Evolution, teleology, intentionality

Daniel C. Dennett

Center for Cognitive Studies, Tufts University, Medford, MA 02155-7068

No response that was not as long and intricate as the two accompanying commentaries combined could do justice to their details, so what follows will satisfy nobody, myself included. I will concentrate on one issue discussed by both commentators: the relationship between evolution and teleological (or intentional) explanation. My response, in its brevity, may have just one virtue: It will confirm some of the hunches (or should I say suspicions) that these and other writers have entertained about my views. For more closely argued defenses of my points, see Dennett (1990a; 1990b; 1990c; 1991a; 1991b).

As **Ringen** notes, I have claimed that mentalistic or intentional explanations are not just similar to adaptationist explanations of evolution but continuous with them; there is just one sort of explanation here, operating according to one set of principles. Ringen thinks this is mistaken, and presents me with a dilemma: I must side either with the neo-Darwinians (who offer to reduce or even eliminate teleology via a mechanistic model of natural selection) or with **Bennett** (who according to Ringen champions a nonreductive, Aristotelian concept of real teleology). The position Bennett presents is more nuanced than Ringen suggests, but he supports a version of Ringen's challenge: He deplores what he sees as my fence-sitting "stance" talk, and urges me to get real. Where I have said that "nothing without a great deal of structural and processing complexity could conceivably realize an intentional system of any interest," Bennett "would replace the last phrase by 'a genuinely intentional system,' leaving 'interest' out of it." He sketches several of his proposals for settling the determinable questions of intentional attribution in ways, he claims, that are independent of my appeals to evolution.

There is a symmetry to **Bennett**'s and **Ringen**'s disagreements with me. Ringen maintains that, contrary to what I have said, the concept of *selection for,* and hence a basis for adaptationist theorizing in biology, can be secured independently of any intentionalizing of the design process – one need not appeal to "what Mother Nature had in mind." Bennett maintains that, contrary to what I have said, assertions about intentional attributions – about what an organism "had in mind" – can be secured independently of any assumptions about the provenance in evolution of the organism in question. If they were both

right, we could have a nonintentional evolutionary theory and a nonevolutionary theory of intentionality.

I continue to think they are both wrong. The apparent differences between adaptationist theorizing in biology and intentionalist theorizing in psychology are due, in my view, to the huge differences in time scale, and – more evident in the discussions of both **Ringen** and **Bennett** – a downplaying of the importance of the implications of the ubiquitous idealizing assumptions in both enterprises. When we grasp the nettle and confront the inelim- inable "practical difficulties" (Ringen) that beset the evo- lutionary theorist intent on distinguishing actual cases of *selection for,* and the parallel practical inability of the intentionalist psychologist to cash out the idealizing as- sumptions that permit talk (in Bennett's example) about a "class of environments . . . unified with help from the concept of food-getting" we see that both enterprises continue to avail themselves – quite appropriately and defensibly – of what Quine called the "dramatic idiom"; the sense-making interpretation-talk of the intentional stance. I claim that since there is just one sort of explana- tion going on in both quarters, the choice Ringen offers me must be rejected: Teleology is neither as illusory as his neo-Darwinians claim, nor as real and irreducible as his Aristotelian Bennett claims.

Ringen renders usefully explicit the vision of real teleology that haunts current thinking both in evolution- ary theory and in philosophy of mind – where I have in mind particularly Dretske's (1986; 1988b) quest for a causal role for meanings. Suppose there were such a thing as a genuinely teleological system, or, equivalently, a *real* (as opposed to approximate or "as-if") intentional system: "Teleological principles provide a basis for predicting what response to new circumstances a system which conforms to them will produce" (Ringen, para. 5; see also Bennett, sect. 7, but note also that Bennett recognizes this to be too idealized, because of the omnipresent possibility of error). Such a system would not just happen to track appropriateness; it would do so in a principled way. It would be *caused*, in Dretske's view, to track meanings in an appropriate way. But there are no such systems, human or otherwise. There are only better and worse approximations of this ideal – which is rather like the ideal of a frictionless bearing, or a perfectly failsafe alarm system. As Ringen points out, the process of natural selection does not quite measure up as a teleological system. Selection itself can only filter, at best supporting the conditional: *If* the appropriate sort of variation is generated, it will be selected. The generation process that provides the candidates for sorting itself is deemed by orthodoxy to be unresponsive to appropriateness. So there can be no guarantee, or anything even close to a guarantee, of genuine "teleological" or meaning-tracking behavior in evolution. I agree, then, with the passage Ringen quotes from Lewontin (1979): "The dynamics of natural selection does not include foresight, and there is no *theoretical* principle that assures optimization as a consequence of selection."

Both **Ringen** and **Bennett** would like to accept the invited contrast of this orthodox view of evolution with a design process controlled by an Intelligent Artificer (or just an intelligent artificer – an everyday, foresightful intentional system such as an engineer). When we look closely at the contrast, however, do we discover anything

but differences in degree? Some engineers are doltish and habit-bound; *if* a particular design solution happens to occur to them, they will adopt it, but there is no guarantee that they will generate the move that we can see in hindsight is the appropriate move in the circumstances. Some engineers are much cleverer, and some have positively brilliant insights into the reasons for and against particular design proposals. How adroit, how flexible, how sensitive must a system be to these reasons for it to be a *real* intentional system?

Bennett's "unity condition" is supposed to answer this question: If "the class of environments is unified with help from the concept of behaving in a manner appropriate" to this or that feature, then we are entitled to attribute that concept to that system, not as a *façon de parler* but literally. But one theorist's unifying concept is another theorist's inflationary shorthand for a mere disjunction of tropisms (cf. Dretske (1986) and Dennett 1987, Ch. 8). Bennett in effect concedes this, for he casts his question in terms of when we may hypothesize that there are going to be more disjuncts than we have observed: "What can make it all right for us to trust an intentional or teleological generalization to lead us from some S-M linkages to predict *others*?" (sect. 7, para. 3). Bennett suggests that either hypotheses about evolution or learning or both could underlie our confidence that one way or another there were mechanisms in an organism (or artifact) that would tend to yield further appropriate linkages. I agree (see Dennett 1990a; 1990b; 1991b), and I do not see (yet) why Bennett claims that his view in this regard is "quite different from" what I have been saying.

I think **Ringen**'s optimism about the independent application of optimality principles in evolutionary theory is similarly undercut. In discussing the case of the sexually reproducing snails' response to the castrating trematode parasite, for instance, he says: "Optimality principles predict that such optimal traits will emerge." I think not. Optimality principles predict that either such optimal traits will emerge or they will not; in the latter case, either the parasites will secure their own extinction by the extinction of their nonadapting hosts, or some semistable exploitation cycle will persist indefinitely. There are no guarantees, only the rationales of hindsight. But do not knock hindsight; one way or another, it is the only sort of sight we can ever count on having. At our best, our adaptive mechanisms lag slightly behind reality, tracking it ever more doggedly, but never giving us a "principle" by which we might predict genuinely teleological activity.

Finally, I will comment all too briefly on some of **Bennett**'s other constructive criticisms and objections. Bennett corrects my interpretation of his views on Quine's 1960 indeterminacy thesis. The view he and Blackburn (1975) hold had not occurred to me, and I have no opinion, yet, on whether, as he claims, determinacy of language is consistent with indeterminacy of thoughts.

Bennett describes my view of intentional attribution as "free-ranging, somewhat haphazard" since it is governed by only "two extremely mild constraints": a rationality assumption and a prohibition of inflationary attributions. He claims that, on the contrary, "a good deal of discipline" can be brought to bear on the project. I have no quarrel with the details of his sketched example (the animal's food-seeking behavior); I just think the considerations he correctly raises are subsumed under my constraints,

which are not mild at all. There is plenty of structure to the reasoning processes that govern the postulation and support of intentional attributions, and it is generated, indirectly, by my minimal constraints.

I think Bennett misunderstands the strategy of my vending machine example. He is not alone, so it is my fault. He is right that the vending machine is even worse than the thermostat as an example of an intentional system – that was deliberate on my part. I wanted to choose an example of a dead-simple, quasiperceptual mechanism (the counterfeit-coin detector) so there would be no controversy about "what we would say"; *of course* there is no deep fact of the matter in this instance about which cases of coin-rejection count as "errors." Any grounds for calling some cases errors and others proper functioning will have to depend on the embedding of the device in a large context of purposes: the purposes of its users. The challenge is then for the believers in deeper facts about content in fancier cases (Twin Earth cases in particular) to show what features of these fancier cases permit us to invoke other principles. I claim they cannot, and I do not see that Bennett's discussion provides any such grounds. Bennett says we do not have to solve the problem of error for the vending machine. We do not *have* to solve the problem of error for anything; we can always ("in principle") eschew intentional discourse and settle for brute physical stance mechanism. But *if* we find it illuminating to adopt the intentional stance (and even in the case of the vending machine, the error-talk is illuminating – just think of the design-improvement process, the invocation of Gresham's Law, etc.), we will find ourselves invoking the minimal but none-too-mild constraints of the intentional stance.

References

Amundson, R. (1989) The trials and tribulations of selectionist explanations. In: *Issues in Evolutionary Epistemology,* ed. K. Hahlweg & C. A. Hooker. State University of New York Press. [JR]

Aristotle (1907) *De anima.* (R. B. Hicks, Trans.). Cambridge University Press. [JR]

Beatty, J. (1980) Optimal design models and the strategy of model building in evolutionary biology. *Philosophy of Science* 47:532–61. [JR]

Bechtel, W. (1986) Teleological functional analyses and the hierarchical organization of nature. In: *Current issues in teleology,* ed. N. Rescher. University Press of America. [JR]

Bennett, J. (1976) *Linguistic behaviour.* Cambridge University Press. (2nd edition [1990] published by Hackett.) [JB, JR]

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

(1991a) How to read minds in behaviour: A suggestion from a philosopher. In: *Natural theories of mind: Evolution, development and simulation of everyday mindreading,* ed. A. Whiten. Basil Blackwell. [JB]

(1991b) How is cognitive ethology possible? In: *Cognitive ethology: The minds of other animals,* ed. C. A. Ristau. Erlbaum. [JB]

(1991c) Folk psychological explanations. In: *The future of folk psychology: Intentionality and cognitive science,* ed. J. D. Greenwood. Oxford University Press. [JB]

Blackburn, S. (1975) The identity of propositions. In: *Meaning, reference and necessity: New studies in semantics,* ed. S. Blackburn. Cambridge University Press. [rDCD, JB]

Block, N. (1980) What is functionalism? In: *Readings in philosophy of psychology,* ed. N. Block. Harvard University Press. [JB]

Brandon, R. (1981) Biological teleology: Questions and explanations. *Studies in History and Philosophy of Science* 12:91–105. [JR]

Burtt, E. (1951) *The metaphysical foundations of modern physical science.* (Revised ed.) Humanities Press. [JR]

Churchland, Paul (1979) *Scientific realism and the plasticity of mind.* Cambridge University Press. [JR]

Cowan, R. (1990) Parasite power. *Science News* 138:200–02. [JR]

Darden, L. & Cain, J. (1989) Selection type theories. *Philosophy of Science* 56:106–29. [JR]

Darwin, C. (1923) *The origin of species by means of natural selection* (Sixth edition). D. Appleton. [JR]

Dennett, D. C. (1974) Why the law of effect will not go away. *Journal of the Theory of Social Behavior* 5:169–87. [JB]

(1978) Why the law of effect will not go away. In: *Brainstorms.* Bradford Books. [JR]

(1987) *The intentional stance.* MIT Press/Bradford Books. [rDCD, JB, JR]

(1988) Précis of *The Intentional Stance* [followed by Open Peer Commentary and Author's Response]. *Behavioral and Brain Sciences* 11:495–546. [JB, JR]

(1990a) Ways of establishing harmony. In: *Dretske and his critics,* ed. B. McLaughlin. Blackwell. [rDCD]

(1990b) The interpretation of texts, people, and other artifacts. *Philosophy and Phenomenological Research* 50:177–94. [rDCD]

(1990c) Dr. Pangloss knows best [reply to Amundson]. *Behavioral and Brain Sciences* 13:581–82. [rDCD]

(1991a) Real Patterns. *Journal of Philosophy* 87:27–51. [rDCD]

(1991b) Do-it-yourself understanding. Center for Cognitive Studies Preprint CSS-90-4. Tufts University, Medford, MA. [rDCD]

Dretske, F. (1985) Machines and the mental. *Proceedings and Addresses of the American Philosophical Association* 59:23–33. [JB]

(1986) *Misrepresentation.* In: *Belief,* ed. R. Bogdan. Oxford University Press. [rDCD, JB]

(1988a) The stance stance. *Behavioral and Brain Sciences* 11(3):511–12. [JB]

(1988b) *Explaining behavior: Reasons in a world of causes.* MIT Press/Bradford Books. [rDCD]

Driesch, H. (1914) *The history and theory of vitalism.* (C. K. Ogden, Trans.) Macmillan. [JR]

Dupré, J. (1987) *The latest on the best: Essays on evolution and optimality.* MIT Press. [JR]

Ghiselin, M. (1969) *The triumph of the Darwinian method.* University of California Press. [JR]

Gould, S. (1980) Caring groups and selfish genes. In: *The Panda's Thumb.* W. W. Norton. (Reprinted in *Conceptual issues in evolutionary biology: An anthology,* ed. E. Sober. MIT Press. [JR]

Gould, S. & Lewontin, R. (1979) The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. *Proceedings of the Royal Society.* B205:581–98. Reprinted in *Conceptual issues in evolutionary biology: An anthology,* ed. E. Sober. MIT Press. [JR]

Holling, C. (1964) The analysis of complex population processes. *Canadian Entomologist* 96:335–47. [JR]

Kant, I. (1990) *The critique of practical judgment.* Hackett. [JR]

Kuhn, T. (1962) *The structure of scientific revolutions.* University of Chicago Press. [JR]

Lehman, H. (1965) Functional explanations in biology. *Philosophy of Science* 32:1–20 [JR]

Lenoir, T. (1982) *The strategy of life: Teleology and mechanics in 19th century German biology.* D. Reidel. [JR]

Lewontin, R. (1979) Fitness, survival, and optimality. In: *Analysis of ecological systems*, ed. D. Horn, G. Stairs & R. Mitchell. Ohio State University. [JR]

Mace, C. A. (1935) Mechanical and teleological explanation. In: *Science, History, and Theology* (Aristotelian Society, Supplementary). 14:22–45. Harrison & Sons Ltd. [JR]

Maull, N. (1977) Unifying science without reduction. *Studies in History and Philosophy of Science* 9:143–62. Reprinted in *Conceptual issues in evolutionary biology: An anthology*. MIT Press. [JR]

Millikan, R. (1984) *Language, thought, and other biological categories: New foundations for Realism*. MIT Press. [JR]

(1989) In defense of proper functions. *Philosophy of Science* 56:288–302. [JR]

Nagel, E. (1977) Teleology revisited. *The Journal of Philosophy* 74:61–300.

Papineau, D. (1984) Representation and explanation. *Philosophy of Science*. 51:550–72. [JR]

Quine, W. V. (1960) *Word and object*. MIT Press. [rDCD, JB, JR]

Ringen, J. (1976) Explanation, teleology, and operant behaviorism: A study of the experimental analysis of purposive behavior. *Philosophy of Science* 43:223–53. [JR]

(1985) Operant conditioning and a paradox of teleology. *Philosophy of Science* 52:565–77. [JR]

(1986) The completeness of behavior theory: A review of *Behaviorism, science, and human nature* by Barry Schwartz & Hugh Lacey. *Behaviorism* 14:29–39. [JR]

Ruse, M. (1971) Function statements in biology. *Philosophy of Science* 38:87–95. [JR]

Shanahan, T. (1990) Evolution, phenotypic selection, and the units of selection. *Philosophy of Science* 57:210–25. [JR]

Skinner, B. F. (1969) *Contingencies of reinforcement: A theoretical analysis*. Appleton-Century-Crofts. [JR]

(1984) The evolution of behavior. *Journal of the Experimental Analysis of Behavior* 41:217–21. [JR]

Sober, E. (1981) Evolutionary theory and the ontological status of properties. *Philosophical Studies* 40:147–76. [JR]

(1984) *The nature of selection*. MIT Press. [JR]

(1985) Methodological behaviorism, evolution, and game theory. In: *Sociobiology and epistemology*, ed. J. Fetzer. D. Reidel. [JR]

(1986) Comments on Rosenberg's review. *Behaviorism* 14:89–96. [JR]

Staddon, J. E. R. (1983) *Adaptive behavior and learning*. Cambridge University Press. [JR]

Stich, S. (1983) *From folk psychology to cognitive science: The case against belief*. MIT Press. [JR]

(1988) Connectionism, Realism and realism. *Behavioral and Brain Sciences* 11(3):531–32. [JB]

Taylor, C. (1964) *The explanation of behaviour*. Humanities Press. [JR]

von Wright, G. (1971) *Explanation and understanding*. Cornell University Press. [JR]

Wimsatt, W. (1972) Teleology and the logical structure of function statements. *Studies in the History and Philosophy of Science* 3:1–80. [JR]

Wright, L. (1976) *Teleological explanations*. University of California Press. [JR]