actuellement comme l'approche la plus ad quate pour fournir une taxonomie des ph nom nes de conscience.

Ainsi que l'a propos  Jackendoff (1987), la notion de "niveaux de repr sentation" se r v le tr s utile dans l'analyse des relations entre les propri t s fonctionnelles et l'exp rience ph nom nale, ainsi que pour une th orie de la conscience fond e sur les notions de modularit  et de modalit  sp cifique, telle que je la propose (cf. Zalla, 1996). Dans son ouvrage *The Modularity of Mind* (1983), Jerry Fodor d fend l'hypoth se selon laquelle il existe une discontinuit  entre perception et cognition. Il d finit un module comme un dispositif de traitement d'information propre   un domaine sp cifique, rapide, automatique, g n tiquement d termin , pr cabi , et computationnellement autonome, par opposition d'une part aux transducteurs et d'autre part aux processus centraux³.

Selon Jackendoff, il y a trois raisons pour lesquelles l'approche computationnelle classique serait insuffisante. Elle n'aborde ni le probl me de l'externalisation des objets de l'exp rience, ni le probl me de la forme d'une exp rience, ni le probl me des propri t s qualitatives des  tats de conscience. Jackendoff propose une version r vis e de cette th orie qui cherche   d couvrir pr cis ment quelles sont les structures d'informations responsables des distinctions qualitatives de l'exp rience, quels sont les niveaux de repr sentation   partir desquels les propri t s ph nom nales sont d termin es. Jackendoff observe tout d'abord que l'exp rience consciente est distincte selon la modalit  sensorielle : l'exp rience visuelle est qualitativement diff rente de l'exp rience auditive ou tactile et chaque modalit  a sa cha ne caract ristique de niveaux de repr sentation. Selon Jackendoff, ce fait sugg re que la sp cificit  modale de l'exp rience est due   des niveaux distincts

de structures d'information, que les interactions des niveaux de structure sont responsables des interactions entre les facult s et enfin que les niveaux centraux conceptuels, largement ind pendants de la modalit  sensorielle, sont le point d'intersection de plusieurs cha nes modales distinctes. Cette hypoth se des niveaux de repr sentation peut fournir une explication de la pluralit  de la conscience, con ue au sens de sp cificit  modale et modulaire.

Cependant, il faut d signer, dans un tel cadre, les niveaux responsables des propri t s qualitatives subjectives. Jackendoff propose une th orie de la forme selon laquelle des niveaux interm diaires de repr sentation pourraient en principe correspondre   des repr sentations conscientes. La conception la plus r pandue consid re la conscience comme correspondant aux niveaux cognitifs centraux,   savoir que les structures conceptuelles sont le lieu privil gi  de la conscience qui fournissent la forme des repr sentations de type conscient. Jackendoff remarque toutefois que dans la production langagi re on est conscient de trois organisations de l'information distinctes et coexistantes : les structures phonologiques, syntaxiques et conceptuelles. Il est clair que la conscience ne peut pas correspondre aux repr sentations primaires, trop pauvres en information. Mais, pour Jackendoff, elle ne correspond pas non plus aux repr sentations centrales qui ne sont pas modalement sp cifi es. Jackendoff propose donc de reformuler la th orie computationnelle classique en postulant des processus attentionnels op rant sur les niveaux interm diaires de repr sentation. Selon cette id e, les processus lents n'ont pas d'effet ph nom nal tandis que le flux de conscience est la projection de structures d'information qui sont cr  es et manipul es par des processus rapides. Les repr sentations centrales sont caract ris es par des processus lents qui ne sont pas modalement sp cifi s. Par contre, les processus rapides op rant sur des structures d'information   des niveaux interm diaires peuvent, eux,  tre   l'origine des propri t s qualitatives subjectives. Le caract re de sp cificit  modale de la conscience ph nom nale est donc la cons quence de la projection de niveaux de structures sp cifiques   chaque modalit  sensorielle. En ce qui concerne la perception visuelle, le niveau de la structure visuelle le plus directement responsable de la forme de la conscience visuelle est ce que David Marr (1982) a appel  l' bauche   2 dimensions 1/2", c'est- dire la repr sentation des surfaces visibles centr e sur l'observateur. Selon Marr les  bauches primaires,   2D 1/2 et   3D restent introspectivement indistinctes. L' bauche primaire ne contient pas assez d'information, ni le caract re unitaire et int gr  propre   la conscience visuelle. Les structures appropri es aux formes de la conscience doivent correspondre aux niveaux o  l'information provenant de deux r tines est int gr e. D'un autre c t , il y a nombre d'arguments contre l'hypoth se plus intuitive selon laquelle la conscience est corr i e   l' bauche   3 dimensions, comme le fait que nos exp riences concernent la surface visible des objets alors que la troisi me dimension est consid r e comme le r sultat d'une inf rence visuelle. Le mod le   3D est trop riche et   ce niveau la modalit  visuelle ne serait pas suffisamment diff renci e des autres modalit s sensorielles. Par contre,   la diff rence de l'exp rience, la m moire de la forme d'un objet peut  tre indiff renci e par rapport   la modalit . En conclusion, Jackendoff soutient que la forme de la

³La th orie modulaire de l'esprit propos e par Fodor (1983) pr voit une typologie   trois composantes fonctionnelles : les transducteurs, les syst mes d'analyse d'inputs et les syst mes centraux. Les transducteurs accomplissent la t che de simples syst mes analogiques qui convertissent de fa on plus ou moins pr cise les stimulations proximales en signaux neuronaux covariants. Les m canismes de transduction sont cens s conserver le contenu informatif des entr es en modifiant seulement le format dans lequel l'information est pr sent e. La fonction des syst mes d'input est de fournir,   partir des sorties des transducteurs, une codification des repr sentations perceptives qui constituent le domaine de op rations des syst mes centraux. Les syst mes centraux qui ont comme entr es les sorties des syst mes d'input (p riph riques), accomplissent des t ches inf rentielles cognitives : ils sont "isotropiques" et "Quiniens" : ce sont des syst mes globaux et holistiques, dont la fonction typique est la fixation des croyances perceptives par des inf rences non-d monstratives.

conscience visuelle - l'apparence des choses - est d'terminée par l'Zbauche ^ 2 D 1/2, alors que la comprZension visuelle - le contenu ou la signification de la conscience visuelle - est d'terminée par l'Zbauche ^ 3D. Plus gZnZralement, pour chaque facultZ, la conscience est d'terminée alors non par les niveaux centraux de la reprZentation, mais par un niveau intermZdiaire entre le niveau pZriphZrique de bas niveau et le niveau central ou conceptuel. Selon l'hypothZse des niveaux intermZdiaires, la conscience linguistique est spZcifiquement causZe, d'terminée ou projetZe par les structures phonologiques; la conscience musicale par les surfaces musicales; la conscience visuelle par les Zbauches ^ 2 dimensions 1/2.

La thZse sur le caractZre modulaire de la conscience trouve appui de la part de la thZorie des niveaux de reprZentation et de la spZcificitZ modale de chaque facultZ sensorielle. Le caractZre unitaire et plurimodal des expZriences dZpend de processus d'intZgration des reprZentations qui se fait au niveau conceptuel du modZle ^ 3 dimensions. Comme les reprZentations sont stockZes en mZmoire avec le mZme modZle tridimensionnel, deux manifestations simultanZes ^ modalitZ diffZrente seront considZrZes comme rZfZrant au mZme objet.

3. Vers une phZnomZnologie expZrimentale

En affirmant que les distinctions phZnomZnologiques sont les projections de structures computationnelles sous-jacentes, Jackendoff rejette l'idZe que les propriZtZs qualitatives ont une efficacitZ causale. Ainsi que nous allons le voir, le caractZre strictement modal de la conscience phZnomZnale m'a amenZ, contrairement ^ la vision ZpiphZnomZnaliste, ^ attribuer un rZle causal aux aspects phZnomZnaux dans le fonctionnement de la cognition. A cet Zgard, je vais mentionner certains travaux effectuZs en psychologie expZrimentale.

Marcia K. Johnson (1988) a entamZ une sZrie d'Ztudes expZrimentales sur le rZle des diffZrences qualitatives dans les reprZentations mentales. Selon ses travaux, les propriZtZs qualitatives des expZriences permettent aux sujets de juger de l'origine des reprZentations mentales, d'identifier ainsi leurs Ztats internes de faZon ^ distinguer les souvenirs des perceptions et des connaissances gZnZrales, les ZvZnements perZus des ZvZnements imaginZs. Sur la base de ces diffZrences, certaines reprZentations plus que d'autres sont accompagnZes par un "sentiment" plus ou moins fiable qu'il s'agit d'ZvZnements rZellement vZcus. C'est la plus grande quantitZ d'information sensorielle et contextuelle qui constituerait la diffZrence entre d'une part les reprZentations perZues et d'autre part les reprZentations imaginZes et les connaissances abstraites. La vZridicitZ de ses propres souvenirs et de ceux d'autrui est jugZe sur la base de la disponibilitZ de ces indices. Dans d'autres expZriences, A.G. Suengas et M. K. Johnson (1988) demandent aux sujets de dZcrire les ZvZnements et d'estimer la qualitZ phZnomZnale des leurs reprZentations. L'intZrZt principal de cette Ztude Ztait d'examiner la nature de la reprZentation en mZmoire en relation avec les processus de jugement et d'attribution. Les auteurs affirment que l'Zvocation et la

verbalisation des souvenirs jouent un rZle critique dans le maintien de la clartZ et dans la qualitZ de ces souvenirs. Les reprZentations d'origine perceptive plus riches en dZtails qualitatifs et en intensitZ des stimuli, restent plus longtemps en mZmoire que les reprZentations imaginZes. L'Zvocation des aspects qualitatifs des expZriences augmente la saillance des caractZristiques sensorielles et contextuelles des souvenirs et des ZvZnements perZus par rapport aux ZvZnements imaginZs ou aux reprZentations abstraites. Quand des ZvZnements purement imaginZs sont largement ZvoquZs, les diffZrences qualitatives entre images mentales et images perceptives s'estompent et les images mentales sont plus aisZment confondues avec les images des ZvZnements perZus. D'une maniZre gZnZrale, la charge Zmotionnelle normalement liZe aux Zpisodes vZcus et les dZtails perceptifs de ces souvenirs engendrent un "sentiment de rZalitZ". Selon les auteurs, l'existence d'un "systZme de contrZle de la rZalitZ" serait ^ la base des fonctions de discrimination et de jugement des diffZrents types de reprZentations mentales. En d'autres termes, des processus d'attribution et de jugement seraient effectuZs grZce ^ des capacitZs mZtareprZentationnelles sur la base de propriZtZs qualitatives associZes aux reprZentations primaires, dZfinies en termes de leur relation sZmantique directe avec le monde (cf. Leslie, 1987 Sperber, 1994). L'hypothZse dZveloppZe par Johnson est qu'un affaiblissement des diffZrences qualitatives dans les reprZentations et des valeurs Zmotionnelles entraZne un fonctionnement dZfectueux du systZme de "contrZle de la rZalitZ". Selon une telle hypothZse, notre capacitZ de jugement dZpendrait en partie des similaritZs et des diffZrences phZnomZnales. A. Baddeley et B. Wilson (1986) ont dZcrit deux patients amnZsiques ayant des difficultZs ^ distinguer leurs expZriences vZcues des ZvZnements imaginZs et ^ dZtecter l'origine perceptive ou imaginaire de leurs reprZentations en mZmoire. L'examen de certains cas d'amnZsie ont mis en Zvidence le fait que les patients amnZsiques manifestent un encodage moins intZgrZ et plus pauvre en dZtails des ZvZnements vZcus par rapport aux sujets normaux. Leurs souvenirs ont une qualitZ dZgradZe, ils se rZvZlent instables, vagues et non spZcifiques. Par consZquent, ces patients sont incapables de composer un rZcit ^ partir de ces souvenirs, de croire ^ ce qu'ils supposent connaZtre et de justifier leur croyances et leurs comportements. Si on examine les cas de patients affectZs du syndrome de la vision aveugle (*blindsight*)⁴, on remarque que ces sujets prZsentent des capacitZs visuelles rZsiduelles en l'absence de toute expZrience phZnomZnale. Par exemple, le patient D. B., ZtudiZ par Larry Weiskrantz (1986) "devinait" l'orientation des lignes et montrait des capacitZs ^ discriminer diffZrentes formes, telles une croix d'un cercle. Lorsque les expZrimentateurs lui demandaient de "deviner", en situation de choix forcZ, la position des stimuli, ce patient avait des performances supZrieures au hasard alors qu'il dZclarait ne pas voir du tout l'objet. Une diffZrence remarquable entre le comportement de ces patients et

⁴ La vision aveugle est provoquZe par une lZsion dans l'aire visuelle primaire V1 qui engendre un scotome, c'est-^-dire l'absence apparente de toute expZrience visuelle consciente correspondant ^ la partie du champ visuel concernZe par le scotome.

celui des sujets normaux est que, sur la base des processus inconscients préservés, les patients *blindsight* ne peuvent commencer de façon spontanée et volontaire aucun comportement visant les objets présents dans la partie aveugle de leur champ visuel (cf. Marcel, 1988). Ils répondent seulement quand ils sont forcés de choisir entre des alternatives données. Comme les sujets amnésiques, ils ne croient pas à ce qu'ils déclarent; ainsi, leurs expériences ne peuvent pas constituer une source de justifications pour leurs croyances.

Ces observations mènent au cœur des propriétés phénoménales des états mentaux et de la distinction entre d'une part les états de croyance concernant nos connaissances abstraites sur le monde et d'autre part nos souvenirs autobiographiques et nos croyances perceptives. À ce propos, le philosophe Sydney Shoemaker (1991) a avancé une hypothèse sur le rôle des propriétés qualitatives de l'expérience. Cet auteur affirme que la capacité de détecter et de distinguer ses propres états internes ainsi que ceux d'autrui n'a une valeur adaptative qu'à condition que l'on puisse communiquer l'information à autrui. Plus exactement, notre capacité à communiquer de manière efficace dépend de notre capacité à véhiculer non seulement ce qu'on entend faire, mais aussi les raisons et les justifications de nos choix et de nos décisions. De la même façon, la valeur adaptative du langage tient en partie à notre capacité de véhiculer l'information concernant les comportements futurs de nos conspécifiques. Il est aussi plausible que les êtres humains aient évolué de manière à être sensibles aux aspects du comportement de leurs conspécifiques qui sont susceptibles de révéler et de prédire leurs actes futurs. La capacité d'accéder à ses propres états internes est le seul procédé permettant de fournir aux autres les indices comportementaux dont ils ont besoin pour parvenir à des prédictions et des explications correctes. Dans l'interaction sociale, l'accès aux raisons de nos croyances et de nos actes est aussi important que l'accès au contenu représentationnel de nos états internes. Ces considérations nous amènent à la question concernant le rôle des qualia dans la connaissance. La richesse de l'information sensorielle et contextuelle de l'expérience, préservée par les croyances perceptives, par les souvenirs autobiographiques et par les images mentales, constitue un domaine inépuisable de connaissances tacites. La disponibilité de l'information concernant la source, les modalités et les circonstances concrètes de l'expérience permet de justifier et de réviser nos croyances et nos actes.

Une telle hypothèse présuppose que la catégorisation sémantique des états mentaux et des expériences s'effectue sur la base des aspects phénoménaux de l'expérience consciente, à savoir que les distinctions qualitatives. Comme l'illustre l'œuvre du psychologue Endel Tulving (1983), d'un point de vue computationnel, un état mental de croyance et un souvenir reflètent le fonctionnement de deux systèmes séparés : un système sémantique pour les connaissances générales et un système épisodique pour les événements personnels; d'un point de vue phénoménal, cette distinction s'exprime par des différences qualitatives dans l'expérience de remémoration. Contrairement à la théorie cognitiviste classique élaborée par Zenon Pylyshyn (1984) suivant laquelle les images mentales ne sont que des

représentations symboliques plus fondamentales, selon notre approche, seules les représentations dans un format pictural sont susceptibles de véhiculer les propriétés phénoménales de l'expérience. Selon nous, la différence de format représentationnel entre ces deux types d'états mentaux, propositionnel pour les croyances concernant les faits généraux et analogique pour les souvenirs personnels et les images, est amplifiée par la marque phénoménale caractérisant ces derniers. La conception des systèmes multiples de connaissance renforce l'hypothèse selon laquelle les aspects qualitatifs pourraient être une source potentielle importante de données permettant la discrimination et l'identification des différents états internes.

4. Conclusion.

En postulant l'existence d'un système distinct d'évaluation, de manipulation et de contrôle des représentations, qui utilise les caractéristiques qualitatives comme source d'information, l'hypothèse de l'efficacité causale des qualia est compatible avec les modèles qui voient la cognition humaine caractérisée principalement en termes de capacités métacognitives opérant sur des connaissances spécifiques de domaine (Cf. Zalla et Palma, 1995). L'intérêt de cette approche consiste en l'ouverture de la recherche expérimentale à l'étude des aspects qualitatifs et subjectifs de l'expérience. L'un des présupposés majeurs du programme de recherche cognitiviste consiste à concevoir la cognition comme une manipulation de symboles basée sur des règles computationnelles. Cette ouverture de la conception classique pourrait remettre en question le présupposé selon lequel le système cognitif interagit uniquement avec la forme physique des symboles et non pas avec leur signification ou avec les états qualitatifs. Dans le débat contemporain, on assiste donc à une tension constante entre l'approche phénoménaliste en première personne et les descriptions et les explications à la troisième personne de la psychologie cognitive et du fonctionnalisme. Le comportement, en tant que source de données pour l'étude de la cognition, a gardé sa place dans le cadre tout à fait nouveau de la recherche contemporaine. Néanmoins, une approche exhaustive de la conscience doit aussi être capable de repérer, dans l'ensemble des faits internes, neurologiques ou mentaux, les facteurs causaux des comportements observés, et d'utiliser les comptes rendus verbaux comme accès indirect à l'expérience subjective. L'objectif est de dégager des concepts explicatifs et de formuler des hypothèses qui expliquent les relations existant entre les caractéristiques phénoménaux des états expérimentiels, les autres états mentaux et le comportement. À travers cette démarche, je crois donc possible de déterminer des rôles fonctionnels pour les propriétés qualitatives de l'expérience dans les processus d'acquisition, d'auto-attribution et d'organisation des connaissances, ainsi que dans certains dysfonctionnements neuropsychologiques.

Face à des programmes de recherche basés sur les sciences "dures", la conscience, réduite à des structures anatomiques ou à des processus neurophysiologiques, est souvent identifiée avec ses bases physiques. Le fonctionnalisme, de son côté, reste ferme dans la conviction que la cognition peut s'accomplir sans conscience car on est loin d'avoir

fourni une th orie explicative compl te et exhaustive qui d montre le r le causal des propri t s ph nom nales. Le fait qu'une large partie de la cognition, y compris des processus cognitifs complexes, se produit de mani re inconsciente, le fait qu'on peut simuler sur ordinateur ces m mes processus, ne prouvent pas que chez l'homme les caract res ph nom naux n'ont aucun r le causal. Selon mon hypoth se, le caract re intentionnel de la repr sentation ne peut pas fournir les informations concernant la cause de la croyance et les contextes  pisodiques d'acquisition d'une croyance. Si cela est correct, les propri t s qualitatives ont un r le fonctionnel dans la production, la r vision, la justification des croyances, en particulier des croyances perceptives. Sur un plan th orique et m thodologique, cela implique que les th ories fonctionnalistes de l'esprit doivent tenir compte des distinctions ph nom nologiques que nous relevons dans l'exp rience consciente. Le fonctionnalisme pr sente n anmoins l'int r t de se situer   un niveau abstrait de description des processus mentaux et de proposer donc d'une approche du mental relativement ind pendante des questions d'impl mentation et de r alisation physique en vertu de laquelle on peut en tirer une alternative   l'approche r ductionniste. C'est donc   partir d'une analyse des limitations qui affectent les th ories fonctionnalistes de l'esprit qu'  notre avis, il faut envisager une configuration possible pour une th orie de la conscience.

Remerciements

Je tiens   remercier flisabeth Pacherie pour ses remarques pertinentes concernant la premi re version de ce texte. Je remercie en outre les relecteurs de la version soumise de ce texte et le public des participants du  me Colloque Jeunes Chercheurs en Sciences Cognitives , en particulier John Stewart.

R f rences bibliographiques

[Baddeley et Wilson 86] Baddeley A. et Wilson B., "Amnesia, autobiographical memory, and confabulation". In D. Rubin (Ed.), *Autobiographical memory*, Cambridge : Cambridge University Press, 1986.

[Block 91] Block N., Evidence against epiphenomenalism, in *Behavioral and Brain Sciences*, 14, 1991, pp.670-672.

[Block 92] Block N., "Begging the question against phenomenal consciousness", in *Behavioral and Brain Sciences*, 15, 1992, pp.183-247.

[Block 94] Block N., "Qualia". In S. Guttenplan (Ed) *A Companion to Philosophy of Mind*, pp. 512-18, Oxford : Basil Blackwell, 1994.

[Block 95] Block N. "On a Confusion about a Function of Consciousness", *Behavioral and Brain Sciences*, n.18, 1995, pp. 227-247.

[Brentano 24-28] Brentano F. *Psychologie vom Empirischem Standpunkt*, 3 vol., Leipzig, F. Meiner Verlag, 1924-28 ; trad. fran . *Psychologie du point de vue empirique*, Paris, Aubier-Montaigne, 1944.

[Erixsson & Simon 93] Ericsson, K.A. et Simon, H.A., *Protocol analysis: Verbal reports as data*. Cambridge, Mass. : The MIT Press, 1894/1993.

[Fodor 83] Fodor J.A., *The Modularity of Mind. En essay on Faculty Psychology*, Cambridge, Mass., and London, England : The MIT Press, 1983 ; trad. fran . *La modularit  de l'esprit*, Paris, Minuit, 1986.

[Kripke 72] Kripke S., *Naming and Necessity*, Oxford : Basil Blackwell, 1972.

[Jackendoff 87] Jackendoff R., *Consciousness and the Computational Mind*, Cambridge, Ma. : The MIT Press, 1987.

[Johnson 88] Johnson M.K., "Discriminating the origin of information". In F. Oltmanns & B. Mahers (Eds), *Delusional beliefs*,. New York : Wiley, pp. 34-65, 1988.

[Leslie 87] Leslie A. M.,  Pretense and Representation : The Origins of  theory of Mind  , in *Psychological Review*, vol. 94, n. 4, 1987, pp. 412-426.

[Marcel 88] Marcel A.J., "Phenomenal experience and functionalism", in Marcel A. J., Bisiach E. (Eds.), *Consciousness in Contemporary Science*, Oxford : Clarendon Press, pp. 121-58, 1988.

[Marr 82] Marr D., *Vision, A Computational Investigation into the Human Representation and Processing of Visual Information*, San Francisco : Freeman, 1982.

[Nagel 74] Nagel T., "What it is like to be a bat", *Philosophical Review*, 83, pp. 435-50, 1974.

[Nisbett & Wilson 77] Nisbett R.E et Wilson T., "Telling More Than We Know: Verbal Reports on Mental Processes", *Psychological Review*, 84, (3), 1977, pp.231-259.

[Pylyshyn 84] Pylyshyn Z., *Computation and Cognition : Toward a Foundation for Cognitive Science*, Cambridge Mass. : The MIT Press, 1984.

[Shoemaker 91] Shoemaker, S., "Qualia and Consciousness", *Mind*, vol. C4, 1991, pp. 507-524.

[Sperber 94] Sperber D., "The modularity of thought and the epidemiology of representation", in Hirschfeld L.A. et Gelman S.A. (Eds), *Mapping the mind. Domain specificity in cognition and culture*. Cambridge : Cambridge University Press, pp. 39-67, 1994.

[Suengas & Johnson 88] Suengas A. G. et Johnson M.K., "Qualitative effects of rehearsal on memories for perceived and imagined complex events.", *Journal of Experimental Psychology : General*, 117, 1988, pp.377-389.

[Tulving 83] Tulving E., *Elements of episodic memory*, Oxford : The Clarendon Press, 1983.

[Weiskrantz 86] Weiskrantz L., *Blindsight. A case study and its implications*, New York: Oxford University Press, 1986.

[Zalla & Palma 95] Zalla T. et Palma A. P., "Feeling of Knowing and Phenomenal Consciousness", in *Behavioral and Brain Sciences*, n.18, 1995, pp. 271-272.

[Zalla 96] Zalla T. , *Le probl me corps-esprit dans les th ories fonctionnalistes et mat rialistes de la conscience. Multiplicit  et modularit  de la conscience*. Th se de Doctorat, en pr paration.