

Daniel LAURIER

## Fonction d'indication et sélection naturelle

### Indicator functions and natural selection

**Abstract :** Fred Dretske has proposed a naturalist analysis of the notion of mental content based on the teleological notion of indicator function, and maintained that indicator functions based on natural selection lack the explanatory force of indicator functions based on individual learning. In this paper, having explained Dretske's basic notions of indication and function, I argue that the two types of indicator functions recognised by Dretske really have the same explanatory value.

Le principe de base du fonctionnalisme, en philosophie de l'esprit, veut que les états mentaux se définissent exclusivement par le fait qu'ils sont des états internes qui occupent un certain rôle fonctionnel, c'est-à-dire, par le fait qu'ils entretiennent telles ou telles relations avec les stimuli, les réponses, et les autres états internes. Cela présente l'avantage de reconnaître que les états mentaux peuvent être des causes internes du comportement, sans pour autant les identifier à des états neuro-physiologiques spécifiques. La marque de commerce du fonctionnalisme, qui est aussi la principale raison de sa popularité, consiste en effet à admettre que le même type d'état mental peut être réalisé par différents types d'états physiques.

Les premiers fonctionnalistes ont eu tendance (1) à assimiler la notion de rôle fonctionnel à celle de rôle causal, c'est-à-dire à supposer que le rôle fonctionnel d'un état interne consiste uniquement dans le fait que cet état interne a telles ou telles relations *causales* avec les autres états internes, les stimuli et les réponses, et (2) à comprendre les notions de stimulus et de réponse en termes proximaux, comme faisant référence à des modifications de certaines terminaisons nerveuses de l'agent. Mais les expériences de pensée de Putnam et Burge ont amené la plupart des fonctionnalistes à reconnaître que le contenu des états mentaux intentionnels dépend au moins en partie des relations que l'agent entretient avec l'environnement physique (ou même social) dans lequel il évolue. Il fallait

donc tenir compte des relations causales entre l'agent et le monde physique.

Il est vite apparu qu'on ne peut pas caractériser le contenu d'une croyance comme étant, par exemple, l'ensemble des situations capables d'en causer la réalisation, sans être obligé d'admettre la conséquence indésirable qu'un état mental ne peut pas être faux. D'un point de vue intuitif, le contenu d'une croyance peut être vu comme le type de situation qui *devrait* être réalisé lorsque cette croyance est réalisée. En ce sens, la notion de contenu comporte un aspect normatif/évaluatif qui est totalement absent de la notion de causalité, mais non pas de la notion traditionnelle de fonction, qui est généralement reconnue comme une notion téléologique. D'où l'idée qu'il pourrait être possible d'expliquer en quoi consiste le contenu d'un état mental en supposant que les états mentaux se définissent moins par les relations causales qu'ils ont ou qu'ils peuvent avoir avec les autres états mentaux ou avec l'environnement, que par les relations causales qu'ils ont comme *fonction* d'avoir avec les autres états mentaux ou avec l'environnement. Il y a loin, évidemment, de la coupe aux lèvres, mais beaucoup pensent apparemment que le fonctionnalisme ne peut espérer rendre compte des propriétés sémantiques des états mentaux qu'à condition de s'appuyer sur une notion téléologique de fonction.

Fred Dretske est l'un de ceux qui ont le plus fait pour promouvoir cette idée, en développant une analyse naturaliste à la fois i) de la notion de contenu mental et ii) de la manière dont le contenu des états mentaux contribue à l'explication des comportements. Comme l'écrit Stich (1990 : 803), l'analyse de la notion de contenu intentionnel proposée par Dretske est à la fois ingénieuse, plausible et d'une simplicité désarmante. Elle repose sur une distinction entre ce qu'un événement ou état de choses indique et ce qu'il a pour fonction (naturelle) d'indiquer, et consiste essentiellement à dire qu'un événement ou état de choses (et en particulier un événement ou un état neurophysiologique) signifie ou représente *ce qu'il a pour fonction (naturelle) d'indiquer*. Après avoir expliqué en quoi consiste exactement la proposition de Dretske, je m'intéresserai principalement, dans les pages qui suivent, à la question de savoir si Dretske a raison de discriminer entre les fonctions d'indication basées sur la sélection naturelle et les fonctions d'indication basées sur l'apprentissage, et de soutenir que les premières n'ont pas la même valeur explicative que les secondes.

## 1. LA NOTION D'INDICATION

Examinons d'abord la notion d'indication. Dretske (1981 : 242 n.10 ; 1986 : 20 ; 1988c : 55-56) laisse clairement entendre qu'il s'agit en quelque sorte d'une version de la notion de signification naturelle que Grice (1957) oppose à celle de signification *non naturelle* qu'il cherche à caractériser. Grice ne définit pas la signification naturelle, mais il met bien en évidence qu'un état de choses ne peut signifier naturellement que p que s'il est vrai que p. Cela vaut aussi pour la notion d'indication chez Dretske. Grice (1957 : 215) prend cependant soin de noter qu'un état de choses peut signifier naturellement quelque chose sans pour autant être un *signe naturel* de cette chose, ce qui laisse entendre que pour lui, un état de choses particulier (par exemple, le fait que Pierre ait en ce moment des rougeurs sur le corps) peut signifier naturellement quelque chose (par exemple, que Pierre a la rougeole) même si le type d'état de choses auquel il appartient n'a pas cette signification naturelle (c'est-à-dire, même si le fait que Pierre ait des taches rouges sur le corps ne signifie pas, en général, qu'il a la rougeole). En d'autres termes, Grice semble être disposé à parler de ce que signifie naturellement un token, et à admettre qu'un token peut signifier naturellement telle ou telle chose même lorsqu'il n'appartient à aucun type ayant cette signification. Ce n'est pas le cas de Dretske, qui insiste (1981 : 192 ; 1990c : 820) sur le fait qu'un token n'indique quoi que ce soit qu'en raison du fait qu'il est d'un type qui indique quelque chose. La relation d'indication est d'abord une relation entre types d'états de choses, et n'est que dérivativement une relation entre tokens. C'est dire qu'il n'y a pas d'indication sans régularité.

L'idée de base<sup>1</sup> est tout simplement que

(1) le fait que r soit G indique (c'est-à-dire, véhicule l'information) que s est F si et seulement si il est nomiquement nécessaire que s soit F, si r est G.

En supposant, par exemple, qu'il y ait une loi en vertu de laquelle un certain type d'éruptions cutanées ne peut être causé que par une certaine

---

<sup>1</sup> Dretske (1986 : 19 ; 1988c : 58-59 ; 1990c : 819) tient explicitement la notion d'indication pour équivalente à celle d'information, qu'il a analysée dans les trois premiers chapitres de son (1981). Dire qu'un état de choses indique que p, c'est tout simplement dire, dans le vocabulaire d'après 1986, qu'il véhicule l'information que p. Dretske renvoie donc systématiquement à son (1981) pour une analyse détaillée de la relation d'indication, bien que certaines remarques, dans les textes plus récents, soient difficiles à concilier avec cette identification. Cette analyse est formulée en termes probabilistes, mais Dretske est apparemment disposé à concéder (1983b : 83) que l'information véhiculée par un signal pourrait aussi être caractérisée de manière satisfaisante sans invoquer la notion de probabilité, directement en termes de régularité ou de dépendance nomique. Je ne discute donc ici que de la formulation non probabiliste.

maladie, alors cette définition permettrait de dire que la présence de ce type d'éruptions cutanées *indique* celle de cette maladie. La définition (1) confère à la notion d'indication un caractère nomique qui en fait une notion intensionnelle, puisqu'il a pour effet qu'un état de choses peut indiquer que s est F sans indiquer que s est H, même dans le cas où tous les F sont H. Un état de choses peut en effet dépendre nomiquement du fait que s soit F sans dépendre nomiquement du fait que s soit H, bien que tous les F soient H, c'est-à-dire, bien qu'il y ait une corrélation (accidentelle) entre être F et être H. Cela mérite d'être souligné, puisqu'à première vue, l'intensionnalité de la notion d'indication constitue un avantage dans la perspective d'une naturalisation de l'intentionnalité<sup>2</sup>.

La caractérisation du paragraphe précédent ne correspond cependant qu'à ce qu'on pourrait appeler l'indication absolue, et ne rend pas compte du fait que, selon Dretske, ce qu'indique un signal est fonction de ce qu'il appelle (1981 : 115) le canal de communication, c'est-à-dire, de l'ensemble des conditions existantes dont dépend le signal et qui ou bien 1) ne génèrent aucune information (pertinente) ou bien 2) ne génèrent que de l'information redondante (*du point de vue du récepteur*<sup>3</sup>). Considérons un voltmètre utilisé pour mesurer la chute de voltage d'une résistance dans un circuit électrique. Une chute de voltage de 7 volts, par exemple, génère un courant dans le circuit interne du voltmètre, lequel produit un champ magnétique qui exerce une force sur un ressort qui actionne une aiguille qui se déplace sur une échelle graduée jusqu'au chiffre "7". Lorsque le voltmètre fonctionne normalement, le fait que son aiguille pointe sur le "7" indique donc une chute de 7 volts dans la résistance à laquelle il est attaché. C'est du moins ainsi qu'il est utilisé. Mais il est clair qu'il n'y a de dépendance nomique entre le signal "7" et une chute de 7 volts que si certaines autres conditions sont satisfaites, comme par exemple que le circuit interne du voltmètre ne soit pas défectueux et que le ressort qui actionne l'aiguille ne soit pas usé, etc. Il y a en fait un nombre indéterminé de configurations d'états internes du voltmètre dans lesquelles le

---

<sup>2</sup> Rappelons qu'en philosophie, l'intentionnalité est la propriété qu'ont certains états mentaux comme la croyance et le désir d'avoir un contenu ou de viser un objet, tandis que l'intensionnalité est la propriété qu'ont certaines constructions linguistiques de ne pas tolérer la substitution d'expressions ayant la même extension. Le lien entre les deux notions tient à ce qu'une construction comme "Pierre croit que la Terre est ronde", qui rapporte un état intentionnel de Pierre, est aussi intensionnelle, puisqu'en substituant "l'arithmétique est incomplète" à "la Terre est ronde", dans ce contexte, on n'obtient pas un énoncé qui a la même valeur de vérité, c'est-à-dire la même extension. Certains auteurs ont d'ailleurs soutenu que l'intensionnalité était une marque de l'intentionnalité.

<sup>3</sup> C'est moi qui souligne.

voltmètre pourrait se trouver sans que le fait que son aiguille pointe sur le “7” n’indique une chute de 7 volts dans la résistance en question. Dans ce cas, ne devrait-on pas dire que le fait que l’aiguille du voltmètre pointe sur le “7” indique (absolument) non pas qu’il y a eu une chute de 7 volts dans telle résistance, mais que ou bien 1) il y a eu une chute de 7 volts dans telle résistance et le voltmètre était attaché à cette résistance, et le circuit interne du voltmètre était dans tel ou tel état, et le ressort qui actionne l’aiguille était dans tel ou tel état, ou bien 2) il y a eu une chute de 10 volts dans telle résistance, et le voltmètre était attaché à cette résistance, et le circuit interne du voltmètre était dans tel ou tel état, et le ressort qui actionne l’aiguille était dans tel ou tel état, etc.? Un moment de réflexion suffit pour reconnaître qu’un signal dépend pratiquement toujours d’une vaste disjonction de conjonctions de conditions (c’est-à-dire, de la disjonction de tous les complexes d’états de choses qui pourraient être nomiquement suffisants pour l’existence de ce signal).

S’il n’est pas nomiquement nécessaire qu’il y ait une chute de 7 volts si l’aiguille du voltmètre pointe sur le “7”, car il y a un nombre indéterminé d’autres situations possibles dans lesquelles cette aiguille pourrait pointer sur le “7”, on devrait en conclure que le fait que l’aiguille pointe sur le “7” n’indique pas une chute de 7 volts. Mais Dretske (1981) refuse cette conclusion, en prétextant que ces autres situations possibles ne sont normalement pas des possibilités authentiques, ou pertinentes. Il soutient en d’autres termes qu’un signal ne véhicule pas d’information concernant toutes les circonstances dont il dépend, et en particulier, qu’il ne transmet pas d’information concernant les conditions du canal de communication. Dans l’exemple du voltmètre, le fait que le voltmètre soit relié à telle résistance, le fait que le ressort qui actionne l’aiguille soit en bon état, etc., font partie de ces circonstances stables qu’il est légitime de prendre pour acquises et à propos desquelles la position de l’aiguille ne transmet aucune information (c’est-à-dire, aucune information nouvelle).

Il apparaît donc que la notion d’indication serait mieux caractérisée en disant que

(2) le fait que  $r$  soit  $G$  indique (véhicule l’information) que  $s$  est  $F$  relativement à la condition  $K$  si et seulement si il est nomiquement nécessaire que  $s$  soit  $F$  si ( $r$  est  $G$  et  $K$ ) et il n’est pas nomiquement nécessaire que  $s$  soit  $F$  si  $K$  (où  $K$  est n’importe quelle proposition décrivant un certain type de situation, c’est-à-dire, le “canal de communication”).

Pareille relativisation<sup>4</sup> de la notion d'indication ne menace aucunement le caractère naturaliste du programme de Dretske, bien qu'elle la rende sans doute plus difficile à manipuler. Elle ne fait pas non plus disparaître toute notion d'indication absolue, puisqu'on peut poser que

(3) le fait que r soit G indique (absolument) que s est F si et seulement si il l'indique relativement à la condition tautologique T.

À partir de 1986, Dretske a commencé à suggérer (1986 : 19 ; 1988c : 57) que l'existence d'une simple régularité non nomique peut être suffisante pour parler d'un lien d'indication. Ainsi, lorsque la sonnerie de la porte d'entrée d'une certaine résidence se fait entendre, cela indique, selon lui, qu'il y a quelqu'un à la porte d'entrée. La sonnerie indique la présence de quelqu'un, en vertu du fait qu'elle ne retentirait pas s'il n'y avait personne à l'entrée, et en dépit du fait qu'il n'y a aucune loi naturelle selon laquelle la sonnerie retentit seulement s'il y a quelqu'un à l'entrée. Cela semble, à première vue, être l'indice d'un changement significatif par

---

<sup>4</sup> Dretske (1981 : 119-120) admet explicitement que la distinction entre la source et le canal de communication, c'est-à-dire entre les conditions dont dépend le signal et à propos desquelles il transmet de l'information et celles dont il dépend et à propos desquelles il ne transmet pas d'information, est fonction de l'information dont dispose le récepteur. Dire que le fait que l'aiguille du voltmètre pointe sur le "7" indique une chute de 7 volts relativement à une certaine condition K revient en effet à dire qu'il indique une chute de 7 volts pour un récepteur qui possède l'information que K. Parler du canal de communication n'est ainsi qu'une autre manière de parler de l'information préalable du récepteur, et dire que ce qu'indique un signal est relatif à une condition d'arrière-fond équivaut en définitive à dire que l'information véhiculée par un signal est relative à l'information préalable dont dispose le récepteur.

Rien ne s'oppose à ce qu'on dise que l'information qu'un signal communique à un certain récepteur (ou l'information qu'il véhicule *pour* un certain récepteur) est relative à l'*information* préalablement présente dans le récepteur (et à laquelle il peut ou non avoir un accès cognitif) ; à condition toutefois de pouvoir faire reposer cette notion relative sur une notion absolue qui puisse servir de base inductive (faute de quoi on obtiendrait une définition circulaire). La notion générale d'information pourrait donc aussi être caractérisée de la manière suivante :

(2') le fait que r soit G véhicule l'information que s est F pour un récepteur qui dispose de l'information que K si et seulement si il est nomiquement nécessaire que s soit F si (r est G et K), et il n'est pas nomiquement nécessaire que s soit F si K.

Il est bon de noter que la notion de "disposer de l'information que K" n'a rien d'intentionnel et n'introduit aucune circularité puisqu'elle peut être inductivement éliminée ; un récepteur dispose de l'information que K, dans le sens pertinent, si et seulement si il est dans un état qui véhicule cette information pour un récepteur qui n'en dispose pas déjà, mais qui autrement dispose de la même information que lui. On obtiendrait une caractérisation de l'information absolue en posant qu'il s'agit de l'information (relative) que le signal véhicule pour un récepteur qui ne dispose d'aucune information préalable (ou dont la seule information est tautologique).

rapport à la position défendue en 1981. Mais quand on y regarde de plus près, on se rend compte que si on peut dire que la sonnerie ne retentirait pas s'il n'y avait personne à l'entrée, c'est seulement en prenant pour acquis que les autres possibilités sont exclues (par exemple, qu'il n'y a pas de court-circuit, que les écureuils qui fréquentent le quartier ne rôdent pas près de la porte, etc.), c'est-à-dire, en faisant implicitement référence à certaines conditions d'arrière-fond. Cet exemple n'est donc pas fondamentalement différent de celui du voltmètre, et aucun d'eux n'oblige à renoncer à l'idée que la relation d'indication repose sur une relation de dépendance nomique. Car si ce n'est pas une loi qu'il y a quelqu'un à la porte si la sonnerie retentit, c'en est vraisemblablement une qu'il y a quelqu'un à la porte si la sonnerie retentit, *et* qu'il n'y a pas de court-circuit, qu'il n'y a pas d'écureuil à la porte, etc. Autrement dit, relativement à l'information préalable dont nous disposons, il est nomiquement nécessaire qu'il y ait quelqu'un à la porte si la sonnerie retentit. Il suffit ainsi d'admettre que la relation de dépendance en question peut être relative, ou que les lois pertinentes peuvent être des lois *ceteris paribus*.

## 2. LA NOTION DE FONCTION D'INDICATION

Selon l'analyse de Dretske, un événement ou un état de choses (type) signifie ou représente (le fait) que p si et seulement si il a *pour fonction* d'indiquer que p, de sorte que le contenu d'une représentation n'est rien d'autre que ce qu'elle a pour fonction d'indiquer. Curieusement, Dretske ne s'appuie, pour défendre cette conception, sur aucune analyse explicite de la notion de fonction, encore que ses illustrations dépendent assez clairement d'une notion étimologique de fonction, qui reste cependant à préciser. Dans sa réponse à Millikan (Dretske 1990c : 824), il reconnaît avoir soigneusement évité d'endosser une analyse théorique explicite de la notion de fonction, mais nie avoir besoin d'une telle analyse. Il lui importe seulement, écrit-il, de pouvoir montrer qu'un indicateur peut acquérir, par un processus naturel, la fonction d'indiquer quelque chose, en un sens du mot "fonction" qui permette de dire que si une chose a pour fonction de faire telle ou telle chose alors elle est *supposée* faire cette chose. Mais il est difficile de voir comment on peut espérer montrer cela sans s'appuyer, ne serait-ce que tacitement, sur une certaine compréhension du concept de fonction.

On s'accorde généralement pour dire qu'une analyse adéquate de la notion de contenu doit (au moins) rendre compte de l'erreur et de l'intensionnalité, c'est-à-dire du fait qu'une représentation peut être fautive et du fait que deux représentations peuvent être nécessairement équivalentes sans pour autant avoir le même contenu. Si on n'utilise qu'une

notion intuitive de fonction, il est assez facile de voir que la position de Dretske satisfait le second *desideratum*, car il est intuitivement clair qu'un état de choses peut avoir pour fonction d'indiquer que p sans avoir pour fonction d'indiquer que q, bien qu'il soit nécessaire que p si et seulement si q (et par conséquent impossible d'*indiquer* que p sans indiquer que q).

La notion intuitive de fonction se caractérise notamment par le fait qu'une chose peut, en principe, avoir une fonction sans la remplir, et même, sans être capable de la remplir. Appliqué au cas de la représentation, ce principe permet de conclure qu'une représentation peut avoir pour fonction d'indiquer que p sans pour autant indiquer que p, ou être capable d'indiquer que p. Mais cela ne suffit pas pour assurer qu'une représentation peut être fautive (et satisfaire ainsi le premier *desideratum*), c'est-à-dire, pour assurer qu'une représentation peut avoir pour fonction d'indiquer que p alors qu'il n'est pas le cas que p ; car du fait qu'un état de choses *n'indique pas* que p, on ne peut conclure qu'il n'est pas le cas que p. Supposons qu'un certain état interne E d'un organisme a pour fonction d'indiquer que p, que cet organisme est dans l'état E au temps t, et que ce token de E (appelons le E<sub>t</sub>) n'indique pas que p, bien qu'il soit le cas que p<sup>5</sup>. Comment faudrait-il décrire pareille situation? Puisqu'il est le cas que p, et puisque l'organisme considéré a, par hypothèse, une représentation du fait que p (puisque'il est dans un état qui a pour fonction d'indiquer que p), il ne semble pas y avoir d'autre choix que de dire que cette représentation est vraie. Ceci montre que le fait qu'une représentation ne remplisse pas sa fonction n'est pas (sous l'analyse de Dretske) une condition suffisante pour que cette représentation soit erronée ; ou en d'autres termes, qu'une représentation fautive est plus qu'une représentation qui ne remplit pas sa fonction (bien qu'elle soit au moins cela). Ces remarques suffisent à montrer que l'hypothèse selon laquelle un état de choses représente ce qu'il a pour fonction d'indiquer n'est pas *a priori* dépourvue de plausibilité et mérite d'être examinée en détail.

Mais il y a apparemment une question préalable à élucider. La proposition de Dretske, je viens de le rappeler, est qu'une représentation est un état de choses qui a pour fonction d'indiquer quelque chose. Mais puisqu'un type d'état n'indique telle ou telle chose que relativement à une condition d'arrière-fond déterminée, de quelle manière faut-il comprendre l'idée qu'un état de choses aurait pour fonction d'indiquer *tout court* telle ou telle chose (car c'est bien de cette manière que Dretske s'exprime)? Il

---

<sup>5</sup> Je reviendrai dans un moment sur le fait, que je note ici au passage, que cela n'est possible que si E (en tant que type) n'indique pas que p.

faut avouer que Dretske ne donne guère d'éléments permettant de répondre à cette question.

On peut toutefois noter que, s'il est vrai qu'un item qui a une fonction doit pouvoir ne pas la remplir, alors un signal (type) qui indique que p relativement à la condition K ne peut acquérir la fonction d'indiquer que p *relativement à la même condition K*, bien que Dretske dise régulièrement que ce sont des indicateurs qui acquièrent des fonctions d'indication, ce qui (ne peut manquer d'être vrai, puisque tout état de choses est un indicateur, et) pourrait laisser croire qu'un tel signal devrait acquérir *cette* fonction. Cela découle immédiatement du fait qu'un type de signal indique que p relativement à une condition K si et seulement si tous les tokens de ce type (en tant que tokens de ce type) indiquent que p relativement à la condition K. Dans les termes de Dretske, cela revient à dire qu'un signal (type) qui indique que p relativement à la condition K ne peut acquérir la fonction d'indiquer que p *relativement à la condition K* (car alors il ne pourrait pas ne pas remplir sa fonction), mais doit acquérir la fonction d'indiquer que p relativement à une condition *plus générale* (incluant la condition K). De la même manière (et pour la même raison), un type de signal qui indique *absolument* que p ne peut acquérir ni la fonction d'indiquer absolument que p, ni celle d'indiquer que p relativement à quelque condition que ce soit. On ne passe ainsi de l'indication à une fonction d'indication qu'en vertu d'une sorte de généralisation inductive. Au moins à titre provisoire, on ne risque apparemment pas grand chose à supposer que dire qu'un état de choses a pour fonction d'indiquer que p revient à dire qu'il a pour fonction d'indiquer que p relativement à toute condition K\* qui ne nécessite pas nomiquement que p.

Comme le remarque Millikan (1990a : 128), en supposant qu'une chose peut avoir pour fonction d'*indiquer* quelque chose, Dretske s'éloigne de l'usage qui veut que la fonction d'une chose soit normalement de *produire tels ou tels types d'effets*, et non d'avoir telle ou telle propriété ou d'être dans tel ou tel état. Car avoir pour fonction d'indiquer que p c'est avoir pour fonction de (ou être supposé) nécessiter nomiquement le fait que p. Il n'est pas clair, à première vue, qu'il y ait beaucoup de sens à dire qu'une chose est supposée en nécessiter nomiquement une autre, sauf si le fait de nécessiter nomiquement cette chose peut être tenu pour responsable de la production de certains effets. Mais il semble possible d'interpréter les analyses de Dretske de manière à ce que cette condition soit satisfaite.

Dretske admet deux types de processus naturels capables de conférer la fonction d'indiquer quelque chose à un indicateur, c'est-à-dire de transformer un indicateur en une *représentation* (en quelque chose qui a un contenu, ou une signification). Il s'agit de la sélection naturelle et de

l'apprentissage. L'approche est étiologique, ou historico-causale, dans les deux cas, bien qu'il y ait selon Dretske des différences significatives entre les représentations fondées sur la sélection naturelle et les représentations fondées sur l'apprentissage. De plus, il est loin d'être clair que Dretske opère avec une notion univoque de fonction. Dans le cas des représentations fondées sur l'apprentissage, l'idée de base est qu'un type d'état n'a la fonction d'indiquer telle ou telle chose que si le fait que les tokens de ce type ont indiqué cette chose explique causalement pourquoi ce type d'état a tels ou tels (ou certains) pouvoirs causaux. Mais la situation est plus trouble dans le cas des représentations fondées sur la sélection naturelle, en raison du fait que Dretske veut soutenir que ces représentations ne peuvent pas être de véritables *croyances*, parce qu'elles ne jouent pas de rôle dans l'explication des comportements.

Il importe, pour apprécier correctement la situation, de résumer brièvement la problématique de Dretske (1988c), quitte à revenir plus loin sur certains points. Le principal objectif de Dretske, dans ce livre, est de montrer que les états intentionnels ont une efficacité causale, qu'ils l'ont en vertu de leur caractère intentionnel, et non pas simplement en vertu de propriétés qui ne sont pas intrinsèquement mentales, et que leur rôle explicatif se distingue donc de celui que jouent les états physiques dans l'explication causale des mouvements corporels.

La question de l'efficacité du contenu se présente principalement, chez Dretske, sous la forme de ce que Kim (1991 : 55) appelle le problème du "syntacticalisme" et Horgan (1991 : 73-75), le problème du soprano. Il est clair que le fait que les sons émis par un soprano aient une signification ne contribue d'aucune manière à expliquer pourquoi ces sons ont fait éclater la coupe de cristal de la comtesse. Seules les propriétés physiques de ces sons sont pertinentes. Or on pourrait être amené à soutenir que c'est là une situation générale, et que le fait qu'une chose ait un contenu (ou le fait qu'elle ait tel ou tel contenu) ne peut tout simplement jamais contribuer à expliquer ou à causer quoi que ce soit, ce qui entre évidemment en conflit avec la conviction du sens commun selon laquelle le fait que nos états intentionnels aient un contenu (et le fait qu'ils aient tel ou tel contenu) joue un rôle essentiel dans l'explication de nos comportements. Quand je dis que Pierre a plongé dans la piscine parce qu'il croyait qu'elle était pleine, je n'entends pas dire simplement qu'il était dans un état interne ayant comme contenu que la piscine est pleine, et que cet état a causé (en partie) son plongeon ; je veux aussi dire que cet état a causé son plongeon *parce qu'il avait comme contenu que la piscine était pleine*. Il serait manifestement difficile de renoncer à cette conviction sans apporter de l'eau au moulin de l'éliminonisme.

Une théorie satisfaisante du contenu doit donc faire en sorte que le contenu qu'elle assigne aux états intentionnels puisse jouer un rôle causal/explicatif dans la production des comportements. Or il n'est pas clair qu'une théorie externaliste du contenu, comme celle de Dretske, puisse satisfaire cette contrainte. Une théorie externaliste du contenu implique en effet que le contenu d'un état intentionnel n'est pas déterminé par ses propriétés intrinsèques (non relationnelles), mais dépend au contraire (en partie) de ses relations avec l'environnement extérieur. Dans la perspective de Dretske, un type d'état a tel ou tel contenu (en partie) en raison du fait que dans le passé, des tokens de ce type ont indiqué tel ou tel état de choses (qui est généralement un état de choses extérieur). Il s'ensuit que deux individus qui auraient exactement les mêmes propriétés intrinsèques pourraient être dans des états qui n'ont pas le même contenu. Mais comme il est plausible de supposer que si deux individus ont les mêmes propriétés intrinsèques alors ils produisent les mêmes comportements, cela veut dire que le contenu de leurs états intentionnels ne peut jouer aucun rôle essentiel dans l'explication de leurs comportements. Selon Dretske, son analyse fournit une solution à ce problème, et permet justement de comprendre comment les propriétés sémantiques des états intentionnels contribuent à expliquer les comportements.

Dretske (1988c) explique et défend dans les deux premiers chapitres une conception du comportement selon laquelle un comportement consiste dans le fait qu'un état ou événement interne d'un système cause un mouvement de ce système ou d'une de ses parties, ou quelque autre changement extérieur. Les comportements sont ainsi identifiés à des *processus* causaux dont le terme initial se situe à l'intérieur du système plutôt qu'à des événements, ce qui veut dire que le comportement se distingue à la fois de la cause interne du mouvement corporel et du mouvement corporel lui-même, qui en sont tous les deux des parties. Cette conception confère donc à la notion de comportement une extension extrêmement large, qui la rend applicable à la fois aux plantes, aux invertébrés, aux machines et aux humains. Une action proprement dite n'est en effet qu'un comportement d'une espèce particulière, un comportement où l'état interne qui cause le mouvement est un état intentionnel. La notion de comportement proposée par Dretske est à vrai dire tellement large qu'elle est menacée de vacuité, puisque comme il vient près de le concéder lui-même, *la* cause initiale ou principale du mouvement d'un système peut pratiquement toujours être située indifféremment à l'intérieur ou à l'extérieur du système, selon les intérêts de l'observateur. Cette stratégie permet cependant de mettre en évidence

des analogies intéressantes entre l'explication des actions et l'explication d'autres types de comportements "plus primitifs".

Si le comportement est un processus, une explication causale d'un comportement devra identifier une cause de ce processus et non pas une cause de son résultat, c'est-à-dire (normalement) du mouvement corporel. C'est pour cette raison qu'identifier les causes internes (par exemple, neurophysiologiques) du mouvement corporel ne revient pas à donner une explication causale du comportement lui-même, mais seulement de son résultat. Une explication causale d'un processus consiste à en identifier soit une cause déclenchante, soit une cause structurante. Une cause déclenchante de la production du mouvement M par l'état interne C, c'est-à-dire du comportement consistant dans le fait que C cause M, n'est rien d'autre qu'une cause de C (et donc indirectement de M) qui est extérieure au système considéré. Par exemple, la chute de la température est une cause déclenchante du fait que le thermostat allume la fournaise, c'est-à-dire du fait qu'un certain état interne du thermostat cause l'allumage de la fournaise. Citer la chute de la température permet d'expliquer pourquoi le thermostat a allumé la fournaise à tel moment plutôt qu'à tel autre. Mais il arrive qu'on veuille savoir pourquoi tel état interne du thermostat a causé l'allumage de la fournaise plutôt que, par exemple, l'ouverture de la porte du garage. On cherche alors une cause structurante du comportement, une cause du fait que tel processus se produise plutôt que tel ou tel autre, c'est-à-dire, une cause du fait que C cause M.

Dretske introduit (1988c : 39) la notion de cause structurante en faisant référence à la multiplicité des facteurs causaux, ou conditions d'arrière-fond, qui sont invariablement impliqués dans tout processus causal. Un état ou événement n'en cause un autre que lorsque certaines conditions sont réalisées ; par exemple, l'état C ne déclenche l'allumage de la fournaise que lorsqu'il est adéquatement relié à l'interrupteur de la fournaise et aux autres composantes du thermostat. Ces conditions d'arrière-fond, ou *leurs causes*, sont ce que Dretske appelle des causes (structurantes) de la production de M par C. Mais ces conditions peuvent aussi être (et sont couramment) décrites comme des causes partielles de M lui-même, et lorsqu'on ajoute le fait que l'opposition entre la cause principale et les facteurs causaux d'arrière-fond est d'ordre purement pragmatique, la distinction entre expliquer un processus et expliquer un événement et la distinction entre une cause déclenchante et une cause structurante paraissent bien minces. En d'autres termes, la cause structurante d'un processus peut aussi, comme Dretske (1991a : 197 ; 1993 : 124, 126 note 7) le concédera plus tard, être décrite comme une cause (d'un certain type)

du résultat de ce processus<sup>6</sup>. Mais Dretske préfère le plus souvent parler de la cause structurante comme de la cause d'un processus, et comme sa théorie du contenu est formulée dans ces termes, je m'en tiendrai à cette pratique.

Le problème de l'efficacité causale des croyances devient ainsi (1988c : 84) celui de montrer que le fait qu'un état interne C ait pour fonction d'indiquer quelque chose peut être une cause structurante du fait que C cause le mouvement corporel M. Dretske suggère qu'une croyance est un type d'état (interne) qui acquiert la fonction d'indiquer quelque chose lorsque le fait qu'il indique cette chose (ou plus exactement, le fait que les tokens passés de ce type ont indiqué cette chose) explique causalement le fait qu'il soit devenu une cause (partielle) de certains mouvements corporels<sup>7</sup>, pour soutenir ensuite que ce point de vue exclut les représentations dont la fonction d'indication est le fruit de la sélection naturelle de la catégorie des croyances. Selon lui, seules les représentations dont la fonction d'indication dérive d'un processus d'apprentissage individuel sont telles que le fait qu'elles indiquent quelque chose explique pourquoi elles causent certains mouvements corporels. Comme il admet explicitement que la sélection naturelle peut conférer des fonctions d'indication, il faut apparemment en conclure qu'il utilise alors le terme "fonction" d'une autre manière (et en particulier de telle manière qu'un type d'état interne puisse avoir une fonction d'indication sans pour autant contribuer à l'explication causale des mouvements corporels, ou sans que le fait qu'il ait cette fonction contribue à l'explication causale des mouvements corporels). Mais il est difficile de savoir de quelle manière il l'utilise exactement.

Cette attitude revient à admettre qu'il y a deux types de contenus intentionnels : ceux qui ont une efficacité causale, et ceux qui n'en ont pas et sont par conséquent épiphénoménaux. Mais si les fonctions d'indication basées sur la sélection naturelle n'ont pas d'efficacité causale, pourquoi devrait-on en admettre l'existence ? Si Dretske tient tellement à montrer que les contenus de croyance ont une efficacité causale, c'est qu'il veut

---

<sup>6</sup> Il s'ensuit que l'analyse "componentielle" du comportement défendue par Dretske, et en vertu de laquelle le comportement est identifié à un processus, a au plus une valeur heuristique, et ne joue aucun rôle essentiel dans sa théorie, contrairement à ce qu'il laisse parfois (1990c : 831 ; 1991a : 216 ; 1991b : 117) entendre.

<sup>7</sup> La question de savoir si, et en quoi, cette suggestion confère réellement une valeur explicative au contenu intentionnel sera discutée plus loin. Je note seulement ici qu'elle consiste apparemment à dire que le fait que C indique F explique à la fois pourquoi C cause M et pourquoi C représente (a pour fonction d'indiquer) F et que cela ne semble pas permettre de conclure que le fait que C *représente* F explique pourquoi C cause M (sans parler de la question de savoir si cela l'explique *causalement*).

réfuter l'éliminationnisme des croyances ; mais il ne semble pas réaliser qu'une autre forme d'éliminationnisme pourrait s'appuyer sur le fait que les fonctions d'indication basées sur la sélection naturelle n'ont pas, selon lui, d'efficacité causale, pour en contester la réalité (ou du moins pour lui contester le droit d'en admettre la réalité).

### 3. LA VALEUR EXPLICATIVE DE LA SELECTION NATURELLE

Voyons de plus près comment Dretske décrit les fonctions d'indication fondées sur la sélection naturelle. Le système auditif d'une certaine variété de papillon de nuit (*noctuid moth*), écrit Dretske, (1988c : 91-94) est "manifestement conçu" en fonction de son principal prédateur, la chauve-souris. Il y a un grand nombre de stimuli acoustiques normalement audibles aux autres organismes auxquels l'oreille de ce papillon ne réagit pas. Elle est cependant capable de détecter les "bouffées" d'ultrasons émises par les chauves-souris. Le comportement du papillon est alors commandé par les stimulations de son appareil auditif : il s'éloigne de la source des ultrasons lorsqu'ils sont de faible intensité, et se met à tourner sur lui-même lorsque celle-ci se rapproche (c'est-à-dire, lorsque l'intensité des ultrasons augmente). Le système nerveux du papillon est ainsi structuré de telle manière que les neurones moteurs qui contrôlent la direction du mouvement sont eux-mêmes contrôlés par des structures qui indiquent la localisation de la source sonore. En d'autres termes, il est structuré de telle manière qu'un certain type d'état (C) de l'appareil auditif qui indique la localisation d'une source d'ultrasons (F) cause un certain type de mouvement corporel (M). Les remarques de Dretske concernant ce type de cas ne vont pas toutes dans le même sens. Il écrit :

What the theory of evolution has to tell us about these cases [...] is that C's production of M is, at least in part, *the result of its indication of F* [c'est moi qui souligne]. M is produced by an indicator of F because such an arrangement confers a competitive advantage on its possessor. (1988c : 92)

mais aussi :

In such cases an internal state, C, which means (indicates) that a hungry bat is approaching and which even (let us say) has the *function* of indicating this (in virtue, let us suppose, of its evolutionary development in this *kind* of moth), *does*, to be sure, cause orientation and wing movements of an appropriate (evasive) sort. C (something that indicates the approach of a bat) causes M (bat-avoidance movements). Nevertheless it is not C's meaning what it does (F) that explains why it causes this (M). In this case the internal state *has* a semantics [...] but the fact that it indicates this, x or is supposed to indicate

this, is irrelevant to an understanding x of why it actually does what it does.  
(1988c : 95)

En d'autres termes, le fait que C cause M devrait être le "résultat" du fait que C indique F, sans pour autant être expliqué par ce fait. À moins de sérieuse équivoque, l'une des deux affirmations citées est manifestement de trop, et il y a fort à parier, vu l'ensemble des commentaires de Dretske sur la question, que ce doit être la première. Mais cela soulève la question de savoir si une fonction d'indication peut résulter de la sélection naturelle sans qu'une affirmation de ce genre ne soit vraie.

Selon l'analyse sélectionniste de Neander (1991a, 1991b), un type d'item X a pour fonction de produire Y si et seulement si le fait que des tokens du type X ont produit Y explique causalement pourquoi le type X a été sélectionné. Cela veut dire qu'un type d'état C a comme fonction d'indiquer que F si et seulement si le fait que des tokens de C ont indiqué F explique causalement pourquoi le type C a été sélectionné. Un item ne peut cependant être sélectionné pour indiquer quelque chose que si le fait d'indiquer cette chose procure un avantage adaptatif à ceux qui le possèdent, et le fait d'indiquer quelque chose ne peut conférer quelque avantage que ce soit que s'il produit des effets. Par conséquent, le fait que les tokens d'un certain type aient indiqué quelque chose ne peut expliquer pourquoi ce type a été sélectionné que s'il explique aussi pourquoi les tokens de ce type ont produit les effets qui les ont rendus avantageux (et pourquoi les effets qu'ils ont produits ont été avantageux). Il s'ensuit apparemment que si un type d'item a pour fonction d'indiquer quelque chose dans le sens sélectionniste de Neander, il doit aussi être le cas que le fait que les tokens de ce type ont indiqué cette chose a été causalement responsable du fait qu'ils produisent certains effets et du fait que ces effets ont été avantageux.

Dans les termes de notre exemple, cela revient à dire que si l'état C doit avoir pour *fonction* d'indiquer F, il faut que le fait que des tokens passés du type C ont indiqué F ait conféré un avantage adaptatif aux papillons qui ont été dans cet état. Mais le fait d'indiquer F ne peut être avantageux que s'il a des effets, ce qui revient à dire, en l'occurrence, que l'état C ne peut être sélectionné pour indiquer F que s'il est aussi sélectionné pour produire M, si les tokens de C qui ont produit des tokens de M les ont produits parce qu'ils indiquaient F, et si la production de ces tokens de M a été avantageuse parce que ces tokens de C indiquaient F. Il s'ensuit que dans cette perspective, le type C ne peut avoir pour fonction d'indiquer F sans avoir aussi pour fonction de produire M, et que la sélection naturelle ne peut lui conférer cette double fonction s'il n'est pas déjà le cas, *pendant la période de sélection*, que les tokens de C qui causent des tokens de M

le font (en partie) parce qu'ils indiquent F. Cela confirme apparemment que la sélection naturelle peut, en théorie du moins, conférer à C la fonction d'indiquer F. Mais peut-elle le faire sans que le fait que C cause M "résulte" du fait que C indique F, comme le souhaiterait Dretske ?

Notons d'abord que même si la réponse à cette question doit être négative, cela ne peut pas être en raison de la sélection naturelle ou de l'évolution que le fait que C cause M résulte du fait que C indique F. La sélection naturelle (ou la théorie de l'évolution) ne peut expliquer pourquoi les tokens de C qui causent des tokens de M le font parce qu'ils indiquent F, puisque c'est ce fait, au contraire, qui explique (en partie, et dans le meilleur des cas) pourquoi il y a sélection de C (pour indiquer F et causer M).

On remarquera aussi que pour que le fait que C cause M résulte du fait que C indique F, il faut, à tout le moins, que C cause M et que C indique F. Or l'état C peut être ou avoir été sélectionné pour indiquer F et causer M, sans qu'il soit (ou ait été) le cas que C indique F ou que C cause M, c'est-à-dire, sans qu'il soit (ou ait été) le cas que *tous les tokens* du type C indiquent F ou causent des tokens du type M. La sélection naturelle "explique" au mieux pourquoi le type d'état C a présentement pour *fonction* d'indiquer F et de causer un mouvement de type M (mais C peut, et doit pouvoir, avoir cette fonction sans indiquer F ou causer M). Il est donc possible que C ait pour fonction sélectionniste d'indiquer F sans qu'il soit le cas que le fait que C cause M résulte du fait que C indique F. Même en supposant qu'écrire "le fait que C cause M résulte du fait que C indique F" est pour Dretske une manière négligente d'écrire "le fait que C cause M résulte du fait que des tokens du type C *ont indiqué* F", la réponse reste la même, puisque C peut avoir pour fonction sélectionniste d'indiquer F sans que C cause M.

Ces remarques semblent donner raison à Dretske, mais ne tiennent pas compte du fait qu'en décrivant le système nerveux du papillon de nuit, celui-ci prend pour acquis que, dans l'état actuel des choses, un état de type C cause effectivement un mouvement de type M. Si on admet cela, et si on suppose que l'état C continue d'être sélectionné pour indiquer F (et causer M), alors il doit être le cas que les tokens de C qui causent des tokens de M le font parce qu'ils indiquent F. Comme par hypothèse tous les tokens de C causent des tokens de M, il s'ensuit 1) que les tokens de C causent des tokens de M parce qu'ils indiquent F (ce que je tiens pour équivalent, à toutes fins pratiques, à dire que le fait que C cause M résulte du fait que C indique F), contrairement à ce que prétend Dretske<sup>8</sup>, et 2) que

---

<sup>8</sup> Toute cette discussion n'a évidemment de sens que si on admet au point de départ qu'un type d'item peut être sélectionné pour indiquer quelque chose, ce qui ne va pas

tous les tokens de C indiquent F (car s'il y avait des tokens de C qui n'indiquent pas F, cela ne pourrait pas être parce qu'ils indiquent F qu'ils causent des tokens de M), ce qui est incompatible avec l'idée que C pourrait ne pas remplir sa fonction. Il faut en conclure que si l'analyse sélectionniste pertinente de la notion de fonction inclut la condition qu'un item ne peut avoir la fonction de produire Y au temps t que s'il est sélectionné, au temps t, pour produire Y, alors (dans la situation considérée) il n'est pas possible que C ait pour fonction d'indiquer F sans qu'il soit le cas que C indique F, et que C cause M parce que C indique F. Cette possibilité ne correspond donc certainement pas aux intentions de Dretske.

Dans le cas où l'état C aurait cessé d'être sélectionné pour indiquer F et causer M, sans pour autant cesser de causer M, est-il légitime de dire que le fait que C cause (maintenant) M "résulte" du fait que des tokens de C ont indiqué F dans le passé ? Il semble bien que oui, puisque s'il est vrai maintenant que les tokens de C causent des tokens de M, c'est en partie parce que dans le passé, des tokens de C ont causé des tokens de M et que cela a procuré un avantage aux individus dans lesquels cela s'est produit. Mais par hypothèse, les tokens de C qui ont causé des tokens de M ne l'ont fait, et cela n'a été un avantage, que parce qu'ils ont indiqué F. Par transitivité, c'est en partie parce que des tokens de C ont indiqué F dans le passé que C cause maintenant M. Notons que cela reste vrai même dans le cas où tous les tokens de C n'indiquent pas F, c'est-à-dire, même s'il n'est pas le cas (maintenant) que C cause M parce que C indique F. Selon ce diagnostic, même une analyse sélectionniste entièrement orientée vers le passé, comme celle de Neander, ne pourrait faire en sorte que C ait pour fonction d'indiquer F sans que le fait que C cause M résulte (en partie) du fait que des tokens de C ont indiqué F.

Mais Dretske conteste précisément ce diagnostic, en invoquant la distinction de Sober<sup>9</sup> (1984 : 147-155) entre une explication sélectionniste (ou "variationnelle") et une explication développementale. Une explication sélectionniste se distingue d'une explication développementale par le fait qu'elle explique pourquoi une population donnée se compose d'individus ayant telle caractéristique, sans expliquer pourquoi chacun des individus qui compose cette population a cette caractéristique, c'est-à-dire, sans supposer que les membres individuels de cette population

---

de soi, et exige qu'on accepte l'idée que le fait d'indiquer quelque chose peut être causalement responsable de la production d'effets avantageux ou constituer lui-même un avantage.

<sup>9</sup> Nagel (1977), Cummins (1975) et Lewontin (1983) défendent une position semblable.

subissent des modifications significatives pendant leur existence. Il s'agit ni plus ni moins, selon la formule de Dretske (1988c : 93), de la différence entre expliquer pourquoi mes amis s'imbibent de martinis en donnant, pour chacun d'eux, les raisons pour lesquelles ils boivent des martinis (explication développementale), et l'expliquer en soulignant que je n'admets quelqu'un dans le cercle de mes amis que s'il boit des martinis (explication sélectionniste).

Une explication développementale explique pourquoi tous les X ont le trait P en expliquant, pour chaque X individuellement, pourquoi il a (ou acquiert) le trait P ; elle prend pour acquis que l'ensemble des X se compose de certains individus déterminés, et que si les propriétés des membres de cet ensemble changent, ce doit être parce que les individus qui le composent se sont modifiés. Une explication sélectionniste, au contraire, suppose que la composition de l'ensemble des X se modifie avec le temps (en ce sens que cet ensemble se compose d'individus différents, à différents moments), mais n'exige pas que les individus qui lui appartiennent subissent des changements. Cela signifie que dans le cas, par exemple, du papillon de nuit, la sélection naturelle (par le biais du fait que des tokens de C ont indiqué F) explique pourquoi la population actuelle de papillons se compose d'individus dont le système nerveux est structuré de telle manière que C cause M, mais non pas pourquoi le système nerveux de tel ou tel papillon particulier est structuré de cette manière. Pour expliquer cela, il faut évoquer le génotype de ce papillon et les lois de l'hérédité et du développement embryologique : ce papillon est structuré de cette manière parce qu'il a hérité des instructions génétiques nécessaires.

Sans mettre en question la validité de cette distinction entre explication sélectionniste et explication développementale, on peut toutefois se demander si elle justifie vraiment la conclusion que le fait que des tokens de C ont indiqué F n'explique pas causalement (en partie) pourquoi C cause M. Il faut d'abord remarquer que Sober lui-même admet que la sélection naturelle peut réellement *expliquer* pourquoi tous les membres d'une certaine population ont un certain trait, et que rien ne permet de douter (comme Dretske (1988c : 93) semble le faire) qu'une telle explication soit une explication causale. Ce que Sober conteste, c'est qu'en expliquant cela, la sélection naturelle explique aussi pourquoi tel ou tel papillon est structuré de telle manière que C cause M. Ce faisant, il remet en question le principe, intuitivement plausible, selon lequel si une chose explique pourquoi tous les X ont le trait P, alors elle explique aussi pourquoi cet X a le trait P (qui en est une conséquence logique). Je n'ai toutefois pas l'intention d'attaquer Sober ou Dretske en me portant à la défense de ce principe général, bien qu'il me semble tout à fait correct. Il

suffit de montrer qu'en l'occurrence, Sober et Dretske ont tort de prétendre que le fait que des tokens de C ont indiqué F ne joue aucun rôle dans l'explication du fait que *ce* papillon, appelons-le Clyde, soit structuré de telle manière que C cause M, c'est-à-dire, dans l'explication du fait que les tokens du type C-dans-Clyde causent des tokens du type M-dans-Clyde.

En disant cela, je n'entends pas suggérer pour autant qu'ils auraient tort de soutenir qu'une explication génétique du fait que Clyde soit structuré de cette manière en soit une explication correcte. Le fait est qu'il peut y avoir plusieurs explications correctes, plus ou moins complètes, de la même chose. En particulier, il semble, comme le soutient Neander (1988c), que le fait que C ait été sélectionné pour indiquer F et causer M soit en partie responsable du fait que Clyde a hérité des gènes correspondants. Clyde tient ces gènes de ses parents, qui les tiennent eux-mêmes de leurs parents, etc. Mais si ces gènes se sont transmis jusqu'à Clyde, c'est certainement en partie parce qu'ils ont conféré un avantage adaptatif à ceux qui les ont eus, c'est-à-dire, parce qu'ils ont été sélectionnés pour structurer les papillons de telle manière que C cause M. Puisque par hypothèse, ces gènes ont été sélectionnés, Clyde ne pouvait pas ne pas en hériter. Et s'ils n'avaient pas été sélectionnés, Clyde n'aurait pas pu en hériter (ce qui ne veut pas dire que Clyde n'aurait pas pu les avoir, par exemple à la suite d'une mutation ; dans ce cas la sélection naturelle n'aurait pas de rôle explicatif), ou aurait eu moins de chance d'en hériter. Il est donc possible de soutenir que la sélection naturelle contribue, dans certains cas, à expliquer causalement pourquoi les membres individuels d'une population ont telle ou telle caractéristique, et en particulier que le fait que des tokens de C ont indiqué F explique en partie non seulement pourquoi C cause M, mais aussi pourquoi C-dans-Clyde cause M-dans-Clyde.

Dans le deuxième passage cité plus haut, Dretske *semble* contester que la sélection naturelle explique pourquoi C cause M, mais le fait qu'il invoque la distinction entre explication sélectionniste et explication développementale suggère qu'en réalité, il conteste seulement que la sélection naturelle explique pourquoi C-dans-Clyde cause M-dans-Clyde (c'est du moins tout ce que cette distinction lui permet de contester, comme je l'ai signalé plus haut). Cela veut dire que même si on lui concédait ce point, il devrait admettre que le fait que des tokens de C ont indiqué F explique en partie pourquoi C cause M, s'il est vrai que C cause M et que C a été sélectionné pour indiquer F et causer M, et par conséquent, que C ne peut pas avoir pour fonction sélectionniste d'indiquer F à moins que le fait que C cause M résulte en partie du fait que des tokens de C ont indiqué F. La question de savoir si la sélection naturelle

explique pourquoi C cause M sans expliquer pourquoi C-dans-Clyde cause M-dans-Clyde n'est tout simplement pas pertinente de ce point de vue.

Si une réponse positive à cette question permet de nier que le fait que des tokens de C ont indiqué F contribue à expliquer le comportement de *Clyde*, elle ne permet pas de nier que ce même fait contribue à expliquer le comportement du *papillon de nuit*. En d'autres termes, cela donnerait raison à Dretske de soutenir que la sélection naturelle n'explique pas le comportement individuel, mais non de soutenir qu'elle n'explique pas le comportement instinctif (c'est-à-dire, pourquoi les membres d'une certaine population ont un certain type de comportement).

Dretske voudrait que la différence entre les fonctions d'indication qui sont fondées sur la sélection naturelle et celles qui sont fondées sur l'apprentissage tiennent à ce que dans le deuxième cas, mais pas dans le premier, le fait qu'un état de type C cause un mouvement de type M (c'est-à-dire, un certain comportement) soit expliqué causalement (en partie) par le fait que des tokens du type C ont indiqué F. Mais il n'a pas besoin de nier que si C a pour fonction sélectionniste d'indiquer F, alors le fait que des tokens de C ont indiqué F explique pourquoi C cause M dans le papillon de nuit, ni même de nier que ce fait explique pourquoi C-dans-Clyde cause M-dans-Clyde, pour établir qu'il y a une différence significative entre les deux types de représentations qu'il veut distinguer. Ce qui importe en effet, c'est que si l'état C tient sa fonction d'indication de la sélection naturelle plutôt que de l'apprentissage individuel, alors le fait que C-dans-Clyde cause M-dans-Clyde *n'est pas* expliqué par le fait que *des tokens de C-dans-Clyde* ont indiqué F (car les tokens de C qui ont indiqué F et qui expliquent pourquoi C-dans-Clyde cause M-dans-Clyde doivent alors appartenir à des ancêtres de Clyde, non à Clyde lui-même). Il pourrait ainsi suggérer qu'un comportement d'un individu est proprement intentionnel lorsque sa cause structurante est elle-même à l'intérieur de l'individu. Il semble clair, en tout cas, que si les croyances ne peuvent pas (ou pas toutes) être des représentations qui tiennent leur fonction d'indication de la sélection naturelle, cela n'est probablement pas parce que ces dernières ne peuvent pas contribuer à l'explication structurante des comportements.

Les principales conclusions qui se dégagent de cette discussion<sup>10</sup> sont que :

---

<sup>10</sup> On notera que, dans toute cette discussion, je laisse entendre (comme Dretske d'ailleurs), que la sélection naturelle opère directement au niveau des types d'états internes isolés. Si c'était réellement le cas il faudrait conclure que si un état C a pour fonction sélectionniste d'indiquer F, alors les individus appartenant à la population concernée possèdent une représentation (éventuellement, un croyance) *innée* de F.

(4) la sélection naturelle ne peut conférer à un type d'état C, la fonction d'indiquer F, sans lui conférer aussi la fonction de produire un certain effet M (à savoir l'effet que les tokens de C qui ont indiqué F ont produit dans le passé parce qu'ils indiquaient F, et en vertu duquel indiquer F a été avantageux)

(5) la sélection naturelle peut conférer à un type d'état C la fonction d'indiquer F (et de causer M) sans que le fait que C cause M résulte du fait que C indique F ou du fait que des tokens de C ont indiqué F (puisqu'elle peut le faire sans que C indique F ou que C cause M)

(6) mais si C cause effectivement M, et si C a pour fonction sélectionniste d'indiquer F (et de causer M), alors le fait que C cause M résulte en partie du fait que des tokens de C ont indiqué F.

J'insiste, en terminant, sur le fait que dans cet article, je n'ai pas examiné la prétention de Dretske selon laquelle son analyse permettrait de rendre compte à la fois de la nature du contenu intentionnel et de son efficacité causale, bien que cette double prétention soit éminemment critiquable. J'ai simplement montré que si son analyse de l'efficacité causale du contenu vaut pour les états dont les fonctions d'indication dérivent de l'apprentissage individuel, alors elle vaut aussi, *mutatis mutandis*, pour ceux dont les fonctions d'indication dérivent de la sélection naturelle<sup>11</sup>.

Daniel LAURIER  
Département de philosophie,  
Université de Montréal

### *Bibliographie*

---

Cela n'est évidemment vrai que dans le sens où les individus ont dans leur répertoire la capacité de former une représentation de F, et non dans le sens où une représentation de F serait inscrite en eux en permanence et de manière innée. Le point de vue selon lequel la sélection naturelle pourrait conférer un contenu à certains états internes ne conduit donc pas nécessairement à la conclusion qu'il y a des représentations innées (dans le sens où on dirait qu'il y a des croyances innées), mais seulement à la conclusion que les individus naissent avec la capacité d'avoir des représentations de telles ou telles choses. En réalité, c'est sur le dispositif qui produit les représentations que la sélection naturelle opère.

<sup>11</sup> Le présent article est extrait d'un manuscrit plus vaste dont j'ai discuté différents aspects avec plusieurs personnes, et notamment avec Éric Audureau, Pascal Engel, Pierre Jacob, Martin Montminy, Élisabeth Pacherie, Claude Panaccio, Joëlle Proust et Richard Vallée. Je leur en suis tous vivement reconnaissant.

- BAKER, Lynne R. (1991) "Dretske on the Explanatory Role of Belief", *Philosophical Studies* 63, pp. 99-111.
- BOGDAN, Radu dir. (1986) *Belief*, Oxford, Oxford University Press.
- BRATMAN, Michael E. (1990) "Dretske's Desires", *Philosophy and Phenomenological Research* 50, pp. 795-800.
- BURGE, Tyler (1979) "Individualism and the Mental", French, P. A. *et al.* dir. (1979) pp. 73-121.
- BURGE, Tyler (1982) "Other Bodies", Woodfield, A. dir. (1982) pp. 97-120.
- CUMMINS, Robert (1975) "Functional Analysis", Sober, E. dir. (1984b) pp. 386-407.
- CUMMINS, Robert (1988) "Comments on Dretske", Grimm, R. H. et Merrill D. M. dir. (1988) pp. 44-55.
- CUMMINS, Robert (1991) "The Role of Mental Meaning in Psychological Explanation", McLaughlin dir. (1991a) pp. 102-118.
- DENNETT, Daniel C. (1987) *The Intentional Stance*, Cambridge (Mass.), MIT Press.
- DENNETT, Daniel C. (1990) "Ways of Establishing Harmony", Villanueva dir. (1990) pp. 18-27.
- DRETSKE, Fred (1981) *Knowledge and the Flow of Information*, Cambridge (Mass.), MIT Press.
- DRETSKE, Fred (1983a) "Précis of *Knowledge and the Flow of Information*", *Behavioral and Brain Sciences* 6, pp. 55-63.
- DRETSKE, Fred (1983b) "Author's Response", *Behavioral and Brain Sciences* 6, pp. 82-90.
- DRETSKE, Fred (1986) "Misrepresentation", Bogdan R. dir. (1986) pp. 17-36.
- DRETSKE, Fred (1988a) "The Explanatory Role of Content", Grimm, R. H. et Merrill D. M. dir. (1988) pp. 31-43.
- DRETSKE, Fred (1988b) "Reply to Cummins", Grimm, R. H. et Merrill D. M. dir. (1988) pp. 55-61 .
- DRETSKE, Fred (1988c) *Explaining Behavior. Reasons in a World of Causes*, Cambridge (Mass.), MIT Press.
- DRETSKE, Fred (1989) "Reasons and Causes", Tomberlin, J. E. dir. (1989) pp. 1-15.
- DRETSKE, Fred (1990a) "Does Meaning Matter ?", Villanueva dir. (1990) pp. 5-17.
- DRETSKE, Fred (1990b) «Précis of "Explaining Behavior : Reasons in a World of Causes" », *Philosophy and Phenomenological Research* 50, pp. 783-786.
- DRETSKE, Fred (1990c) "Reply to Reviewers", *Philosophy and Phenomenological Research* 50, pp. 819-839.
- DRETSKE, Fred (1990d) "Putting Information to Work", Hanson dir. (1990) pp. 112-124.

- DRETSKE, Fred (1991a) "Dretske's Replies", McLaughlin dir. (1991a) pp. 180-221.
- DRETSKE, Fred (1991b) "How Beliefs Explain : Reply to Baker", *Philosophical Studies* 63, pp. 113-117.
- DRETSKE, Fred (1993a) "Mental Events as Structuring Causes of Behaviour", Heil et Mele dir. (1993) pp. 121-136.
- DRETSKE, Fred (1993b) "The Nature of Thought", *Philosophical Studies* 70, pp. 185-199.
- FODOR, Jerry A. (1987) *Psychosemantics*, Cambridge (Mass.), MIT Press.
- FODOR, Jerry A. (1990) *A Theory of Content and Other Essays*, Cambridge (Mass.), MIT Press.
- FRENCH, Peter A., UEHLING, T. E. et WETTSTEIN, H. K. dir. (1979) *Midwest Studies in Philosophy* 4, Minneapolis, U. of Minnesota Press.
- GODFREY-SMITH, Peter (1989) "Misinformation", *Can. J. of Phil.* 19, pp. 533-550.
- GODFREY-SMITH, Peter (1991a) *Teleonomy and the Philosophy of Mind*, thèse de doctorat, U. of California, San Diego.
- GODFREY-SMITH, Peter (1991b) "Signal, Decision, Action", *J. of Phil.* 88, pp. 709-722.
- GODFREY-SMITH, Peter (1992) "Indication and Adaptation", *Synthese* 92, pp. 283-312.
- GRICE, Paul (1957) "Meaning", Grice (1989) pp. 212-223.
- GRICE, Paul (1989) *Studies in the Way of Words*, Cambridge (Mass.), Harvard U. Press.
- GRIMM, R. H. et MERRILL, D. M. dir. (1988) *Contents of Thought*, Tucson, U. of Arizona Press.
- HANSON, Philip P. dir. (1990) *Information, Language, and Cognition*, Vancouver, U. of British Columbia Press.
- HEIL, John et MELE, Alfred dir. (1993) *Mental Causation*, Oxford, Clarendon Press.
- HORGAN, Terence (1991) "Actions, Reasons, and the Explanatory Role of Content", McLaughlin dir. (1991a) pp. 73-101.
- JACOB, Pierre (1993) "Un moteur peut-il être sémantique ?", *Dialogue* 33, pp. 527-544.
- KIM, Jaegwon (1990) "Explanatory Exclusion and the Problem of Mental Causation", Villanueva dir. (1990) pp. 36-56.
- KIM, Jaegwon (1991) "Dretske on How Reasons Explain Behavior", McLaughlin dir. (1991a) pp. 52-72.
- LAURIER, Daniel (1994) "Pangloss, l'erreur et la divergence", *Journal of Philosophical Research* 19, pp. 343-368.
- LEWONTIN, Richard (1983) "Darwin's Revolution", *New York Review of Books* 30, pp. 21-27.

- LOEWER, Barry (1987) "From Information to Intentionality", *Synthese* 70, pp. 258-317.
- LOEWER, Barry et REY, Georges dir. (1991) *Meaning in Mind : Fodor and his Critics*, Oxford, Blackwell.
- McLAUGHLIN, Brian (1987) "What is Wrong with Correlational Psychosemantics", *Synthese* 70, pp. 271-286.
- McLAUGHLIN, Brian dir. (1991a) *Dretske and his Critics*, Oxford, Blackwell.
- McLAUGHLIN, Brian (1991b) "Belief Individuation and Dretske on Naturalizing Content", MacLaughlin dir. (1991a) pp. 157-179.
- MILLIKAN, Ruth G. (1984) *Language, Thought and Other Biological Categories*, Cambridge (Mass), MIT Press.
- MILLIKAN, Ruth G. (1990a) "Compare and Contrast Dretske, Fodor, and Millikan on Teleosemantics", Millikan (1993) pp. 123-133.
- MILLIKAN, Ruth G. (1990b) « Seismograph Readings for "Explaining Behavior" », *Philosophy and Phenomenological Research* 50, pp. 807-812.
- MILLIKAN, Ruth G. (1991) "Speaking up for Darwin", Loewer, B. et Rey, G. dir. (1991) pp. 151-164.
- MILLIKAN, Ruth G. (1993) *White Queen Psychology and Other Essays for Alice*, Cambridge (Mass.), MIT Press.
- NAGEL, Ernest (1977) "Teleology Revisited", *J. of Phil.* 74, pp. 261-301.
- NEANDER, Karen (1988) "What Does Natural Selection Explain ? : Correction to Sober", *Phil. of Science* 55, pp. 422-426.
- NEANDER, Karen (1991a) « The Teleological Notion of "Function" », *Aust. J. of Phil.* 69, pp. 454-458.
- NEANDER, Karen (1991b) "Functions as Selected Effects : the Conceptual Analyst's Defense", *Phil. of Science* 58, pp. 168-184.
- PAPINEAU, David (1987) *Reality and Representation*, Oxford, Blackwell.
- PUTNAM, Hilary (1975) *Mind Language and Reality*, Cambridge, Cambridge U. Press.
- SOBER, Elliott (1984a) *The Nature of Selection*, Cambridge (Mass.), MIT Press.
- SOBER, Elliott dir. (1984b) *Conceptual Issues in Evolutionary Biology*, Cambridge (Mass.), MIT Press.
- STALNAKER, Robert (1984) *Inquiry*, Cambridge (Mass.), MIT Press.
- STAMPE, Dennis W. (1990) "Desires as Reasons", *Philosophy and Phenomenological Research* 50, pp. 787-793.
- STICH, Stephen P. (1990) "Building Belief : Some Queries about Representation, Indication and Function", *Philosophy and Phenomenological Research* 50, pp. 801-806.
- TOMBERLIN, James E. dir. (1989) *Philosophical Perspectives 3 : Philosophy of Mind and Action Theory*, Atascadero, Ridgeview.

- TUOMELA, Raimo (1990) "Are Reason-Explanations Explanations by Means of Structuring Causes ?", *Philosophy and Phenomenological Research* 50, pp. 813-818.
- VILLANUEVA, Enrique dir. (1990) *Information, Semantics and Epistemology*, Oxford, Blackwell.
- WOODFIELD, Andrew dir. (1982) *Thought and Object*, Oxford, Oxford U. Press.