

Intentionality and Evolution

Joëlle Proust

Behavioural Processes, 35, 1-3, 1995, pp. 275-286.

Abstract

Three conditions are, according to Dretske, necessary for representations to be causal in an organism's behavior, and typically not fulfilled by phylogenetic representations: a) The system must be able to pick up present information ; b) it must use that information to satisfy its various needs and purposes ; c) the piece of information to which the system responds must be available to it in some central way.

Condition c is first discussed: it is argued that, in some plausible interpretation of "system as a whole", at least some types of information can shape behavior in a teleological way without being available to the system "as a whole". Then condition b is shown not to provide in itself any way of establishing the privileged status of learning over other kinds of flexible, information-driven processes in goal-oriented behaviors. A third section examines the notion of present versus past causal roles, which are supposed respectively to belong to informational contents and to genetic plans. To this, it is objected that learned correlations share with phylogenetic representations such a temporal gap and that informational content did play a role in recruiting a particular phylogenetic representation to control motor output.

Keywords

Belief ; causality ; evolution ; information ; intentionality ; representation.

The notion of intentionality as philosophers currently understand it does not have to do with goal-directedness, but with Brentano's concept of intentionality as aboutness (Brentano, 1924-28, 102), i.e. with representational capacity. One of the prominent questions raised by the development of cognitive science is to explain intentionality in a scientifically acceptable way, in other words, to "naturalize" it. It is important to note at the outset that, when philosophers commit themselves to this project, they should be ready to accept any constraints and results delivered by the different sciences that are relevant to such an ambitious project, such as biology, ethology, experimental psychology and the neurosciences. The most promising way towards naturalizing intentionality is presently teleosemantics: a crucial factor allowing an animal to represent states of the environment consists in the fact that certain internal states or mechanisms have a definite function.

According to the most general formulation of Dretske's own definition (Dretske, 1988, 84), a representation is an indicator, natural or conventional, whose function is to indicate what it does. An indicator is one of two *relata* in some direct or indirect causal chain, consequent on the other. While fire causes smoke, smoke indicates fire. Analogously, a certain pattern on the retina, or a certain neuronal vector indicate some external condition. Now an animal able to detect, let us say, a predator, from a fixed set of features or from some kind of sensory pattern, will have the corresponding representation insofar as the latter controls appropriately its present or future motor output, acquiring thereby a definite function. The predator representation, therefore, will carry an informational content, in virtue of the fact that it indicates an external condition (such as a predator present in the vicinity); and it will control the appropriate motor output that it was recruited to control according to the content it carries.

At this point, most naturalistic philosophers should agree, because no substantial claim has yet been made. For according to the notion of information involved, any two events which are related by some kind of regular connection, like the event of rain and the formation of puddles, or fire and smoke, or presence of a predator nearby and a neural vector being activated in the prey's brain, are conversely related by a reciprocal indication or information-carrying relation: puddles indicate rain, smoke indicates fire, a relevant brain state indicates predator.

Now the fact that this information carrying indicator is correlated with some kind of motor output - orienting behavior, tracking, etc. - and turned into a representation with a definite function, can be seen as a consequence of learning, or as a consequence of biological structure. Interestingly enough, philosophers seem to have strong diverging intuitions as to which processes are responsible for recruiting an indicator in a given representational function. While some philosophers, such as Millikan, are only interested in biological functions and balk at taking the causal role of informational content into account (Millikan, 1993, 126 sq.), others, such as Dennett (1969, 46 sq.), are happy with the idea that a representation is nothing else than an information bearer that evolves in a motor control. But others, such as Dretske, insist that representations should have something more significant to do, namely to allow the organism which has them "to steer". In this view, intentionality becomes a property of reason-guided behavior, which requires not only that a particular internal state release some appropriate motor program, but also that the meaning of the representation be "relevantly engaged in the production of the output" (Dretske, 1988, 94). The meaning must be one "to or for the animal in which it occurs" (95).

In what follows, I will examine three arguments advanced by Dretske (1988) intended to narrow-down mental content to types of representations that are available to an individual system taken as a whole. If my responses to these arguments are sound, this should tend to suggest that the cake of intentionality could be sliced in a different way, and, in particular, that the role of evolution may have been underestimated in the analysis of internal states whose content is causally efficacious.

According to Dretske, the causal-explanatory fact through which genetically determined behaviors relate to external conditions is not located at the appropriate level for intentional content to play a causal role. In a genetically determined behavior, the internal state responsible for motor output has been recruited not because of the presence in the environment of properties that the individual learned to use, making thus the internal indicator into a representation, but because the past continued link between events having those properties and some corresponding beneficial reactions fixed in the species a mechanical, hard-wired sequence in which the representational contents played no role.

For example, noctuid moths have receptors designed to respond to high-frequency sounds, the latter being normally emitted by a bat. The response is either turning away from the source, in the case of low frequency sounds, or diving and spiralling in the case of high-frequency sounds. Evolution theory explains why internal states of the frequency detectors cause the animal to turn away or to dive: this mechanism confers on the moth a competitive advantage. But it is not able to explain why this particular moth does whatever it does ; using Sober's distinction (Sober, 1984), Dretske says that the explanation here is selectional , not developmental. No particular fact about this moth can help explain why it does what it does. The reason is that the relation of indication in this case rests between a type of internal state and a type of condition which has been part of the moth's environment long enough to constitute a selective pressure for an appropriate genetically induced response. If the external environment suddenly changed, the moth would still respond to high frequencies in the same way, and the relationship between this internal state token and the present state of the external world would fail to explain why this moth does what it does. In other words, present informational content is in this case causally idle. This kind of behavior is a tropism, which fails to give reasons for the token behavior, because "what the indicators indicate is irrelevant to what movements they produce" (Dretske, 1988, 94). Again, the system does not move as it does now because of what C (= the relevant internal state) means now about external condition F, but because of some mechanical or feedback process which cannot be modified by learning. One cannot then, according to Dretske, attribute the behavior of the organism to any reason it could entertain.

What Dretske means in the moth's case summarized above is not that there is no informational content in an internal state, eg. in the moth's auditory vector correlated to high frequency emissions of a bat, but that this content plays no role in the moth's present behavior. It cannot be seen as a reason that the moth in question can entertain. So let us try to spell out the conditions that make a content relevant in causing behavior.

To qualify as mental content, Dretske says, information must be cognitively available to the subject, and not simply present "objectively" in its receptors' relevant states. This makes the noctuid moth's internal state "not relevantly engaged in the production of output". Although it has "meaning of the relevant kind, this is not a meaning it has to or for the animal in which it occurs" (1988, 94-5).

This set of prerequisites for some internal state to have causal content can be reconstructed through the following line of reasoning. First, as we saw above, an internal state should according to Dretske be granted a causally efficacious meaning only if that meaning is “instrumental in shaping the behavior” (1990a, 14). Second, three conditions at least must be present for such a shaping to happen. a) The system must be able to pick up present information (1978, 115). This condition is fulfilled only in systems able to learn. Evolutionary solutions - such as tropisms, of fixed action patterns - precisely lack this feature: they evolved as a result of past correlations between a type of system and a type of environment. Correlatively, causality in these processes is secured not by any informational content, but by genes (1989, 12). b) It must use that information to satisfy its various needs and purposes, i.e. respond behaviorally to the content of its internal states. c) A system can respond to information only if that information is available to it in some central way: “You earn no cognitive credits for the detective capabilities of your parts - not unless the results are made available to you for modulation of your response” (1978, 113).

I will not dispute the first step of the reasoning, and will accept that intentional states should have some distinctive causal efficacy in shaping the behavior. What I will question is the relevance of conditions a, b and c. I will first address condition c, and suggest that, in some plausible interpretation of “system as a whole”, at least some types of information can shape behavior in a teleological way without being available to the system “as a whole”. In the following section, I will argue that condition b does not provide in itself any way of establishing the privileged status of learning over other kinds of flexible, information-driven processes in goal-oriented behaviors. Finally I will examine in a third section the notion of present versus past causal roles, which are supposed respectively to belong to informational contents and to genetic plans.

1) Availability of the information to the subject.

Let us first turn to the distinction between an information being “available to the whole system”, and an information being available to one of its parts. There is indeed an important distinction to be drawn between purely physical indicators and cognitive ones. We can suppose that the condition of a sun-burnt skin indicates exposure to sun ; such an indicator does not qualify as a cognitive indicator, and in particular cannot be described as “sun-memory”. (Dretske, 1978, 125, fn 3). This distinction cannot, on pain of circularity, be substantiated by appeal to what the subject knows. What is crucial in this distinction is not that a subject should be aware of a piece of information affecting her behavior (still less that her conscious belief be true), but, Dretske says, that this piece of information could modulate the subject’s behavior.

Let us suppose that I have an information-bearing (external or internal) state, such as a sunburn, which prevents me from doing some planned work. I indeed have a causally efficacious informational state. Does my sunburn qualify as mental content ? It fails do so for two reasons:

first, the informational content of the state is not intrinsically causal, just like the meaning of the soprano's song is not intrinsically causal in shattering a glass ; second, the indicator does not have a fortiori the function of indicating what it does: the neurophysiological consequences of solar radiation on my alertness, did indeed causally affect my general activity, but they were not as such "recruited" to do what they did.

While one can accept the requirement that a "cognitive" type of detector "should make this information available to the system of which it is a part for purpose of shaping its responses", it is unclear whether this requirement involves the relevant piece of information being made available to the system "as a whole" (ibid. 114). How exactly are we to decide what a "whole" system amounts to ? Dretske cannot be suspected of requesting from a cognitive state that it be conscious (1978, 112) ; aside from the imprecision of such a property, it is well known that the responses of a subject, whether human or animal, can be influenced by signals that do not reach awareness. In the absence of an output-independent distinctive property - such as conscious accessibility - , the question of deciding whether such signals have been made available to "the system as a whole" runs the risk of being trivialized: some information is available to a system "as a whole" just in case that system responds to it. And a system responds to some piece of information just in case the latter was made available to it "as a whole". In fact, the example introduced above of the noctuid moth shows that such a system may respond to some external cue through the activation of some internal state without the latter having meaning for the animal. Therefore, the trivial answer is not even true.

Philosophers usually consider that the appropriate level is the one in which beliefs and desires can be described as reasons guiding a subject's action. So does Dretske: an internal representation controlling behavior will fail to qualify as a belief if it is not "relevantly engaged in the way it steers" (1988, 94). And it is relevantly engaged if "the structure's indicator properties figure in the explanation of its causal properties", i.e. if "what it says (about external affairs) helps to explain what it does". (1988, 94). Therefore "having meaning to or for the animal" seems to be the distinctive feature of a belief (or a desire), i.e some indicator property having an intrinsic causal role.

A subject failing to identify one of its beliefs or desires as being the cause for some action should not in itself raise a difficulty for the "system-as-a-whole" approach, insofar as being available to the subject, as noted above, does not equate with being consciously accessible by the subject. In what follows I would rather defend a different view: an information can well be cognitively efficacious without being available to the system as a whole, in the sense that the system may rationalize its actions using a set of representations distinct from those involved in the causation of the actions. The fact that a subject may in good faith advance false reasons for her

own behavior puts in serious doubt the very existence of some flow-chart box where all information would be put together. More importantly, it suggests that behavior may be caused in a variety of ways, while the subject only has the belief-desire mode at her disposal to explain her own (and others') behavior (a similar point is raised by Stampe (1990, 791 sq.).

Let us illustrate this difficulty with one example from Nisbett & Wilson (1977). In a study conducted in a store under the guise of a consumer survey, subjects were asked to say which one among four identical pairs of nylon stockings was the best quality and to explain their choice. Most subjects (75%) declared that the stockings farthest to the right were the best, thus falling prey to a well-documented, if poorly understood, position effect. All subjects offered various rationalizations to explain their choice, while firmly denying that position could have influenced them. Is that kind of belief one entertained by the system "as a whole" ? The belief that every subject had as to which property was guiding her choice is clearly mistaken: had the chosen pair been located elsewhere, it would have probably failed to be chosen. Nor can we deny that her choice was grounded in representations. For the position bias has the characteristics of an indicator for a spatial property whose function is to control the behavior. It should then be taken to be intentional in the strict sense of being a bona fide mental representation, i.e. an internal state that controls motor output in virtue of its informational content, and that has this very function. Now it could be submitted either that the position bias involves a tacit belief to the effect that "right is best" or that it belongs to some lower-level, non propositional format of representation ; which option to choose certainly cannot be decided on purely conceptual grounds, and may seem at first blush to be independent of the issue of intentionality. I will try to show that it is not.

Were it required from any truly cognitive representation that it should be available to the system as a whole, it is clear that neither the causally efficacious, nor the reported propositional attitude content would do. For the one is available to the system in the sense that it controls its action, while the other is available to it in the sense that it controls the report of its action. In other words, each candidate does not qualify as having the right kind of global availability supposedly necessary for cognition. Therefore the issue turns out to be indeed quite central to determining what is a causally efficacious meaning.

The question of determining revised criteria for intentional behavior opens a difficult dilemma. On the one hand, one wants to make intentionality coincide, no matter what, with a belief-desire psychology, in order to make sense of rational behavior. Then however, one will have to admit that a significant number of the informational states that regularly control motor outputs are not intentional in that sense. The price to pay will be that a number of responses will turn out not to belong to either intentional or non-intentional activity (in the Nisbett-Wilson type of cases). On the other hand, one wants to make intentionality coextensive with representations at large in the sense already indicated. Then however, the price to pay will be a willingness to discount the reasons given by a subject, and to have an active informational content override a spurious and idle self-attributed representational content.

Now selecting the second horn of the dilemma may be found more sensible, in particular if one wants to give an account of intentionality which applies across species. The requirement that information be “available to the system” then only means that this information is represented and transformed in ways analogous to the familiar functional flow-chart. Each relevant new input should in principle be able to influence present or future activity in multiple functional ways (memorizing, planning, etc.) if that input is used in a cognitive representation. Information that would be only local (i.e. not available to any mental function) would fail to achieve a cognitive status. But understood this way, the requirement looks rather vague and possibly circular as long as we don’t have an independent, non-representational account of those functions. One might articulate it more exactly by using computational norms of what it takes for information to be of a non-local, i.e. a non purely proximal type (see on this issue Proust, forthcoming).

The point of this section is to suggest that there may be more to mental causation than what the belief-desire level may capture ; if this is recognized, then the very requirement that a meaning be one “to the whole system” dissolves.

But most authors preferred to explore the first horn of the dilemma, and tried to make belief-desire psychology coextensive with sensitivity to consequence. Let us now turn to this second condition, which might help us define more precisely how information should contribute to shaping motor output.

2) The pragmatic view: intentionality as sensitivity to consequence.

As Dretske (1978, 115) himself suggests, a promising way of understanding the expression “system as a whole” is to look at the organism engaged in its environment ; we can propose that an organism has intentionality if and only if it behaves appropriately given its needs and purposes, i.e. rationally. Or we can make the weaker claim that an organism may have intentionality if it meets some specified criteria concerning the use being made of various kinds of input information to control motor output, even though it may not act according to its own best interest. This pragmatic view is proposed, among many other authors, by Dennett (1969) for discriminating “intelligent” from “non-intelligent” information storage. “The criterion for intelligent storage”, Dennett says, “is the appropriateness of the resultant behavior to the system’s needs given the stimulus conditions of the initial input and the environment in which the behavior occurs”. (p. 46-7).

Hayes & Dickinson (1993) develop the same kind of proposal as Dennett’s and claim that only when a response is sensitive to its causal consequence can it be called intentional (in the broad sense of cognitive-representational used here). They argue that even intentional systems frequently fail to display sensitivity to consequences in particular types of responses. For example, an approach response can be learned by rats which, upon hearing a tone, approach the bowl even

when presentation of food is omitted (Holland, 1979). In that case, “the animals can never experience a contingency that would support the appropriate causal belief for approach during tone”. (p.109).

It is tempting to say that the rats do not behave under the belief that [tone + approach] predicts [food], for the evidence goes just in the other direction. We can assume (taking the second horn of the dilemma) that they act under no particular belief to this effect, but rather that their movement is caused by some innate representation, here a fixed action pattern triggered by the tone as a predictor of food. This example would be another piece of evidence that the belief-desire level is not causally sufficient to account for motor outputs, even at the level of meaning-triggered behaviors. But we may as well, as Hayes & Dickinson do (taking the first horn of the dilemma) conclude from this that only some of the rats’ or of the humans’ motor outputs qualify as intentional. When a rat presses a lever, he does it because he believes that he will get food as a result. But when he cannot refrain from approaching a bowl, he does it with no belief, i.e. in a non-intentional way.

This second condition for joining the intentional club has to do with a feature of goal-oriented systems. By definition a goal-oriented system develops a course of action until a certain final state is reached, which we shall call the target event. Sensitivity to consequence can be seen to be involved at least in three different independent ways in goal-oriented behaviors.

- A - It can be understood as what allows an organism to interrupt its course of action, either
- (i) because the target event, or some event covarying with it, is reached, or
 - (ii) because some precondition of the target event is missing.

As illustrated by Jean-Henri Fabre’s celebrated description of the Chalicodome’s egg-laying behavior, fixed action patterns may not be amenable to any other interruption besides those of type (i). The covarying event may be reduced in that particular case to the completion of the relevant action pattern. Although such an end-signal may be reliable in most evolutionary contexts, it is clearly not in a laboratory experiment such as Fabre’s.

B - It can be taken as a capacity of an organism engaged in some action to extrapolate (Rosenblueth et al., 1943), i.e. to compute later properties of the target element from current perceptual data. A missile homing on a target can use negative feedback to reduce its distance from the target ; a predator is usually able to modify its planned motor pattern using visual or auditory feedback.

C - It can be identified with learning, i.e. a capacity of modifying one’s current behavior in the light of former outcomes of actions of the same type. Learning involves something more than A and B types of sensitivity to consequence: it requires a capacity for changing - as opposed to interrupting or adapting - courses of action.

Dretske obviously favors C as being the only viable candidate for internal states having a causally efficacious meaning. Although A and B may be present in evolution-designed action

patterns, they supposedly fail to display the kind of sensitivity which allows an animal to act from his beliefs. What is distinctive of C is, for Dretske, the fact that a historical encounter with an external condition modifies the later dealings of an individual organism with that condition. Only then, it seems, is information genuinely causal.

But what about B, or Aii ? Is information not, in these cases, causally responsible for the organism being able to eventually reach its target ? Does not some recently acquired information contribute to modifying adequately the subject's course of action ? The only response which Dretske can offer is that evolution-based indicators allowing to accomplish A- or B-types of flexible behaviors are not brought to existence by the (present) meanings of the relevant internal states, but by "solutions now encoded in, and explained by, the genes". (1989, 12). A satisfactory argument cannot thus be completed from condition b alone: we need to defer to condition a to get a full evaluation of Dretske's anti-evolutionary argument.

3) Individual and generic explanations.

In order to block the attribution of intentional content to wired dispositions such as fixed action patterns and tropisms, we are left with Dretske's ultimate point: phylogenetic representations do have a content, but that content is not mental, insofar as it does not have an intrinsic causal role in the behavior of an individual animal.

This argument dwells upon a classical analysis of philosophers of biology to the effect that natural selection cannot explain why organisms have the properties for which they are selected. (Dretske, 1988, 92, Cummins, 1984). In particular, natural selection cannot account for any intentional property causing behavior in some individual animal, for, even though evolution-selected representations may control behavior, they do not do so in virtue of their content, but by blind feedback selective processes.

Why is the moth doing what it does when its auditory receptors track some high frequency pattern ? To have intentional content, receptors should capture distal information. (or, in Dretske's words, information about the present predator), "whereas the control circuitry in this moth has nothing to do with what this [internal state] C indicates" (p. 92)).

Dretske's strategy consists in deliberately conflating two claims. One belongs to the philosophy of biology, and states what evolution theory can or cannot explain. According to him, it can explain facts about general features of the species, not particular features of each and every member of the species ; the latter explanation can only be given in terms of genetic mechanisms. The other claim, philosophical, spells out what intentionality should involve, i.e. agency in a present, ever changing context, involving a competence for distal representations. Dretske argues that a system with no capacity for assessing how the world is at the moment cannot be said to control "on line" its motor outputs. It is comparable to some elementary feedback mechanism, acting like a thermostat according to pre-established values of proximal stimulations.

I will briefly address the first condition, which would require much longer considerations. My general argument will be that while the philosophical point about distality is entirely sound, and should be analysed in much greater detail if, as I believe, it is the key to intentional capacities (Proust, forthcoming), the first point (on evolution theory) should be re-examined in the light of recent discussions, the upshot being that a causal role may be given to selected as well as to learned meaningful internal states.

Let us summarize the central objection in Cummins' version. Cummins discusses a widely held claim about biological function (Cummins' analysis applies as well to the functions of artefacts, but I will ignore this possible extension, which does not play any role in Cummins' article). The claim is that "the point of functional characterization in science is to explain the presence of the item that is functionally characterized" (Cummins, 1984, 386). Such a teleological strategy is exemplified by Millikan's notion of a proper function, insofar as the item having such a proper function is supposed to be there because of its having performed whatever its job was in some self-replicating lineage. Expressed in our words, Millikan's proper function explains the representational capacity of an indicator by the fact that the possession of this function by some cerebral mechanism made possible the survival and the reproductive success of a system having it.

The very idea of explaining the presence of a function, like the intentional capacity of an indicator, by its survival value, is for Cummins the result of a deep mistake. It is not uncommon to reason from the existence of a function to the presence of structures and processes that carry out the function ; but such an inference to the best explanation, (as the one which derives, for example, chlorophyll from photosynthesis), does not amount to an explanation. Whereas one may infer function from appropriate structures, one may not derive appropriate structures from function.

Evolutionary theory is often understood as allowing one to make the latter inference. The reproductive value seems to represent a causal factor explaining why some kind of structure performs some kind of useful function. But in fact, as Cummins points out rightly, the processes which have a particular function "are totally insensitive to what that structure does". The mistake about evolution theory responsible for the failure to appreciate this fact is a confusion between genetic change and effects of the presence of functional structures and processes. Whatever happens in the phenotypes' lives, whether beneficial or not, cannot alter the genetic plan. The causes that affect this plan are various, including the random genetic drift, pleiotropy, allometry etc. But the beneficial character of some change in the context in which the plan is exercised will not alter the plan ; it will only alter the frequency of the trait within the species.

We can nevertheless derive some second order causal consequences from a function being well fulfilled by some structure. "Natural selection cannot alter a plan, but it can trim the set" (Cummins, 1984, 394). But again this trimming factor does not allow one to explain why, for example, contractile vacuoles occur in certain protozoans. It only allows to explain, Cummins

says, why the sort of protozoan incorporating contractile vacuoles occur. What natural selection explains, thus, is not why a particular element is there (an explanation which would be given in terms of the function it helps fulfill), but how a particular structure having some particular function contributes to the capacities of the containing system.

Drawing consequences from the trimming activity of selection pressures implies a change in level, from individual characteristics to population features. Indeed it is at that level that selection pressures operate. They determine the relative fitness of individuals having or not having some particular trait, given assumptions on stability and heritability. (Sober, 1984, 151) Therefore we are entitled to say, for example, that the opposable thumb was beneficial in a certain population, and, being heritable, became selected. We can also say that the functional character of the thumb explains its frequency in the population, which is not to say that it explains its presence in the gene pool. The difference is a difference in causal-explanatory level. The property of having an opposable thumb is not the same when considered at the level of the genetic mechanisms explaining its phenotypical expression in an individual and at the level of the evolutionary mechanisms which allowed that property to be selected for, against other alternative traits (ibid., 152).

Dretske concludes from Sober's argument that no individual intentional property can be selected because of the fact that it has some specific property, such as indicating F. Because the selectional cause is indirect, second-level, it cannot explain in any specific way why any particular organism has this property. It can only tell why such a property, once present by way of some genetic mechanism, did spread in the population. If the argument is sound, the conclusion one should have to draw is that only at a population level can intentionality be explained on evolutionary grounds. Men and other higher mammals would then qualify as having a selected competence for believing, but only individual correlations F-C in an organism, i.e. only specific historical exposures of a particular system to some relevant condition, could explain the recruitment of particular intentional states as causes for motor output. But is this conclusion forced upon us? I will try to present briefly two arguments to the contrary.

A recent discussion in the field of philosophy of biology (Neander, forthcoming) suggests that the argument summarized above should not lead one to conclude that natural selection has nothing to say about the phenotypic or genotypic properties of individuals. As Neander aptly shows, natural selection is a cumulative selection process, in which each sequence alters the outcome of future sequences. When Sober, Cummins and Dretske insist that mutation, and not selection, brings genetic sequences into being, they do so because they focus on an isolated selection step, while ignoring "what happens when selection is followed by further genetic recombination and mutation: preceding selection can dramatically alter the probability of subsequent gene sequences occurring" (Neander, forthcoming). Selection does not work simply by "pruning the tree of life" or "trimming the set": it changes the probability of specific subsequent mutations by actually eliminating less fit types. If it turns out to be successful, Neander's attempt at introducing

developmental causality inside selectionist pressures would allow us to recognize that content may be causal even in evolution-selected representations.

What then of the last Dretsian objection against phylogenetic representations having causal powers: does the past character of their intentional properties make them unfit to cause present behavior, in contradistinction with present learned correlations ? In Dretske's approach, as we saw above, an indicator gets its function to represent what it indicates through a process of "recruitment", in virtue of which that indicator becomes harnessed to some motor output (by operant conditioning). As some critics have emphasized, (Stampe, 1990), the indicator tokens which were decisive for the internal state acquiring this particular representational function belong to the past history of the organism ; learning makes present informational content into a behavioral cause just in case we assume that the learning situation did not change, i.e. that the indicator's present token continues to covary with an external condition of the same type. If it were not the case, as Stampe notes, Drestke's account would be "vulnerable to the claim that learning processes are no more relevant to the causal efficacy of resulting current tokens of the state type, than are selectional processes" (*ibid.*, 793). Nothing in Dretske's response contradicts Stampe's point: "Obviously, he says, the fact that a present token of C is causing M is not to be explained in terms of what it indicates since it may not indicate anything at all relevant" (*ibid.*, 830). Is it not precisely this argument that was used against phylogenetic representations having a presently relevant causal content ?

Now the problem deepens when one realizes that selectionist pressures do not belong to the past any more than learning processes do: they are still being exerted on each phenotype, by determining the specific properties that raise its propensity to reproduce. If one adopts a propensity theory of fitness, which we have independent reasons to do, we must also acknowledge that the reliability of an indicator is a property that, among others, constitutes the fitness of an individual organism carrying it as part of its phenotype. When Dretske insists that in a phylogenetic representation, "at no stage of the process is the fact that F is being indicated by C part of the explanation of why C is causing M" (1990, 829), he ignores the fact that at least at some stage, the informational content of C must have been helpful to the phenotype's reproductive capacity ; and only if C was reliably correlated to F could C happen to be helpful to survival and reproduction. Therefore, after all, some instrumental intentional properties (that C indicates reliably F, and helps S mate, run away, etc.) did contribute to causing -not the production, but - the selection of indicators having C-type internal states, just as learning did contribute to causing - not the production, but - the selection of similar indicators in higher organisms.

Conclusion

My point in this paper was to clarify and evaluate Dretske's apparently powerful arguments to the effect that no content playing an adequately causal role can be inherited and propagated by

natural selection. I hope that it is now clear that there are no compelling arguments to this effect. First, intentional-causal powers may have to be granted to structures other than internal states of the belief-desire variety, jeopardizing the “system as a whole” scheme. Second, sensitivity to consequence can be secured in goal-directed behaviors in ways more primitive than, and independent from, learning. Third, the distinction between developmental and selectional arguments might not be as clear-cut as maintained by Sober, Dretske and Cummins. Moreover, the analogy between evolution and learning might be stronger than Dretske would lend us to believe. It has to be left to another paper, however, to show how an informational model of intentionality can fill the new agenda.

Acknowledgements

I express my gratitude to Mark Cladis for his linguistic help. I also thank Fred Dretske, Karen Neander, Elisabeth Pacherie and Georges Rey for helpful discussions, as well as an anonymous referee of this journal for his/her comments.

References

- Brentano, F., (1924-1928, *Psychologie vom Empirischen Standpunkt*, 3 vol., Leipzig, Felix Meiner Verlag.
- Carpintero-Garcia, M. (1995), “Dretske on the causal efficacy of meaning”, *Mind and Language*, forthcoming.
- Cummins, R. (1975), Functional Analysis, *Journal of Philosophy*, 72: 741-760 ; reprinted in E. Sober (ed.), *Conceptual Issues in Evolutionary Biology*, Cambridge, MIT Press, 1984.
- Dennett, D.C. (1969), *Content and Consciousness*, London: Routledge and Kegan Paul.
- Dickinson, A. & Balleine, B., (1993), Actions and Responses: The dual psychology of behavior, in N. Eilan, R. McCarthy & B. Brewer (eds.), *Spatial Representation*, Oxford, Blackwell, 277-293.
- Dretske F.,(1978), “The Role of Percept in Visual Cognition”, in C. Wade Savage (ed.), *Minnesota Studies in the Philosophy of Science*, vol IX: *Perception and Cognition ; Issues in the Foundations of Psychology*, 107-127.

- Dretske, F., (1986), "Misrepresentation", in *Belief*, R. J. Bogdan ed., Oxford, Oxford University Press.
- Dretske F.,(1988), *Explaining Behavior, Reasons in a World of Causes*, Cambridge, MIT Press.
- Dretske F., (1989), "Reasons and Causes", in J.E. Tomberlin (ed.), *Philosophical Perspectives*, 3, *Philosophy of Mind and Action Theory*, Ridgeview Publishing Company, 1-15.
- Dretske F., (1990a), Does Meaning Matter ? in *Information, Semantics and Epistemology*, E. Villanueva (ed.), pp. 7-17, Oxford, Blackwell.
- Dretske F., (1990b), Seeing, Believing and Knowing, in *Visual Cognition and Action*, D. Osherson (ed.), vol. 2, Cambridge, MIT Press.
- Dretske F., (1990c) "Précis of "Explaining Behavior: Reasons in a World of Causes", *Philosophy and Phenomenological Research*, vol. 1, 4, pp. 783-786 & 819-839.
- Dretske F., (1991), How beliefs explain: Reply to Baker, *Philosophical Studies*, 63, 113-117.
- Dretske F., (1993) The nature of thought, *Philosophical Studies*, 70, 185-199.
- Dretske F., (1994), Modes of perceptual representation, in R. Casati, B. Smith & G. White, (eds.), *Philosophy and the Cognitive Sciences*, Proceedings of the XVIth Int. Wittgenstein Symposium, Vienna, Hölder-Pichler-Tempsky, pp. 147-157.
- Fabre, J.-H., ([1879] -1989), *Souvenirs entomologiques*, Paris, Editions Robert Laffont.
- Fodor, J. (1975), *The Language of Thought*, New York: Thomas Y. Crowell, Co. Reprinted by Harvard University Press, 1979.
- Fodor J. (1990), *A Theory of Content and Other Essays*, Cambridge, MIT Press.
- Gallistel, R.C., (1980), *The Organization of Action: A New Synthesis*, Hillsdale, New Jersey.
- Gallistel R.C., (1990) *The Organization of Learning*, Cambridge, MIT Press.
- Godfrey-Smith, P., (1991) Signal, Decision, Action, *Journal of Philosophy*, 88, 709-722.
- Hearst, E. & Jenkins, H.M., (1975), *Sign tracking: The stimulus-reinforcer relation and directed actions*. Austin, Texas: The Psychonomic Society.
- Heyes, C., & Dickinson, A., (1993), The Intentionality of Animal Action, *Mind and Language*, vol. 5, 1, 1990; reprinted in M. Davies & G.W. Humphreys (eds.), *Consciousness*, Oxford, Blackwell.
- Hineline, P.N. & Rachlin, H., Escape and avoidance of shock by pigeons pecking a key, *Journal of Experimental Analysis of Behavior*, 1969, 12, 533-538.

- Holland, P.C., (1979), Differential effects of omission contingencies on various components of Pavlovian appetitive conditioning, *Journal of Experimental Psychology: Animal Behavior Processes*, 5, 178-93.
- Kim J., (1991), Dretske on how Reasons Explain Behavior, in B. McLaughlin (ed.), *Dretske and his Critics*, Oxford, Blackwell.
- Millikan, R. (1984), *Language, Thought, and other Biological Categories*, Cambridge, MIT Press.
- Millikan, R., (1993), *White Queen Psychology and other essays for Alice*, Cambridge, MIT Press.
- Neander, K., (forthcoming), “Pruning the tree of life”, *British Journal for the Philosophy of Science*.
- Nisbett, R.E. & Wilson, T.D., (1977), Telling More Than We Can Know ; Verbal Reports on Mental Processes, *Psychological Review*, 84, 3, 231-259.
- Proust J. (1993), Les rapports de l’esprit et du corps: des interactions entre structure et fonction, in R. Klibansky & D. Pears (eds.) *La philosophie en Europe*, Paris, Gallimard, pp. 641-670.
- Proust, J. (1994), Naturalizing Intentionality through Learning Theory, in R. Casati, B. Smith & G. White, (eds.), *Philosophy and the Cognitive Sciences*, Proceedings of the XVIth Int. Wittgenstein Symposium, Vienna, Hölder-Pichler-Tempsky, pp. 233-245.
- Proust J., (forthcoming), “Descripteurs distaux et externalisme”, Proceedings of the Conference “Esprit, representation, contexte” (Neuchatel 1993), in *Dialectica*.
- Rosenbluth, A., Wiener, N. & Bigelow, J., (1943), Behavior, Purpose and teleology, *Philosophy of Science*, 10, pp.18-24.
- Sober, E. (1984), *The Nature of Selection*, Cambridge, MIT Press.
- Stampe, D. (1990), “Desires as Reasons - Discussion Notes on Fred Dretske’s “Explaining Behavior: reasons in a World of Causes”, *Philosophy and Phenomenological Research*, vol. 1, 4, pp. 787-793.
