The argument:

(*i*) Theories: S1 subserves P1; S2 subserves P2.

(*ii*) Therefore: V4 plus P2 produces V2.

(*iii*) Observations: B1 exists.

(*iv*) B1 subserves B2.

(*v*) Proposition: V4 without B1 produces V3.

(*vi*) Because: V4 without S2 produces V3.

The following part is ambiguous:

Either (Option I)

(*vii*) Therefore: V4 without S2 produces V3 (permanent).

(*viii*) Thence: V4 without S2 plus B1 produces V1.

Or (Option II)

(*vii*) Therefore: V4 plus S2 without B1 produces V3 (permanent).

(*viii*) Thence: V4 plus S2 plus B1 produces V1.

Criticism of the argument:

1. P1 and P2 are not defined, nor is the relationship between them. We do not know how we could recognise either of them, whether they are alternative parallel functions or whether P2 subsumes P1. We do not know the conscious status of P1 or P2.

2. (*ii*) cannot be held to follow from (*i*) without the relationship between V4 and S1 and S2 being specified. For example, if (*ii*) were to follow from (*i*) this would imply (at least) that V4 was brought about by the absence of S1 so that S2 on its own produced V2 (blindsight). This would also imply that attentional vision and blindsight are equivalent. No evidence is offered to support any of these suppositions.

3. The logical status of (*vi*) is unclear. Is it an observation, theory, or mere assumption? (v) cannot follow from (*vi*) unless S2 subserves B1. If it does then this implies that P2 and B1 are equivalent. (*ii*) would therefore imply that V2 (blindsight) is due to B1 (voluntary attention shifts) in conjunction with V4 (scotoma). Blindsight, by definition, is not voluntary attention shifting. It is voluntary in the sense that subjects voluntarily make the movements, but because they are "guessed" movements it is difficult to see how they could be construed as attention shifting. If, by some semantic gymnastics, one did so construe them, then it would be impossible to distinguish between blindsight and normal vision.

4. *Option I*

If S2 subserves B1, then how can V4 without S2 plus B1 exist? If it did, why should it produce V1 and not V2? Such a proposition is incompatible with (*ii*).

Option II

Why should V4 plus S2 plus B1 produce V1 and not V2. Again this is incompatible with (*ii*).

The experimental paradigm. The reaction time paradigm described by Lutzemberger and his coauthors would be a useful way of indicating the knowledge of a stimulus prossessed by a subject for two reasons. First, as they indicate, it is not prone to response strategy changes. Second, it is also not prone to the differential sensitivity problem. We expect that these commentators are right about the minimal light scatter, but a better safeguard, as we indicated in the target article, would be to present stimuli bilaterally in the blind and sighted fields and compare stimuli presented in the blindspot and normal parts of the blind field.

Our only serious quibble is with the conclusions that could be drawn from data showing spatial summation across the hemianopic boundary. It cannot be claimed that this is an unconscious process for the reasons given above. Nor can it be claimed that it is an effect of nonstriate cortex without evidence of lesion localisation. With statistically nonsignificant results and a possible trend towards such summation in only some patients, the extent of the lesions becomes crucial. For example, did the patients showing evidence of summation have only striate damage whereas the others included extrastriate damage (impacts our Proposition 1) or, as is *equally* likely, given our knowledge of their lesions, did these patients have some residual striate cortex (impacts our Proposition 2). The existence of performance variation on its own could thus be taken to either support either of two theories depending on which particular theory one chooses to adopt and depending on the integrity or otherwise of a brain structure for which no evidence is offered.

References

Campion, J., Latto, R. & Smith, Y. M. (1983) Is blindsight an effect of scattered light, spared cortex, and near-threshold vision? *Behavioral and Brain Sciences* 6:423–86. [LL]

Diaper, D. M. (1982) Central backward masking and the two-task paradigm. Unpublished Ph.D. Thesis, Cambridge University, England. [JC]

Dixon, N. F. (1971) *Subliminal perception. The nature of a controversy.* McGraw-Hill. [JC]

Haber, R. N. (1983) The two visual system hypothesis loses a supporter. *Behavioral and Brain Sciences* 6:453–54. [LL]

Marcel, A. J. (1982) Is cortical blindness a problem of visual consciousness or visual function? Paper presented at the Fifth International Neuropsychological Society European Conference, Deauville, France. [LL]

Marzi, C. A. (1983) The neural basis of perceptual equivalence of visual stimuli in the cat. In: *Advances in vertebrate neuroethology*, ed. J. P. Ewert, R. R. Capranica & D. J. Ingle. Plenum Press. [LL]

Pizzamiglio, L., Antonucci, G. & Francia, A. (1984) Response of the cortically blind hemifields to a moving visual scene. *Cortex* 20:89–99. [LL]

Posner, M. I. (1980) Orienting of attention. The VIIth Sir Frederic Bartlett Lecture. *Quarterly Journal of Experimental Psychology* 32:3–25. [LL]

Schmielau, F. & Marzi, C. A. (1983) Mirror symmetric facilitation contralateral to the blind spot in normal subjects. ARVO *Abstracts:* 186. [LL]

Singer, W., Zihl, J. & Pöppel, E. (1977) Subcortical control of visual thresholds in humans: Evidence for modality specific and retinotopically organized mechanisms of selective attention. *Experimental Brain Research* 29:173–90. [LL]

Zihl, J. (1980) "Blindsight": Improvement of visually guided eye movements by systematic practice in patients with cerebral blindness. *Neuropsychologia* 18:71–77. [LL]

Zihl, J. & Werth, R. (1984) Contributions to the study of blindsight – 1. Can stray light account for saccadic localisation in patients with post-geniculate defects? *Neurospychologia* 22:1–11. [JC]

Commentary on **Daniel C. Dennett (1983) Intentional systems in cognitive ethology: The "Panglossian paradigm" defended. BBS 6:343–390.**

Abstract of the original article: Ethologists and others studying animal behavior in a "cognitive" spirit are in need of a descriptive language and method that are neither anachronistically bound by behaviorist scruples nor prematurely committed to *particular*

"information-processing models." Just such an interim descriptive method can be found in *intentional system theory*. The use of intentional system theory is illustrated with the case of the apparently communicative behavior of vervet monkeys. A way of using the theory to generate data – including usable, testable "anecdotal" data – is sketched. The underlying assumptions of this approach can be seen to ally it directly with "adaptationist" theorizing in evolutionary biology, which has recently come under attack from Stephen Gould and Richard Lewontin, who castigate it as the "Panglossian paradigm." Their arguments, which are strongly analogous to B. F. Skinner's arguments against "mentalism," point to certain pitfalls that attend the careless exercise of such "Panglossian" thinking (and rival varieties of thinking as well), but do not constitute a fundamental objection to either adaptationist theorizing or its cousin, intentional system theory.

Aristotle, final cause, and the intentional stance

Aaron Ben-Zeev
Department of Philosophy, University of Haifa, Haifa 31999, Israel

Dennett's (1983) view of explanation and intentionality is similar, in some respects, to Aristotle's view of explanation and the final cause. Indicating the similarity can help us better understand the issues at hand.

Aristotle assumes that everything can be explained in light of four explanatory factors ("causes" as they are inadequately termed): material, formal, efficient, and final. The *material cause* refers to a lower level of description than that of the explanandum, namely, the elements out of which something is made. The *formal cause* refers to a higher level – that into which something is made. This cause refers to the essential functions (or the general regularity) typical of the species in question. Discussing the species is on a higher level than discussing the individual animal. The *efficient cause* is that by which something is done, and the *final cause* is that for the sake of which something is done. Both of the last causes usually refer to the same level of description. The movement of one billiard ball is on the same level of description as the movement of another ball which is its efficient cause. Similarly, the mature tree is on the same level of description as the young tree (the former is the final cause of the latter).

The material and formal causes are not separated in time or space from the explanandum; they are different aspects of it. In the case of the efficient and final causes, it is not the level of description that differentiates them from the explanandum. Rather, they refer to events that are separated – in time or space – from the explanandum. Whereas material and formal causes refer to what constitutes a certain thing, efficient and final causes refer to causes and reasons of a certain process of change. The relations in the former are not causal relations in the usual sense, since there is no actual separation between the cause and the effect. They are, rather, a relation of support (the lower-level entities support the higher-level properties) or of realization (the higher-level properties are realized in the lower-level entities). The efficient cause is identified with what modern science usually regards as a cause. The scientific notion of causal relations refers to relations within a certain level of description.

The above general framework is quite suitable for describing mental states. Thus one's state of fear may be described by referring to (a) a stimulation of the physiological mechanism (material cause – a reference to a lower level of description); (b) a particular mental event, for example, imagining one's boss (efficient cause – similar level), or one's future promotion (final cause – similar level); (c) the general character (essence) of the agent (formal cause – higher level). In light of this analysis, physiological entities are not efficient causes for mental states but rather a supportive basis. Mental states are realized in physiological processes (Ben-Zeev 1983). Whereas the relation of causality indicates spatial or temporal separation between the cause and effect, the relations of support and realization imply a reference to two different aspects of the same thing. Causality, in its strictest sense, is a relation typical of the same level of description.

The above Aristotelian contentions can solve quite a few problems in the philosophy of mind. They are also close to Dennett's approach, for example, to his insistence upon having different types and levels of explanation (e.g. his denial that mechanistic [or causal] explanations take priority over, and render false, intentional explanation), and to his attempt to find some sort of continuity between human and nonhuman behavior. The similarity between the two approaches is even deeper when we compare Aristotle's notion of final cause with Dennett's intentionality.

Aristotle's final cause ("for the sake of") has two senses – the purpose for which something is done (the desired object) and the beneficiary for whom something is done. For our discussion the former sense is more relevant. It denotes that which the activity is directed at. Intentionality has a similar meaning: It is the feature of being directed at. If Aristotle is right in assuming that everything can be described by referring to a final cause, then in a certain sense everything may be described in light of an intentional stance. Indeed, Aristotle ascribes intentional states – such as desires – to inanimate things too. For instance, he describes the fall of a stone as expressing its desire to reach its natural place. Dennett seems to take a similar stand when he ascribes intentional stances to chess-playing computers, thermostats, and lightning bolts (see also Lloyd 1983).

Indeed, it seems that each phenomenon can be explained by referring to at least three levels of description; lower, similar, and higher. But can we attribute a final cause or an intentional stance to each phenomenon? Yes and no. If by final cause we merely refer to a future state of the system (the state at which it will probably arrive in light of its regularity), then each phenomenon is directed, in a sense, at that future state. If intentional denotes merely this minimal sense, then it can be ascribed to computers, thermostats, and lightning bolts. But the more interesting use of intentionality is that which refers to a special relation of "being directed at." That relation is typical of mental states found in human beings. (The New Science bitterly opposed Aristotle's use of this relation in connection with inanimate things. It has been suggested that this opposition is an important factor in the rapid advancement of the New Science.)

Dennett ascribes to the above second notion of an intentional *state* one typical feature: Its description exhibits referential opacity. He also ascribes two features to the intentional *system:* It is rational and it has different grades.

The feature of referential opacity is indeed absent from other kinds of "being directed at." Thus the following description is adequate.

1. The thermostat is set at 20°C.
2. The most comfortable temperature for humans is 20°C.

And, contrary to the feature of referential opacity, we may conclude here that

3. The thermostat is set at the most comfortable temperature for humans.

Hence, contrary to Dennett, thermostats – and other inanimate things – do not have intentional states (in the above second sense).

Referential opacity stems, I believe, from some cognitive capacity of the agent. By this capacity the agent is able to know its object through various aspects. I may, for instance, conceive of the thing in front of me as a chair, a comfortable sitting place, a

favorite piece of furniture, and so forth. The agent may know only some, but not all, of these possible aspects. This partial knowledge is the basis of the referential opacity. Knowing that the chair is a comfortable sitting place does not entail knowing that it is my favorite piece of furniture as well. It is plausible to assume that the cognitive capacity of animals has this feature of "knowing x as y." Hence, we may ascribe to them intentional states too.

The explanatory factor of a final cause is best used, I believe, in connection with some cognitive capacity. Only agents who are able to know circumstances beyond the present ones can be described as believing in and desiring those circumstances. Other kinds of systems merely tend toward (in the minimal sense of being directed at) such circumstances.

The two features Dennett ascribes to intentional systems are connected with cognitive capacities as well. The different grades of intentional systems stand for different grades of cognitive capacity. Take, for example, Dennett's first and second order of intentional systems:

1. X believes that P.
2. X believes that Y expects X to believe that P.

The second-order intentional system exhibits more complex knowledge than the first-order system. The more possible aspects the system is able to grasp, the higher its intentional order is. The grade of intentional order is indeed (as Dennett suggests) a good criterion for distinguishing human intentionality from that of animals.

That kind of distinction is close to Wittgenstein's contention that though animals can go beyond the present circumstances, they do it in a much less determined manner: "We say a dog is afraid his master will beat him; but not, he is afraid his master will beat him to-morrow. Why not?" (Wittgenstein 1958, p. 650). In a somewhat similar vein, Aristotle indicates that the difference between lower and higher animals' imagination concerns its determinacy. The imagination of lower animals is indeterminate (*De Anima*, 433b31f).

The feature of rationality that Dennett ascribes to intentional systems is problematic, since it is not clear what rationality means. If rationality merely means a certain degree of cognitive capacity (and the ability to act accordingly), then this feature does not add much to what has been characterized so far. If rationality means something else, then I am not sure I can see its connection to intentionality.

An intentional stance involves the agent's ability to refer to various possible aspects of the object. This involves a complex cognitive capacity which enables the agent to take into consideration possible future events. Consequently, there are both the past and future dimensions for describing the activity of an agent with such a capacity. In a somewhat similar manner, Aristotle's explanatory factors of efficient and final causes refer to those two temporal dimensions (the dimension of the present is expressed in the material and formal causes). However, unlike Dennett (and Aristotle), I believe that the reference to the intentional stance (or the final cause) is meaningful only in regard to agents with some cognitive capacities. The greater the complexity of those capacities, the more relevant and suitable are the considerations concerning the intentional stance (and the final cause).

The intentional stance reexamined

Radu J. Bogdan

Department of Philosophy, Tulane University, New Orleans, La. 70115

Dennett (1983) takes the intentional stance to be a provisional and heuristic strategy of asking questions about mind and behavior, the right answers to which will eventually be given by nonintentional hardware and design theorizing. The intentional stance, then, is more like a preliminary and rough map of a territory whose final mapping is bound to be quite different. A basic assumption Dennett is making is that (i) the intentional and nonintentional maps share some (mental) territory which (ii) the final, nonintentional map of the mind will reconceptualize and explain in its own terms, definitively. The assumption need not be granted. The rest of this commentary explains why.

First, a trivial point: Some ancient attempts notwithstanding, not everything under the sun is worth explaining intentionally. Even for the imaginative and fearful cave people, once a phenomenon acquired a safe and familiar regularity, it opened up the possibility of a nonintentional story. The same is true of psychological matters: Not everything cognitive or behavioral is worth explaining intentionally. Tropisms, reflexes, instincts, any rigidly patterned processes in general, appear to be unlikely candidates for intentional explanation. This is a documentable fact. What it shows is that, far from being a mere heuristic ploy, as Dennett often maintains, the intentional stance manifests ontological sensitivity, as it were, to certain aspects of the mental. What could these aspects be?

Regularity, or rather the lack of it, has emerged as a distinguishing mark. But this is much too vague. Dennett's insightful comparison between intentionalism and adaptationism can help in making the story more specific. Consider an adaptive development, a creative response by the species to novel challenges. Two aspects of it are relevant here. One is that, from an exclusively internalist angle on the species' hardware and design, an adaptive development necessarily appears as *improvisation*, that is, one that is unexplainable as a development, because underdetermined, by hardware and design laws. To understand such an improvisation and its effects, one cannot rely only on knowledge of the species' hardware and (past) design. This brings us to the second aspect. To identify an adaptive development and explain it, to capture the factors and regularities, if any, responsible for it, one must go beyond hardware and design and thus abandon (what I have called) the internalist approach. One must also consider facts and regularities of the environment, of its interaction with hardware, and so forth. This is because the adaptive development we try to understand is, in some sense, sensitive to, and has a raison d'être in, aspects that transcend the internal territory of hardware and design. As an explanatory strategy, then, adaptationism seems to be needed, in one form or another, when nature, both outside and then inside a species of organisms, breaks some previous laws, and when, as a consequence, the species has to improvise in order to figure out and adjust to the new laws of the game.

The human mind is to a large extent adaptive. With respect to content, thoughts, beliefs, or plans appear as improvisations with information *if* looked at exclusively from the angle of their hardware and design instantiations, with their principled features and laws. The intentional stance is after cognitive contents. That stance seems to work best when, tracking contents, it characterizes those relevant aspects of mind and behavior that are neither hard-wired nor rigidly programmed, in other words, aspects that cannot be fully explained internally. These are the aspects that, in content, characterize the *central* states and functions of the mind, such as beliefs, thoughts, and the like. Consider, for example, a configuration of beliefs. It is not hard-wired, for there is no unique hardware realization or localization for beliefs, and it is not rigidly programmed, for a human mind displays no unique doxastic design. It follows that there are no internal laws for beliefs, for it takes hardware and design constraints to instantiate internal lawfulness, but, as just shown, there are no such constraints that can uniquely fix and hence explain a configuration of beliefs. Or consider the information carried by a thought verbally articulated. No program constraints, for example, whether grammatical, conceptual, or otherwise computational, can uniquely determine that information, for the same information is compatible with several permissible program or formal structures, just as one and the

same formal structure may, in different contexts, carry different units of information.

If so much is true of the central mind, then intentionalism, as an explanatory strategy, indeed has something in common with adaptationism. If it is to figure out the contents of central mental states, in order to explain from them, the intentional stance must go beyond the hardware and design facts and laws and look for some other external facts and laws. It is this latter territory (of whatever central states are sensitive to, whether causal interactions, contexts, social practices and conventions, or something else) that the intentional stance attempts to map. Its raison d'être, then, is more than just heuristic, it is ontological. This does not mean that the cognitive reality of thoughts or beliefs is more than that of hardware states governed by some design strictures. It only means that the design (in particular, the program) capabilities are such that neither hardware laws nor program strictures can account for what a central output (thought, belief, utterance) is about and hence it cannot account for what role that output plays in cognition and behavior.

But, then, obviously, the intentional and nonintentional stances do not map the same territory, not essentially anyway. Their concerns overlap only to the extent to which, say, a belief or thought is necessarily matter (hardware) as well as form (program) and content (information); but whereas the nonintentional disciplines are essentially interested in matter and form, the intentional approach cares only for content, not any sort of content, only the central sort. So the first half of Dennett's assumption, that is (i), does not look good, unless of course there is some way to bridge the gap between content and form, which is what the second half, that is (ii), is all about. Yet I do not see how this gap can be bridged. This is a problem that the computationalist view of cognition, which *is* a design view, also faces, as I have argued elsewhere (Bogdan 1983). A nonintentional analysis can easily read noncentral content off its forms because the design of the latter and often their very hardware uniquely fix the information those forms can carry. This is how scientists look at what peripheral transducers have access to and process. They do not need the intentional stance to figure that out. The same is true, as Dennett suggests, for bees and countless other species of cognitive systems.

The problematic gap, as I see it, is between *central* content and *its* forms. At one point, Dennett writes that the intentional characterization of a cognitive system provides guides or landmarks about a system's representational competence which the design theorist can then reformulate and explain nonintentionally. This is probably what often does and will happen in the study of cognition. However, one should not read too much into this. Any content, since form, can be *described* in a design, hence nonintentional vocabulary. But can it be similarly *explained*? I wonder. Explanation presupposes some generic features as well as regularities, and neither Dennett nor anyone else, as far as I know, has convincingly shown that the generic features and the regularities of central content, if any, can be matched or absorbed by those of design. One possible confusion should be avoided at this point. The design stance can redescribe what the intentional stance identifies. If that were all the commerce between the two stances, then indeed the intentional stance would play only a heuristic role and the design stance could fully accommodate what is thus heuristically suggested to it. But the intentional stance does more than that. It identifies central contents in ways that are conducive to explanation and rationalization. It is this latter task, I maintain, that the design theorist cannot handle.

So we end up with a paradox. On the one hand, the intentional stance does the heuristic work of bridging the content–form gap precisely when it is least needed, that is, when the content is noncentral and is thus uniquely fixed by its form. On the other hand, the intentional stance cannot discharge its assigned heuristic task precisely when most needed, that is, when there is no obvious (I mean: internal) way from form to content. This does not mean that the intentional stance does not tell the design theorist something about the competence a system has. It does. It can tell him what *sort* of competence to expect. Specifically, it can tell him that if intentional attributions and explanations look essential, then there are good chances that the underlying competence is very complex and has program features that ensure that the contents they make possible are, for example, highly flexible, multiply embedded in various cognitive attitudes (the iterations Dennett is talking about), mentioned as well as used, and so forth. This in turn will be an indication that the resulting contents cannot be uniquely determined by the internal laws of design. Among the program features with such effects on content we may list generative power, metarepresentability, quotation and so on [for more general discussion, see Stabler: "How Are Grammars Represented?" *BBS* 6(3) 1983].

This will be similar to the predicament of a biologist who discovers a species programmed for continuous adaptation. Knowledge of that program would tell the biologist what it is possible for the species to do but not what it does and why, given specific circumstances, just as knowledge of the grammatical competence of a speaker would tell the linguist (qua design theorist) what forms are possible but not what informational content they carry in particular speech situations. The central mind may well behave like an organism designed for continuous adaptation.

Dennett on cognitive ethology: A broader view

Bo Dahlbom

Department for the Theory and Philosophy of Science, Umeå University, 901 87 Umeå, Sweden

1. Romanticism is the fashion in science these days. Deep-rooted ideas central to our Enlightenment tradition are being questioned. There is a trend away from atomism, empiricism, functionalism (adaptationism), and gradualism toward holism, innatism, structuralism, and saltationism. People like Chomsky, Fodor, Neisser, Kuhn, and Bohm have paved the way for this Romantic vogue by attacking the old fashion for its "emptiness," "simplifications," and the like, using pretty much the standard method of introducing and selling a new fashion. Chomsky's critical review of Skinner's *Verbal Behavior* (Chomsky 1959) was an admirable, early press release announcing the new trend.

The recent commotion in evolutionary theory brought about by Eldredge, Gould, Lewontin, Stanley, and others (Eldredge & Gould 1972; Gould & Eldredge 1977; Gould & Lewontin 1979) is only another instance of this Romantic trend. Gradualism, atomism, and adaptationism, pillars of the synthetic theory of evolution (and of the Enlightenment), are being questioned. Saltationism (punctuated equilibria rather than gradual phyletic evolution), holism ("Baupläne" rather than atomic traits), and structuralism (attention to the structures of organisms in addition to the operation of natural selection) are defended. No one can fail to see the work of fashion in this trend – the similarities of ideas sweeping through linguistics, philosophy of science, psychology, physics, and biology, to name a few disciplines, and the similarities of the attacks against the old fashion. So why does Dennett choose to place Gould and Lewontin in a camp with Skinner, of all people, rather than with Chomsky, Kuhn, and the like, where they clearly belong? [See *BBS* special issue on the work of B. F. Skinner, *BBS* 7(4) 1984.]

Dennett (1983) casts suspicion on the criticism of adaptationism by comparing it to Skinner's critique of mentalistic psychology. I agree that there are similarities here, but they are based on the kind of properties we always find when the proponent of a research approach argues against a rival approach. Such a comparison of styles of expression does not

cast any more doubt on the substance of Gould's and Lewontin's alternative to adaptationism than a similar comparison would on Chomsky's alternative to Skinner's theory of verbal behavior. It only gets in the way of appreciating an important element – the Romantic trend – in our recent intellectual history. Identifying this trend, Dennett might have realized that this is not the time to try to sell adaptationist programs to biology.

Dennett follows adherents of the synthetic theory in protesting that Gould and Lewontin (1979) attack a straw man. There are no adaptationists in Gould's and Lewontin's sense, and there have never been any. But this response expresses a misunderstanding, I think, of the way science works. As in all changes of fashion, the change proposed by Gould and Lewontin is not so much the result of new evidence or new ideas. It is rather a question of what evidence and what ideas to consider important. It is one thing to admit (and pay lip service to) the existence of "constraints" in the evolution of organisms, and something completely different to make it a central notion in one's research program.

2. What is the advantage to ethology of speaking, as Dennett suggests, about intentional systems rather than about problem-solving systems? Viewing organisms (populations, ecosystems) as problem-solving systems would provide us with the kind of "interim descriptive method" Dennett thinks needed. Furthermore, this language is already available in ethology, being pretty much a standard way of speaking in evolutionary theory.

There are some obvious similarities between the standard problem-solving approach and Dennett's intentional system theory. Both rely on a notion of information in need of further analysis (in spite of Dretske 1981; and its multiple review in *BBS;* Dretske 1983). Both make use of the notion of intentionality as understood by Husserl: Problem solvers and intentional systems view the world from a point of view. This means that the information available to the system is restricted by location, cognitive capacities, interests, and the like. These similarities can be taken to indicate that a problem-solving approach is just what Dennett calls "the intentional stance" under a different name. This is even suggested by Dennett, I think (1983, p. 349), but this is not an answer he should be happy with. For then one would want to know what is new about Dennett's offer to ethology.

There are also differences between the two approaches, however, but they do not seem to favor intentional system theory. The language of that theory is our commonsense intentional idiom. A problem-solving approach is not committed to this idiom, but may very well choose to use it. This option of choice is no disadvantage, considering the unanalyzed complexities of our commonsense idiom. Intentional system theory makes us attend to differences in intentional order in a way that the problem-solving approach does not. But I am skeptical (as are several earlier commentators) about the value of this attention in the field of ethology. I doubt that ethologists will find much use for these notions, except perhaps when dealing with monkeys and a few other species. The cases that will benefit from a use of these notions need not be pointed out, but can be handled in terms of differences in available information within a problem-solving approach.

3. The intentional stance provides characterizations of organisms (or systems in general) that assume rationality. This means that when we adopt the intentional stance toward a chess-playing computer (Dennett 1971; 1978a), for example, we can forget about its physical hardware and design, and simply ask what would be the rational move for anyone to make in the given situation. Indeed, the beauty of the intentional stance is that we need not know anything about the hardware or the design to predict the system's behavior. The concept of rationality relied upon is that of the Enlightenment. To be rational is to be unprejudiced and "logical." But as the Romantic tradition has tried to teach us, we cannot avoid prejudging. Attributing rationality to someone typically means assuming that he operates under the same constraints we do. The intentional stance works well when there is common ground for understanding. But it will fail when such common ground is lacking (a bitter lesson anthropologists have had to learn).[1] And when the intentional stance fails the "Enlightened" person is apt to deny rationality, to "demote" the organism as Dennett puts it. But this is a dangerous practice as exemplified by the standard Enlightenment reaction to foreign cultures. Since the American Indians (to name only one example) did not share our culture, we Europeans thought they were animals.

Viewing animals as intentional systems means viewing them as persons. This is the result of using our folk psychological idiom of beliefs, desires, and the like in talking about intentional systems. But thinking of intentional systems as persons makes it difficult to handle the cases when an individual acts as part of a larger "superpersonal" intentional system, such as a society or a social institution. Dennett hesitates to acknowledge the existence of such superpersonal intentional systems, preferring to speak vaguely about "free-floating rationales" (Dennett 1983, p. 351, but compare his 1981a). We owe our understanding of bureaucratic organizations to Weber (1964; 1968) and his theory of rationality. Viewing bureaucracies as intentional systems makes us realize that the behavior of the bureaucrat cannot be understood unless we understand the system of which he is a part. In our culture calling a person a "bureaucrat" is to demote him. We find him irrational because we cannot appreciate the rules ("constraints") he follows. Paradoxically, it is the rationality of the bureaucratic system that makes the individual belonging to it seem "irrational." Rather than demoting insects like *Sphex ichneumoneus* (digger wasps) for their instinctual behavior, perhaps we should think of them as bureaucrats. If animals are bureaucrats, we should not deny them intentionality, just because they seem irrational. As members of superpersonal intentional systems their behavior cannot be understood, nor can their intentional status be estimated unless these systems have been identified.

Dennett attributes the predictive success of the intentional stance to the fact that systems (persons, computers, organisms) are rational in the Enlightenment sense. In view of Weber's theory, I think the sucess of the intentional stance owes more to the fact that systems are partly habitual ("traditional" in Weber's sense) and partly follow a system of superpersonal rules (being "rational" in Weber's sense). But then the intentional stance works only if we have correctly identified the habits and rules in question.[2]

4. There is a general theory of problem solving incorporating the Enlightenment concept of rationality, discussed and defended by Dennett (1975; 1978a). The important features of this theory – in its extreme version – are that attempts at solution are chosen without prejudice (randomly), and that the problem is solved in a number of small steps (gradually). We find this theory exemplified in such seemingly diverse quarters as Thorndike's theory of learning by trial and error (the law of effect); empiricist methodology, including Popper's falsificationist program and his idea of piecemeal social engineering; classical liberal economic theory; and the synthetic theory of evolution. We might call it "the Enlightenment theory," or in this context perhaps better "the adaptationist theory of problem solving." Intentional system theory is, through its assumption of rationality (in the previously indicated sense), committed to this theory, to adaptationism as Dennett notes.

The Romantic alternative attends more to the discovery side of problem solving than to the justification side, stressing the process of choosing plausible solutions from a background knowledge about the problem situation, defining the problem in terms of that background knowledge. Hegel's theory of cultural development, Marx's theory of historical materialism, cognitive theories of problem solving and perception (Neisser 1967, etc.), and modern theories of scientific growth (Kuhn 1962, etc.), all exemplify this alternative. Its central features are that attempted solutions are prejudiced (constrained) by the background, and for this very reason often constitute structurally

motivated, substantial "leaps" toward a solution of the problem. Thus, the Romantic theory corrects the gradualism and adaptationism of the Enlightenment by introducing saltational and structural elements.

In its allegiance to the Enlightenment notion of rationality, intentional system theory is committed to gradualism and adaptationism. A problem-solving approach, however, is neutral on these issues, and can be used by adherents of either of the theories outlined, constituting a basis of agreement from which to argue. Offering intentional system theory as a "language of description" for ethology, Dennett must answer the criticism leveled against the synthetic theory by Gould and Lewontin. Admitting the force of their arguments, he is willing to introduce "constraints" into his theoretical approach, but he fails to see, or is unwilling to accept, the move whereby such "constraints" take on a central role and call into question the basis of his whole theory. Just pause to think how different Dennett's theory would be if "constraint" rather than "rationality" were its central notion.

5. Danto (1983) proclaims the autonomy of intentional system theory as a theory. But Dennett is more modest. He does not endorse reductive physicalism, as seen from what he has to say in response to Churchland (1983). But he is still a physicalist of sorts, as indicated by his claim that intentional stance predictions "have no predictive *hegemony* over design stance or physical stance predictions." In a tug of war between the physical (or design) stance and the intentional stance concerning what a person believes or desires, the intentional stance must always give. The physical stance is treated as the "infallible protocol" against which to test intentional attributions.[3] I fail to see why intentional attributions may not sometimes be firm enough to make us revise our physical characterization. In denying this, Dennett shows an unwillingness to take the last step away from reductionism, introducing intentional system theory as an equal citizen beside physics in the land of science. Why be so modest?

Bennett (1983) asks for clarification of the fundamental concepts of intentional system theory. Dennett interprets this (tentatively) as meaning the formulation of "flow charts and systems of rules to be followed by (but not necessarily represented in) the organism." (Why?) And he says that if this is what Bennett has in mind, he would "heartily concur." But the elaboration of flow charts would spell disaster for Dennett's program. Such a development of the theory could at this stage only be speculative and arbitrary, and would make intentional system theory share the fate of what is called "general systems theory." In the hands of Ashby, Wiener, and others, this theory was a fruitful and illuminating way of isolating a few interesting general properties of "systems." It soon lost our interest, however, as the result of the addition of ever more unfruitful and boring complexities. Who among us, except the system theorists themselves, care about this theory today? Were Dennett to start drawing flow charts he would end up ruining his program in the same way general systems theory has been ruined by silly drawings of boxes with interconnecting arrows.

To be successful, Dennett's enterprise will have to steer clear both of general information-processing models of the general systems theory variety and of the particular information-processing models Dennett warns against. It should model itself on the way historians and anthropologists study cultures rather than on the way psychology concocts models of information processing. It should be "idiographic" rather than "nomothetic." And it could use "the Sherlock Holmes method" as suggested by Dennett. In recent discussions of the methodology of history and anthropology, this method has received growing attention (Ginzburg 1980). Since experiments are relatively uninteresting to historians, this discussion has centered on Holmes's proficiency in locating and interpreting "clues," of finding the telltale, superficially unimportant signs that reveal the character of a person or the structure of a crime. Historians will have to look for clues where they can find them. Anthropologists and ethologists are in a position to provoke them in the way described by Dennett.

6. When ethology turns cognitive, one would expect it to start asking serious questions concerning the "cultures" of animal species. What Dennett in effect offers in his program for ethology is a skeleton version of the methods of functionalist anthropology. But Lévi-Strauss (1955; 1966) has argued that functionalism tends to ignore the fact that thinking, wherever it takes place, is constrained by social structures. To learn about a foreign culture we should study its kinship systems, its productive systems, its rituals, and so on, and be particularly observant of the way these systems are being upheld by concrete structures of living.

It is not too difficult, I think, to find analogies between human cultures, and conceived by Lévi-Strauss, and the structures of other species' ways of life. Lévi-Strauss's insistence on the concrete nature of cultures makes such a move plausible. Once the notion of culture is decoupled from the notion of consciousness, it should be possible to extend it in ways interesting to ethology. Such a study of cultures is "holistic." It sees a culture as a way of life built around a few central principles expressing themselves in different ways in the various activities of the population. In its holistic nature the structuralist research approach differs from functionalism. The move from functionalism to structuralism in anthropology is another example of the recent trend away from the Enlightenment to Romantic ideas.[4]

NOTES

1. Dennett's little fairy tale opening his excellent paper on the frame problem (Dennett 1984) is in its way a good illustration of this. There Dennett shows how difficult it is to program a (perfectly rational) robot to do exactly what we want it to do. It is difficult simply because it is so difficult to say exactly what we want it to do. With people we know, it is easy because we can rely on a background of shared (physical, cultural, moral) constraints. We can assume they are "rational." But robots are not like the people we know. And neither are animals.

2. Ghomsky (1959) showed that Skinner's attempts to predict verbal behavior in terms of operant conditioning and reinforcement presuppose severe constraints on the alternative verbal responses available, constraints that Skinner takes for granted without discussion. In Dennett's (1983) response to Skinner's (1983) commentary he gives us a version of this kind of criticism. But isn't the intentional stance lacking in precision in the same way as Skinner's theory? It works only by relying on a shared background of (unidentified) constraints.

3. Thus Dennett is unwilling to depart too far from the position attributed to him by Rosenberg (1983). But wasn't Dennett's (1969) position more complex and subtle than that? Rosenberg quotes from section 10, pp. 85–86. Does not the concluding paragraph of that section on pp. 89–90 indicate a more benevolent attitude to the autonomy of the intentional stance?

4. If there is anything at all in the idea of a cognitive ethology besides a fancy name, ethologists will have a lot to gain, I think, by taking sociology and anthropology much more seriously than they have done as yet. So far they have read Lévi-Strauss (1969) only to denounce him (van den Berghe 1983; but see Bateson 1983a; 1983b for a more generous attitude).

Beyond Burrhus and behaviorism: Dennett defused

Thomas Gray

Department of Psychology, Concordia University, Montreal, Quebec, Canada H3G 1M8

Dennett's (1983) target article was addressed to cognitive ethologists, but I think he will forgive a psychologist from butting in with a few remarks.

Dennett seems earnestly concerned with finding good ways to explain behavior. As he demonstrated in *Brainstorms* (Dennett 1978a), he is willing to take into account contributions from even *physiological* psychology. His chapter called "Skinner Skinned," where he continued flogging a retired horse was a

disappointment, but much of the book was sensible. In *Brainstorms*, and in the *BBS* article, I detect a kindred spirit. For one thing he too has actually read Gould and Lewontin's (1979) "Spandrels of San Marco" article. The main problem, I think, is that Dennett's friends in psychology have somehow let him down.

Dennett, who still seems intent on tilting at Behaviorism (capital B variety), is using an outdated cast of characters, and his stage directions need some editing.

Perhaps it is time to banish "behaviorist" and related terms. The label is too closely associated with muscle-twitchism to be useful, especially in interdisciplinary journals such as *BBS*. (The word should certainly not be used in the *New York Review of Books* because of the apoplexy it elicits in the Humanists.) Dennett should also be reminded that there are many other "behaviorists" besides Skinner. Hull, Tolman, Thorndike, and Guthrie do get their names in the index of *Brainstorms*, but, strange as it might seem given Dennett's concern to expose the errors of empty-organism theorizing, you will search in vain for a reference to Hebb.

As I flipped back and forth between Dennett's *BBS* article and the book containing his earlier brainstorms I got the funny feeling that I was a crypto-intentionalist. Had I been assuming something equivalent to an intentional stance all along? Shades of Moliere's Monsieur Jourdain!

I wonder what Dennett would have contributed to the present-day attempts to understand behavior had he chosen to dissect Hullian or Tolmanian behaviorism? Could he have helped us in our tottering attempts at manipulating intervening variables and hypothetical constructs? Could he have aided us – as we sought to derive organizing power from these "temporary formulations" – to avoid their "gratuitous and incautious overextension"? One of his psychological colleagues should have suggested that he take more seriously what some post-Hullians have had to say about these issues. Our use of Hebb-like Conceptual Nervous System terminology, with its enormously influential cell-assembly concept, might have been much more refined with Dennett's help. Dennett might also have taken Bindra more seriously (see Bindra, especially 1976 and 1978). If only Dennett had taken a look at some of the newer (would you believe 20 years already?) approaches to conditioned associations he would have found some behaviorists writing reports about "surprise" and "selective attention" in rats (see, e.g., Kamin 1969). (This latter behaviorist is defending the freedom and dignity of potential victims of overzealous hereditarians. Lewontin, 1981a, p. 13, has in fact praised him for his "brilliant muckraking in the byres of IQ studies."

To ask researchers, such as some of my colleagues, who are busy finding out the *nature of the mechanisms* that make beliefs, wants, desires, and rationality possible, to now adopt Dennett's intentional stance would be like asking biologists to consider how wonderfully useful the gene concept would be on the day after Watson and Crick (1953) announced to the scientific community that "it has not escaped our notice . . ." It would be like Hofstadter's (1979) Tortoise and associates coming up with the pinball-machine metaphor of the brain while in the next room Kandel or Alkon was describing what happens in the synapses and cell membranes of systems of neurons as they are modified by experience (see, e.g., Carew, Walters & Kandel 1981; or Farley & Alkon 1980). I know there is a long way to go, but at least we have found out that the nervous system is not made of components available in Radio Shack, and we know it is not neatly laid out in diamond shapes, squares, and ellipses with lines connecting them.

I don't want to be misunderstood here, and I particularly don't want to be mistaken for a *vulgar*, nonromantic reductionist. I am concerned, though, that crude reductionists are seen lurking behind trees (perhaps with the Behaviorists) ready to take all the fun and "romance" out of life. In a book review of Lumsden and Wilson's *Genes, Mind, and Culture* (1981) (not on my list of favorite books), Lenwontin, for example, is pleased that "reductionist science" has failed to explain complex nervous systems. Nor, he adds, "has cognitive function been reduced to a collection of single neuron firings" (Lewontin 1981b, p. 23). Where are these researchers who hold such a primitive view of neurophysiological explanation? [See also *BBS* multiple book review of Lumsden & Wilson's *Genes, Mind and Culture BBS* 5(1) 1983.]

Dennett is dejected that all the romance will be taken out of life. He seems saddened that the bees' body-removal behavior is *just* a response to the smell of oleic acid. Many people think it's a marvel. Would he languish in existential despair if it were shown that thinking is *just* a function of exquisitely modulated neuro-bio-chemo-hormo-physico-mechanico-uncle-Tom-Cobbleigho responses? Does he really think the pleasure I get from the taste of chocolate-chip and raisin cookies will go away if he (or I) understand the basis of the response? Leonardo's painstaking analyses of muscle attachments did not, it appears, make it impossible for him to capture the spirit and grace of animal movement. Does Dennett really believe that we should take seriously those people (if there are any) "who suppose that since colour can be explained in terms of the properties of atoms which are not coloured, nothing is coloured"? (Dennett 1978a, p. 65).

Take heart, I hear there is both charm and color at the subatomic level.

Author's Response

When does the intentional stance work?

Daniel C. Dennett
Philosophy Department, Tufts University, Medford, Mass. 02155

Ben-Zeev raises the issue of referential opacity – the failure of substitution of codesignative terms *salva veritate* – and claims that the application of the intentional stance to thermostats (an example of mine) doesn't exhibit it. This shows, he thinks, that the use of the intentional stance in application to such things as thermostats is fundamentally different from its use in application to things – people and at least some creatures, presumably – that have genuine intentional states. This would leave the concept of an intentional state in need of some other account than mine, and Ben-Zeev proposes that "the reference to the intentional stance . . . is meaningful only in regard to agents with some cognitive capacities. The greater the complexity of those capacities, the more relevant and suitable are the considerations concerning the intentional stance."

But this is not an alternative account; it is my account obscured and weakened by its reliance on an unanalyzed notion of an agent with a cognitive capacity. **Ben-Zeev's** point about substitution failure is important, but slightly misses the mark. Finding a case where substitution works is not the issue; after all, in many cases of attributing intentional states to human beings, one may substitute codesignative terms *salva veritate*. The proper question is: Does explanation of the behavior of thermostats (or bees or birds) *ever* require attributions that resist substitution? And here the answer is interestingly indirect. Thermostats are *virtually* "oblivious" to substitution, but not quite, and we can easily imagine more

sensitive (wily) thermostats, attributions to which were clearly sensitive to substitution. But thermostat designers have not felt the need for such sophisticated devices for the most part. Similarly, consider the bee, which surely does not need to recognize or distinguish her oleic-acid-exuding dead sister qua health hazard, or even qua corpse. The bee has a very minimal "cognitive capacity" – to use Ben-Zeev's term. But when we go to explain why this phenomenon exists in nature, why bees should have this proclivity built in, our explanation will single out the dead bee under the marked description; it was qua *health hazard,* and not qua anything else, that dead bees were "recognized" by the evolutionary process itself (Mother Nature). The rationale of the behavior (if not the individual bee's rationale, then a free-floating rationale) is nevertheless expressible only in the referentially opaque language of intentional explanation. So whatever a "cognitive capacity" is, if its presence is marked by an appeal to referentially opaque explication, then natural selection is itself an "agent with cognitive capacities."

As **Bogdan** says, "the adaptive development we try to understand is, in some sense, sensitive to, and has a raison d'être in, aspects that transcend the internal territory of hardware and design."

Ben-Zeev's assertions that my view is "close to" Wittgenstein, and "similar, in some respects," to Aristotle, must be true, of course. With a sufficiently relaxed standard of similarity, affinities can be found between the views of almost any two philosophers on any subject – Hegel and Aristotle, say, or Sartre and Quine. (I have long yearned to write the rather obvious paper entitled "How Sartre's 'transparency' is just Quine's 'opacity.'") In this instance I do not see anything particularly striking or useful or worth quarreling about in the comparisons, so I will resist the temptation to "compare and contrast" as they say in final exam essay questions.

Bogdan makes the point that I think gets obscured by **Ben-Zeev**'s proposal to define intentionality in terms of cognitive states (and not vice versa). What makes something a (central) cognitive capacity or contentful state is that it is "unexplainable, . . . because underdetermined, by hardware and design laws." That is not to say that it is not *in principle* fully predictable by, say, a Laplacean ominiscient scientist working with nothing but "hardware and design laws," but that any such nonintentional (mechanistic, atomistic, local) explanation would miss something important: that peculiar relatedness to remote conditions, real or implied, that is most familiarly recognized as *aboutness* – what philosophers call intentionality. It almost looks like "action at a distance," but of course it is not. The indirect bearing of the Eiffel Tower on my thought about the Eiffel Tower, like the indirect bearing of the toxicity of those ancestral bee corpses that weren't removed from their hives on the current behavior vis-à-vis corpses of today's bees, is not the sort of relation that can be illuminated by a mechanistic, nonintentional account, however voluminous.

Bogdan sees as an implication of this that "the intentional stance cannot discharge its assigned heuristic task precisely when most needed, that is, when there is no obvious (I mean: internal) way from form to content." That is, for the best truly "central" cases of content, where the intentional stance is our only grip on the phenomenon, we cannot expect the intentional stance to point out the path to its own elimination in favor of design stance accounts. I guess that is true, for both psychology and evolutionary theory, and it amounts to a very mild sort of "irreducibility" claim. Not that mind is irreducible to brain, or that intentionality is inexplicable in mechanical terms, or that adaptation cannot be the result of (nothing but) evolution by natural selection (and genetic drift and other clearly mechanical processes), but just that the only sense we will ever be able to make of the play and interaction of "central" intentional states will be the explanations we make from the intentional stance. Other accounts may be true, and predictive, but won't explain everything that needs explaining.

Dahlbom offers six numbered points, to which I will respond in turn.

1. His account of the current trend away from Enlightenment values toward a Romantic vision of science provides a novel, valuable and, I believe, largely correct perspective on contemporary controversies. He is right that I am unfashionably bucking the Romantic trend, but then I have long been branded a "verificationist," "reductionist" opponent of Chomsky and Fodor, for instance, so my sympathies should not surprise Dahlbom. But why, he asks, do I willfully place Gould and Lewontin, arch-Romantics, with Skinner, the embodiment of the Enlightenment creed? Because I realized that their arguments – not just the style but the substance – were the same. The joint theme is that both the intentional stance and adaptionism make a "question-begging" appeal to optimality when the proper way for science to proceed here is to do unadorned mechanical history of actual selection. Just as Quine and Skinner abjure borrowing intelligence (intentionality), Gould and Lewontin abjure borrowing optimality of design. And since rationality *is* optimality of cognitive design, one can look at the intentional stance as just a special case of adaptationist thinking. I lump Gould and Lewontin with Skinner because in spite of their ideological differences elsewhere, here they are saying the same thing about the same issue, and they have all overstated their case.

Dahlbom's point about paying lip service is important. The truth about *which* ideas to emphasize *when* in science would probably lie in the boring middle ground, were it not psychologically important to researchers and theorists to have a somewhat radical and intolerant conviction about how everything fits (will jolly well be made to fit) into one elegant vision. So we can expect the adaptationists to pay mere lip service to "constraints," just as their opponents pay mere lip service to adaptation, and with any luck, like Jack and Mrs. Sprat they will lick the platter clean – which is probably more than we could hope for from a herd of mealy-mouthed pluralists.

2. Yes, what **Dahlbom** calls the problem-solving approach and intentional system theory are one and the same thing under different names. I am not claiming to have a whole new way of doing science to offer to ethologists; I am just pointing out, as philosophers are wont to do, the conceptual obligations and privileges of a way of doing business that is already familiar, if often undertaken under false or strained pretenses. And surely Dahlbom is right that the higher-order intentional

attributions will find their main utility among the higher animals if anywhere, but one must remember that the higher-order attributions are also useful in analyzing relationships that are *not* "appreciated" by the individual organisms – the low-nesting birds, or the bees, for instance – but only by the evolutionary process that created and preserves the regularity in those relationships.

3. The intentional stance will fail, **Dahlbom** thinks, when common ground is lacking – when the target creatures are not enough like (human) *persons*. I am unconvinced (cf. Stich 1981, and my reply to Stich, Dennett 1981, and Stich 1983). It is no doubt a more difficult exercise of the imagination to "think like a Martian" – or a beaver or a coyote – but so what? And I do not see that it is made more difficult to think of superpersonal organizations when one thinks of individuals as intentional systems. A bureaucrat *is* an intentional system; to think of someone as a bureaucrat is not to demote him but to identify a type of intentional system to which the individual belongs. Bureaucrats – when they act true to form – do exactly what it is rational to do if one is in the bureaucrat's particular (and ubiquitous) predicament. If you or I were stuck being bureaucrats (if the options of rebellion or obstreperousness were particularly unattractive – thanks to our having to see three children through college, say) we would be stuck believing as bureaucrats do, and behaving as bureaucrats do, for under those grubby circumstances that is what it is rational to do. Some of the distinctive features of Weber's ideal types may be merely habitual or traditional – manifesting a sort of drag that creates a gap between actual practice and ideally rational practice – but the core of every such system is a rational practical reasoner. So I don't see the problem Dahlbom poses as looming large at all.

4. **Dahlbom** nicely describes the way "the Romantic theory [of "prejudiced" trial and error] corrects the gradualism and adaptationism of the Enlightenment by introducing saltational and structural elements." A good thing, but haven't we already seen that this is a matter of emphasis – like the grain problem that bedevils (and vitiates) the debate over punctuated equilibrium? What looks gradual from a bird's-eye view looks like fits and starts midst stasis from close up. Of course there has to be some biasing structure to constrain the trial and error process. Dahlbom suggests that a "move whereby such 'constraints' take on a central role" would "call into question the basis" of my theory. That is true; any theory of learning or development or evolution that gives the *central* role to what I call the constraints will be strongly opposed to mine, and to adaptationism, but such a theory will have the huge task of explaining (and not merely paying lip service to) the excellence of design and aptness of thought so normal in our world – without ever appealing to adaptationist trends. I am less convinced than Dahlbom that Gould and Lewontin (or Kimura or anyone else in biology) are pointing the way to such a theory. Simply describing – let alone explaining – these phenomena has traditionally depended on intentional language, with its assumptions of rationality or optimality built in. Skinner learned, to his discomfort, that he simply couldn't describe the domain of his field without lapsing into the suspect "mentalistic" vernacular with its tacit appeals to rationality. As the antiadaptationists are learning, it is equally quixotic to set oneself the purist goal of an account of evolution that is shriven of all Panglossian formulae.

Dahlbom's points 5 and 6 develop themes that require more thought from me, but so far as I can see I can agree with him. His reminder of the fate of general systems theory sends a salutary chill down my spine. Forewarned is forearmed.

My recommending a postbehaviorist vision to ethologists has provoked **Gray**, who sees me "using an outdated cast of characters," focusing on Skinner and ignoring such latter-day behaviorists as Hebb, Hull, Tolman, Bindra, and Kamin. Is Kamin really a behaviorist? Are Carew, Walters, and Kandel? I think Gray might take his own advice: "Perhaps it is time to banish 'behaviorist' and related terms" – if we are intent on referring to work on behavior that is only indirectly linked to the tradition of Watson, Thorndike, and Skinner (and Hull, Tolman, etc.).

What is **Gray**'s point? I guess it is that *non*cognitivist, (neo-?)behaviorist psychologists have something to offer ethologists that I have overlooked. I don't agree. It seems to me that Hebb, Bindra, and others managed at best to demonstrate how difficult and barren such approaches were, even when pursued with energy and brilliance. (Gray says I "might also have taken Bindra more seriously." I have, in Dennett 1978b, a commentary on Bindra's article in *BBS* which I am content to let serve as my summary of what was wrong not only with Bindra's approach, but with the other late behaviorists' attempts at theory-construction.) Nor do I think those who really are making progress on the "nature of the mechanisms" (Carew, Walters, and Kandel are a fine example) have much to say to ethologists yet, since the transfer from *Aplysia* or insects to birds and mammals is such a long journey that most of the good baggage must be abandoned along the way. [See also Hoyle: "The Scope of Neuroethology" *BBS* 7(3) 1984.]

Gray says that "we have found out that the nervous system is not made of components available in Radio Shack." This is presumably a Bronx cheer directed at AI, but it misses its mark and strikes some caricature inhabiting Gray's imagination. I cannot think of a single proponent of AI, no matter how fanatic or radical, whose views are challenged by this remark. And when Gray observes that "Dennett is dejected that all romance will be taken out of life" he convicts himself of a rather heroic misreading of my playfully labeled scale from romantic to killjoy. True-blue Behaviorists are only first-order intentional systems; they have beliefs and desires (we all do – behaviorism is false), but they don't believe that they or anyone else does. And so one telling symptom of behaviorism, not surprisingly, is obliviousness to humor.

References

Bateson, P. P. G. (1983a) Rules for changing the rules. In: *Evolution from molecules to men*, ed. D. S. Bendall. Cambridge University Press. [BD]

(1983b) Uncritical periods and insensitive sociobiology. *Behavioral and Brain Sciences* 6:102–3. [BD]

Bennett, J. (1983) Cognitive ethology: Theory or poetry? *Behavioral and Brain Sciences* 6:356–58. [BD]

Ben-Zeev, A. (1983) Toward a different approach to perception. *International Philosophical Quarterly* 23:45–64. [AB-Z]
Bindra, D. (1976) *A theory of intelligent behavior.* Wiley. [TG]
(1978) How adaptive behavior is produced: A perceptual-motivational alternative to response-reinforcement. *Behavioral and Brain Sciences* 1:41–91 [TG]
Bogdan, R. J. (1983) Fodor's representations. *Cognition and Brain Theory* 6:237–49. [RJB]
Carew, T., Walters, E. & Kandel, E. (1981) Associative learning in Aplysia: Cellular correlates supporting a conditioned fear hypothesis. *Science* 211:501–4. [TG]
Chomsky, N. (1959) A review of Skinner's *Verbal Behavior. Language* 35:26–58. [BD]
Churchland, P. S. (1983) Dennett's instrumentalism: A frog at the bottom of the mug. *Behavioral and Brain Sciences* 6:358–59. [BD]
Danto, A. C. (1983) Science as an intentional system. *Behavioral and Brain Sciences* 6:359–60. [BD]
Dennett, D.C. (1969) *Content and consciousness.* Routledge & Kegan Paul. [BD]
(1971) Intentional systems. *Journal of Philosophy* 68:87–106. Reprinted in Dennett *Brainstorms*, 1978a. Bradford/MIT Press. [BD]
(1975) Why the law of effect will not go away. *Journal of the Theory of Social Behavior* 5:169–87. Reprinted in Dennett, *Brainstorms*, 1978a. Bradford/MIT Press. [BD]
(1978a) *Brainstorms.* Bradford/MIT Press. [BD, TG]
(1978b) Requisition for a Pexgo. *Behavioral and Brain Sciences* 1:56–7. [DCD]
(1981a) Three kinds of intentional psychology. In: *Reductionism, time, and reality*, ed. R. Healey. Cambridge University Press. [BD]
(1981b) Making sense of ourselves. *Philosophical Topics* 12: 63–82, reprinted in *Mind, brain, and function*, ed. J. Biro & R. Shahan, 1982. University of Oklahoma Press. [DCD]
(1983) Intentional systems in cognitive ethology: The "Panglossian paradigm" defended. *Behavioral and Brain Sciences* 6:343–90. [AB-Z, RJB, BD, TG]
(1984) Cognitive wheels: The frame problem of AI. In: *Minds, machines and evolution*, ed. C. Hookway. Cambridge University Press. [BD]
Dretske, F. I. (1981) *Knowledge and the flow of information.* Bradford/MIT Press. [BD]
(1983) Précis of *Knowledge and the flow of information. Behavioral and Brain Sciences* 6:55–90. [BD]
Eldredge, N. & Gould, S. J. (1972) Punctuated equilibria: An alternative to phyletic gradualism. In: *Models in paleobiology*, ed. T. J. M. Schopf. Freeman, Cooper and Co. [BD]
Farley, J. & Alkon, D. (1980) Neural organization predicts stimulus specificity for a retained associative behavioral change. *Science* 210:1373–74. [TG]
Ginzburg, C. (1980) Morelli, Freud and Sherlock Holmes: Clues and scientific method. *History Workshop* 9:5–36. [BD]
Gould, S. J. & Eldredge, N. (1977) Punctuated equilibria: The tempo and mode of evolution reconsidered. *Paleobiology* 3:115–51. [BD]
Gould, S. J. & Lewontin, R. C. (1979) The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. *Proceedings of the Royal Society of London* B205:581–98. [BD, TG]
Hofstadter, D. (1979) *Gödel, Escher and Bach: An eternal golden braid.* Basic Books. [TG]
Kamin, L. (1969) Predictability, surprise, attention, and conditioning. In: *Punishment and aversive behavior*, ed. B. A. Campbell & R. M. Church. Appleton-Century-Crofts. [TG]
Kuhn, T. S. (1962) *The structure of scientific revolutions.* University of Chicago Press. [BD]
Lévi-Strauss, C. (1955) *Tristes tropiques.* Plon. English trans. *A world on the wane*, by John Russell. Hutchinson, 1961. [BD]
(1966) *The savage mind.* Weidenfeld and Nicolson. [BD]
(1969) *The elementary structures of kinship.* Beacon Press. [BD]
Lewontin, R. (1981a) The inferiority complex. *New York Review of Books* 28:12–16. [TG]
(1981b) Sleight of hand. *Sciences* 21:23–26. [TG]
Lloyd, D. (1983) The scope and ingenuity of evolutionary systems. *Behavioral and Brain Sciences* 6:368–69. [AB-Z]
Lumsden, C. J. & Wilson, E. O. (1981) *Genes, mind, and culture.* Harvard University Press. [TG]
Neisser, U. (1967) *Cognitive psychology.* Appleton-Century-Crofts. [BD]
Rosenberg, A. (1983) Content and consciousness versus the intentional stance. *Behavioral and Brain Sciences* 6:375–76. [BD]
Skinner, B. F. (1983) A better way to deal with selection. *Behavioral and Brain Sciences* 6:377–78. [BD]
Stich, S. (1981) Dennett on intentional systems. *Philosophical Topics* 12:39–62, reprinted in *Mind, brain, and function*, ed. J. Biro & R. Shahan, 1982. University of Oklahoma Press. [DCD]
Stich, S. (1983) *From folk psychology to cognitive science.* Bradford/MIT Press. [DCD]
van den Berghe, P. L. (1983) Human inbreeding avoidance: Culture in nature. *Behavioral and Brain Sciences* 6:91–123. [BD]
Watson, J. & Crick, F. (1953) Molecular structure of nucleic acids: A structure for deoxiribose nucleic acid. *Nature* 171:737–38. [TG]
Weber, M. (1964) *The theory of social and economic organization.* Free Press. Reprinted in Weber (1968). [BD]
(1968) *Economy and society.* Bedminster Press. [BD]
Wittgenstein, L. (1958) *Philosophical investigations.* Macmillan. [AB-Z]

Erratum

The commentary by Schwartz (1985) had the following coauthor: Myrna Schwartz, Department of Psychology, University of Pennsylvania, Philadelphia, Pa. 19104. The citation should accordingly be: Schwartz, B. & Schwartz, M. (1985) Organic insight into mental organs. *Behavioral and Brain Sciences* 8:30–31.