Hist. Phil. Life Sci., 27 (2005), 249-280

Natural Selection and the Conditions for Existence: Representational vs. Conditional Teleology in Biological Explanation

John O. Reiss

Dept. of Biological Sciences
Humboldt State University
Arcata, CA 95521, USA

ABSTRACT - Human intentional action, including the design and use of artifacts, involves the prior mental representation of the goal (end) and the means to achieve that goal. This representation is part of the efficient cause of the action, and thus can be used to explain both the action and the achievement of the end. This is intentional teleological explanation. More generally, teleological explanation that depends on the real existence of a representation of the goal (and the means to achieve it) can be called representational teleological explanation. Such explanations in biology can involve both external representations (e.g., ideas in the mind of God) and internal representations (souls, vital powers, entelechies, developmental programs, etc.). However, another type of explanation of intentional action (or any other process) is possible. Given that an action achieving a result occurs, the action can be explained as fulfilling the necessary conditions (means) for that result (end), and, reciprocally, the result explained by the occurrence of those necessary conditions. This is conditional teleological explanation.

For organisms, natural selection is often understood metaphorically as the designer, intentionally constructing them for certain ends. Unfortunately, this metaphor is often taken rather too literally, because it has been difficult to conceive of another way to relate natural selection to the process of evolution. I argue that combining a conditional teleological explanation of organisms and of evolution provides such an alternative. This conditional teleology can be grounded in existence or survival. Given that an organism exists, we can explain its existence by the occurrence of the necessary conditions for that existence. This principle of the 'conditions for existence' was introduced by Georges Cuvier in 1800, and provides a valid, conditional teleological method for explaining organismal structure and behavior. From an evolutionary perspective, the conditions for existence are the range of boundary conditions within which the evolutionary process must occur. Moreover, evolutionary change itself can be subjected to conditional teleological explanation, because natural selection theory is primarily a theory about the relation between the conditions for the existence of organisms and the conditions for the existence of traits in populations. I show that failure to distinguish representational from conditional teleological explanation has confused previous attempts to clarify the relation of teleology to biology.

KEYWORDS - Cuvier, function, design argument, evolution.

This paper is about how biology can rid itself of illegitimate teleological thinking. Such thinking, based on the metaphor of natural

selection as the designer of biological systems, is persistent, perhaps even more so among philosophers of biology than biologists. I first make a distinction between two types of teleological explanation, which I call representational and conditional explanation. I argue that these are two entirely distinct explanatory modes, and that conditional teleological explanations are valid both for representational systems and for systems to which we are not willing to ascribe representations and attendant intentions or purposes. I argue further that a false determinism also can result in teleological thinking, in which achieved ends are explained as the necessary result of laws that are in themselves insufficient to account for the ends explained; I call this teleological deterministic explanation. I point out that conditional teleological explanation of organisms can be grounded in organismal existence or life, a perspective first developed by Georges Cuvier some 200 years ago. Finally, to place my views in context, I examine a few representative views of other authors from the perspective I have developed.

Intentional and Representational Teleological Explanation

Human intentional action is perhaps the psychological source of teleological thinking in general. We commonly say, for example, 'I went to the store to buy bread' and consider this an explanation of our actions. Almost everyone agrees that from a scientific standpoint this is not problematic, and involves no reverse causation, or cause that comes after its effect (a common concern of critics of teleology). My desire for bread, combined with my belief that I can get bread at the store (both of which occur prior to my going to the store) explain the act of getting there and buying the bread. This is *intentional teleological explanation*, explanation of future-directed action based on positing an *intentional system* with a prior mental representation of the goal of the action and the means to get there, as well as the actual means to carry out the act, as part of the efficient cause of the action.

Intentional teleology is also involved in explaining the production and use of artifacts. If I make something, say a clock, my desire for a device to tell time (end) together with my belief that a device built in a certain way will allow me to tell the time (means), and my ability to actually implement these means, explain the production (and thus the existence) of the clock.

If we treat the intentional teleology involved in acts and the making of artifacts as the paradigm for teleology in general, the issue for teleological conceptions in biology is whether nature can have

intentions, or at least something analogous to intentions, and thus to what degree we can unite explanations of intentional acts and artifacts with explanations of instinctive behaviors of organisms and of organisms themselves.

There are really two distinct questions here. One is the question of whether representations of goals and the means to achieve them exist outside of humans. If such representations exist, and we can determine what they are, then we can explain the behavior of systems containing them with reference to such representations. Such explanations might be called representational teleological explanations. Intentional explanations are merely a special case of this general type of explanation – the case in which the representations are made to an internal consciousness. Explanations based on internal representations of the goal state (souls, vital powers, entelechies, developmental programs, etc.) have been common in the history of biology.

However, the question I am primarily interested in here is entirely different from considerations of such representational teleology operating within organisms. It is the question of whether the evolutionary process that resulted in the organisms we see around us today can be used to explain features of those organisms, in particular, their apparent complexity, in a way analogous to the use of the representation (intention) in representational (intentional) explanation. Though these questions are distinct, they have often been confused, no doubt because part of organismal complexity is of course behavior that appears to us 'adaptive' or 'goal-directed,' and thus dependent on the existence of some internal representation of a goal state.

Conditional Teleological Explanation

In the same examples used to illustrate intentional or representational teleology another sort of teleological explanation is possible. Given that I do go to the store and buy bread, my going to the store (means) can be explained as a necessary condition[1] for buying bread (the end result). Moreover, the means can be broken up into parts, and the partial actions that make up this complex action – getting in my car, starting it, following a particular route, etc. – can all

[1] By necessary condition I don't mean that there was no other possible way to get bread, but rather that in the context of the actual act, going to the store was instrumental in getting the bread in a way that other activities of mine, such as chatting with a friend I ran into there, were not.

also be explained as necessary conditions for the completion of the specific act of which they were a part, and thus for the result. Given that I do make a clock (means) that has hands that move in a certain way (end), my making it was a necessary condition for it's having hands that move that way. Moreover, my giving the parts of the clock a certain shape and certain relationships can be explained as necessary conditions for the movement of the hands in just that way. This is what I will call conditional teleological explanation. Conditional teleological explanation is a form of causal explanation, but it is not causal in the same way that explanations based on physical laws are - a specific end result must be proposed for the significance of the preceding actions to be clear.

Importantly, once an overall goal, end or result is given or assumed, conditional teleological explanation does not depend on any intention or representation. If I go to the store and buy bread, my going to the store can be used to explain my buying bread, whether I intended to buy it or not. If I make a clock that has hands that move in a certain way, their movement can be explained by my making it that way, even if they don't move the way I intended. Such explanations of actions do not seem terribly informative, but this is because they are explanation of entire actions. What makes conditional teleological explanations useful is that the means-ends relationship can be broken up into smaller parts, thus making them *functional explanations*. The partial actions that make up a complex action can each be explained as means to the complex action as an end, the actions of parts of an artifact can be explained as means to the action of the artifact as a whole as an end. Getting in my car, starting the engine, etc. can be explained as means to the end of buying bread, having gears of a certain size and relation, etc., can be explained as means to the end of the clock's movement. In the case of artifacts, the independence from the intentions of an agent in such functional explanations is especially clear; given a clock that has hands that move in a certain way, the contribution of the parts to the overall motion of the clock can clearly be examined with no knowledge of its creator's intentions (though some end, e.g., a specific type of motion, must be assumed). As has been pointed out by McLaughlin (2001), the problem of functional explanation is not so much the problem of future direction, so much as it is the problem of the relation of the part to the whole, of reductionism vs. holism. In conditional teleological functional explanations the whole is taken as explanatorily prior to the parts, and the parts are explained as conditions for the whole (and the whole as conditioned by the parts).

In fact, it is the means-end relationship involved in conditional teleology that informs intentional teleology. If I desire bread, it is my belief that a certain set of conditions must be satisfied as means to that end that informs my actions. If I desire a clock, my belief that a certain structure will work to tell the time informs my actions. In intentional action we want something to happen, and we do what we think is necessary (we supply what we think are necessary conditions) to bring it about. A mental representation of conditional teleological relations thus plays a role in intentional action (cf. Rosen 1985). However, such representations are not the same as the actual conditional teleological relations. Shipwrecked on an island, we can pray to God, believing that this is a necessary condition for our salvation, but this doesn't mean that when we are saved, our prayers were of any use.

More generally, any system that behaves purposively may encode a representation of the goal state and have the means to get there, e.g., the set point of a thermostat and the mechanism that turns on the furnace when the temperature falls below it. But again, there is no necessary connection between the representation and achieving the goal: if the thermostat sends a current to turn on the furnace, but the furnace is broken and the house warms up by the sunlight coming through the windows, the representation had no actual role in the result.

Thus while conditional teleological explanation focuses on future ends, it is free from any assumptions about desires, beliefs, or representations. Moreover, because the ends are known to have occurred, no problem of the 'missing goal object' occurs. When a prior representation of the end and the means to get there agrees with the end achieved and the means actually used, representational and conditional teleological explanations are both possible, but this is not always the case. When the goal represented is frustrated (the case of the 'missing goal object') only representational teleological explanation is possible. On the other hand, when a set of events results in conditions necessary to a purpose being fulfilled with no role of a corresponding representation, only conditional teleological explanation is.

Teleology and Necessity

What is it that makes an explanation of a system or process teleological rather than simply nomological, i.e. occurring in

accordance with natural laws? This distinction turns out to depend on the type of necessity involved. As first recognized by Aristotle, teleological explanations (both representational and conditional) involve only a 'hypothetical necessity,' rather than the 'absolute necessity' involved in deterministic or law-like explanations (Randall 1960). In other words, teleological explanations must be explanations of events that are contingent (for us), because if the event is determined by known laws – and associated boundary conditions – there is nothing left for the end to explain that hasn't already been accounted for by these laws. Hypothetical necessity invokes only necessary conditions for a process to occur, whereas absolute necessity invokes sufficient conditions. Thus in order to buy bread it is necessary that I go to the store, but it is not necessary that I go to the store to buy bread in any absolute sense; I am not determined (at least by laws we are familiar with) to go. Confusion between these two modes of necessity, in addition to confusion between representational and conditional teleology, has also resulted in scientifically objectionable teleological explanations.

If one doesn't recognize the distinction between absolute and hypothetical necessity, one effectively treats teleological explanations as if they were deterministic. This fallacy might be called *deterministic teleology*. Such deterministic teleology can occur in two ways. First, it can occur when explicit representations of ends are considered to determine the result, although the causal connection between the representation and the end is not clear; this is *deterministic teleology* proper. For example, I might treat my representation of buying bread as the deterministic cause of my going to the store, although I can't trace the deterministic causal chain forward from my mental representation through nerve impulses to muscular movements that result in my getting to the store.

A second, more subtle deterministic teleology can occur when known mechanisms are considered to account for a particular result, even though it can't be derived from them; this might be called *teleological determinism*. To give a historical example, the young, precritical Kant presented a cosmogony in which he explained the origin of the structure of the solar system on the basis of Newton's laws, although he did not derive this structure from the laws (Kant 1981).[2] Such a false determinism is a representational teleological

[2] In contrast to Kant, Newton himself wisely refrained from considering his laws adequate even to account for the stability of the solar system, because he could not demonstrate that stability from them.

explanation in disguise; when an achieved end is viewed as necessary (in an absolute sense), even though we can't derive that necessity from natural laws, an implicit representation is used to complete the determination of the system. The location of this representation is often rather vague; it may be conceived of as in the laws themselves, the matter, or the mind of God. It is clear that such teleological determinism confuses the absolute necessity of the laws with the hypothetical necessity of the results.

In summary, we see that there are two primary dangers to be avoided in the use of teleological explanations in science: 1) the danger of attributing explicit intentions (or at least representations of the goal state) to nature where they don't exist, and 2) the danger of confusing a teleological with a deterministic explanation, and thus promoting a false, teleological determinism.

A Taxonomy of Teleology

Two types of teleology are commonly recognized: *purposive* and *functional* teleology (Beckner 1967; Woodfield 1998). This distinction has been generalized from the two major situations where teleological explanations have been considered appropriate: intentional behavior and the production and use of artifacts. *Purposive* or *goal-directed* teleology involves the behavior of a system with respect to a goal of the system as a whole. As is widely recognized, purposive or goal-directed teleology is *intrinsic* or *internal* teleology, in that the end that we recognize is a future state of the system itself. It is thus *monadic* (Boylan 1986), in the sense that it is unitary, and incapable of further division. By contrast with purposive teleology, *functional* teleology involves behaviors of parts that contribute to some activity of a larger system to which they belong. Functional teleology is *external* or *extrinsic* teleology, in that the function of the part is always relative to an activity of a larger containing system. For this reason Boylan (1986) called this type of teleology *systemic* teleology.

The distinction between purposive and functional teleology can be summarized by noting that purposive teleology focuses on ends, functional teleology on means. However, all teleology, whether functional or purposive, necessarily involves reference both to an end (goal) and to the means to that end, and it is clear that there must be more going on here. To clarify the situation, we must make another important distinction, which cuts across the distinction between purposive and functional teleology in general. This is the distinction

between teleological *systems* and teleological *explanations*. One needs to distinguish between positing the actual existence (in some sense) of ends in the world and our use of ends as principles of understanding.

These combined distinctions can be most easily summarized in the form of a table, relating teleological (functional and purposive) systems and teleological (conditional and representational) explanations (Table 1).

Table 1. *A taxonomy of explanations for functional and purposive systems*

	System	
Explanation	**Functional** (interacting parts with characteristic behavior)	**Purposive/Representational** (contains internal representation of goal)
Deterministic (behavior explained by properties of parts and laws of interaction)	**Valid** (if laws and properties can be found and boundary conditions specified)	**Valid** (if laws and properties can be found and boundary conditions specified)
Teleological Deterministic (properties of parts used to explain behavior, but they don't)	**Invalid** (implicit representation used to complete determination)	**Invalid** (implicit representation used to complete determination)
Representational (Internal) (internal representation of behavior/goal explains it)	**Invalid** (no such representation exists)	**Valid** (though may be difficult to establish which behavior/goal is represented internally; purposive functions determined with respect to overall goal)
Representational (External) (external representation of behavior/goal explains both behavior and system structure)	**Valid** (if system has been designed; design functions determined with respect to overall goal)	**Valid** (if system has been designed; design functions determined with respect to overall goal)
Conditional/Functional (system behavior explained by necessary conditions of parts and their interaction)	**Valid** (conditional functions determined with respect to a particular behavior)	**Valid** (conditional functions determined with respect to a particular behavior)

We can see that at least five types of explanation are possible for the behavior of such systems, though not all are valid:

1) *Deterministic Explanation*. In a deterministic/reductionistic explanation the behavior of the overall system can be inferred from the properties of the parts and the laws of their interaction, together with appropriate boundary conditions. Such explanation is obviously valid, given that we can define the parts and their properties in a way that the system properties can indeed be inferred from them and can specify the relevant boundary conditions. For example, we can construct a theoretical model of a pendulum clock, which derives its operating behavior from such facts as the laws of gravity, the ratios of the gears, etc., together with boundary conditions such as the local gravitational field. Similarly, we can construct a theoretical model of a thermostat, an electrical circuit controlling a furnace, etc., in which the operating behavior is derived from properties of the metal, laws of electrical conduction, etc., together with boundary conditions such as ambient external temperature, rate of heat gain or loss, etc.

2) *Teleological Deterministic Explanation*. As discussed above, when one attempts to give a deterministic explanation but does not specify the relevant laws and boundary conditions in a way that allows determination of the overall system behavior, one is guilty of a false, teleological determinism. Such an 'explanation' is at best a sketch of the possibility of a deterministic/reductionist explanation, not such an explanation itself. This type of pseudo-deterministic explanation is invalid both for functional and for purposive systems.

3) *Internal Representational Explanation*. In this mode of teleological explanation, the behavior of the system is explained on the basis of an internal representation of the behavior or goal within the system. This makes no sense for systems that are merely functional and have no such representation of a goal, but for purposive (representational) systems it does make some sense to explain their behavior with respect to this goal. Thus it makes sense to say that the thermostat turned the furnace on in order to bring the house back up to temperature, or that I went to the store in order to buy bread. Within such an explanation, the *purposive function* of parts of the system can be determined with respect to the overall goal. Thus the function of my pulling the car out of the driveway is to head toward the store, with respect to my overall goal of going to the store to buy bread. As with all teleological explanations, such an explanation is hierarchical, in that functions can be considered purposes with respect to lower level (smaller) parts of the system: my turning the key to start the car has a purposive function not only with respect to the overall purpose of buying bread, but with respect to the subsidiary purpose (function) of pulling the car out of the driveway.

4) *External Representational Explanation.* In this mode of teleological explanation, the behavior and structure of a system are explained on the basis of an external representation of this behavior and structure. This is of course the explanation based on intentional design. It is valid both for functional and purposive/representational systems, given that they are in fact designed (i.e., we can explain the behavior of both a car engine and a guided missile by the relevant intention, together with the assumption that the designer accurately understood the behavior that would result from his or her design). With respect to the overall representation of the intended behavior of the system, *design functions* are the intended functions of the parts. Again, these are hierarchical explanations: the design function of the engine is to provide power for the car to move down the road (intended system behavior or goal); the spark plugs have the design function of igniting the gas not only within the larger system of the car, but within the smaller system of the engine.

5) *Conditional Teleological (Functional) Explanation*. In this mode of explanation the overall goal or system behavior one is attempting to explain must be given; given this behavior, the necessary conditions for the behavior are objective facts that explain its occurrence. Such an explanation is of course valid both for merely functional and for purposive/representational systems. For example, given that the engine is running, and this is the behavior of the system we are interested in, with respect to that behavior the (conditional) function of the spark plugs is to ignite the gas. The fact that the spark plugs are igniting the gas helps explain the running of the engine. Likewise, given that the thermostat does turn on the furnace, and the relatively constant temperature of the house is the system behavior we want to explain, with respect to this behavior the (conditional) function of the thermostat is to turn on the furnace. Such conditional explanation has often been confused with explanation based on external representations (design), presumably because in the case of artifacts, the system behavior we choose to focus on is the design function of the system for humans.

What is clear from this overview is that both functional and purposive systems are generally subject to the same sort of explanations. The major difference, of course, is that only purposive systems can have their actions explained on the basis of internal representations of the goal state. Much of the confusion in discussions of teleology in biology has come from confusion between purposive and functional explanations of purposive systems, because the same system undergoing the same behavior can be understood both

purposively (based on an internal representation of its goal) and functionally. In fact, the behavior of a designed purposive system like a homing torpedo can be explained in four distinct, but valid ways: 1) deterministically, 2) based on the internal representation of the goal state, 3) based on the external representation of the goal state (design), and 4) functionally.

I have focused on artifacts so far, rather than organisms, because understanding organisms is more difficult. Given that we are trying to achieve a naturalistic explanation, and that a deterministic explanation of organisms is beyond our ability, only internal representational and conditional/functional explanations are possible. I have already noted that internal representational explanation of organisms is problematic, because it is difficult to determine what goals are internally represented; in any case, such explanation does not really concern us here. Instead, we need to focus on functional explanations. The big problem for functional explanation of organisms is of course determining an overall end to which the functions of the parts can be relativized. In the case of artifacts, this end is derived from the intention/purpose of the designer or user, but such relativism to a goal is clearly anti-naturalistic in the case of organisms – we no longer want to base functions on God's design. Alternatively, we can just relativize the functions to any activity we see the organism carrying out. But this creates the problem that there is no way to distinguish between 'true' functions and false ones. There is no way to say that pumping blood is a function of the heart, but making heart sounds is not. Thus functions become entirely subjective (Cummins 1975).

Instead, I would argue that one can make objective functional/conditional explanations of biological systems by grounding them on the continued existence of the system itself. Organisms, and living systems in general, are systems that depend on the activities of their parts for their continued existence (survival). Thus the activities of the parts can be relativized to that continued existence as their end result. This is the principle of the conditions for existence, to which I now turn.

The Principle of the Conditions for Existence

Georges Cuvier (1769-1832) was the founder of vertebrate paleontology and the leading comparative anatomist of his generation. Born in Montbéliard, then in the Duchy of Württemberg, and educated in Stuttgart, he weathered the French revolution as a tutor in Normandy, and soon after moved to Paris, where he became

director of the Musée d'Histoire Naturelle (Outram 1984). His theoretical stance was generally quite conservative; he is well known for stressing the primacy of fact over theory, and he used his *éloges* of other scientists to denigrate their speculative activity (Appel 1987; Coleman 1964). His rejection of evolution was also framed largely in these terms, and is not surprising considering the speculative excesses of the evolutionists of his day (e.g. Lamarck, Geoffroy St. Hilaire).

For Cuvier, the principle of the 'conditions for existence' (*conditions d'existence*) was the fundamental guide for work in natural history (see Letteney 1999 for an excellent discussion of Cuvier's principle). Cuvier's emphasis on this principle was in keeping with his entire approach to zoology, which always involved an Aristotelian stress on the functional unity of the organism. The principle was first broadly introduced in his *Leçons d'Anatomie Comparée* (Cuvier 1800, i., 47, translation adapted from several sources):

> Of what use would sensation be to us, if muscular force did not help it, even in the most trifling circumstances? What use could we make of touch, if we could not carry our hands toward the palpable object? and what should we behold if we could not turn our eyes or head at pleasure? It is on this mutual dependence between the functions, and on this reciprocal aid, that the laws are founded which determine the affinities of the organs of animals.... For it is evident, that a suitable harmony between those organs which act upon each other, is a necessary condition for the existence of the being to which they belong; and that if one of these functions were modified in a manner incompatible with the modifications of others, that being could not exist.

Cuvier elaborated further on this principal in his *Règne Animal* (Cuvier 1817, i., 6):

> Natural History nevertheless has a rational principle that is exclusive to it and which it employs to great advantage on many occasions; it is the *conditions for existence*, or, vulgarly, *final causes*. Since nothing can exist that does not fulfill the conditions that render its existence possible, the different parts of each being must be coordinated in such a way as to render possible the existence of the being as a whole, not only in itself, but also in relation with other beings, and the analysis of these conditions often leads to general laws which are as certain as those which are derived from calculation or from experiment.

As Cuvier recognized, this a teleological principle, but if one considers it in light of the distinctions drawn above between representational, deterministic and conditional teleology, it is clear that it involves only a conditional teleology. If we observe that an organism exists, then it must be *possible* for it to exist, but this does

not mean it was *designed* to exist, or that it *had* to exist.[3] The problem for a naturalistic teleology has always been an objective standard of value, for all teleological thinking involves not just ends but valued ends. What Cuvier recognized is that existence itself is the only end that objectively exists in nature.[4] The existence of organisms is a solid fact on which to ground our (conditional) teleological explanations of that existence. Cuvier's principle was later to be influential on the positivist philosopher Auguste Comte and the physiologist Claude Bernard (McClellan 2001).

The Meaning of the Phrase 'Conditions for Existence'

My translation of the *conditions d'existence* as 'conditions *for* existence' is unorthodox – almost all English translations have instead rendered it literally, as the 'conditions *of* existence' (though Lee 1833 translated it as 'condition to existence' and the English translation of Corsi 1988 as 'prerequisites of existence'). I have translated it thus because this gives the term less ambiguity: the 'conditions *of* existence' can in fact mean two quite different things.

The first of these meanings, and that of Cuvier, is that of the necessary conditions *for* the existence of an organism. These conditions are a characteristic of the organism. For an animal *obtaining enough food* is a condition of existence in this sense, where 'enough' is obviously relative to the particular organism in question. The second possible meaning is that of the environmental conditions, or circumstances, *in which* an organism exists. The *types of other organisms present* in the environment of an animal are conditions of existence in this sense. The reason I have preferred to translate the phrase as 'conditions for existence' is that this latter meaning, entirely different from Cuvier's (Russell 1916, 34), is thereby excluded.[5]

It is interesting to note that Spencer (1864, 445), in his discussion of the 'survival of the fittest', implicitly recognized the distinction

[3] Although Cuvier himself appears to have believed in supernatural design, he was also familiar with, and sensitive to, Kant's critique of the design argument, as well as the critiques of *philosophes* like Buffon and d'Holbach. Letteney (1999; see also Asma 1996) convincingly argues that Cuvier went beyond Kant's strictures in his use of the principle; however this does not mean that the principle itself is unsound.

[4] Perhaps it is here that he was most influenced by Kant, who argued that existence is not a predicate even in his precritical philosophy.

[5] Asma (1996), in an otherwise excellent treatment, unfortunately follows Darwin in giving it this latter meaning.

between the conditions for existence (his 'conditions *to* life') and the environmental conditions (his 'conditions *of* life'); the destruction of the 'diseased and feeble' through their 'failure to *fulfil* some of the conditions *to* life, leaves behind those which are able to *fulfil* the conditions *to* life, and thus keeps up the average *fitness* to the conditions *of* life' (my emphasis).

Representational Teleology, the Argument from Design, and Natural Selection

The classical *argument from design* posited an external representational explanation for the structure and behavior of organisms: God designed organisms like we design a clock, and His desire to make just that organism explained the existence of particular organisms. Likewise, specific features of the organism (such as valves in the veins) could be explained in terms of their design function, i.e., by their intended contribution to the overall effect they were designed to produce (life). That God's plans worked just as intended – that intention and actuality coincided – could not be doubted. After all, an omniscient and omnipotent Creator could not have failed to achieve his purpose, or only accidentally hit upon something that works.

We know now that organisms were not designed, they have evolved. The existence, structure, and behavior of organisms are thus to be explained by the evolutionary processes that led up to it. But this knowledge leaves the biologist in a real muddle over teleology. Did evolution somehow eliminate teleology from biology? Or did it naturalize it? Does it make sense to talk of function (or even purpose) without a designer?

Many people, beginning with Darwin himself, have tried to answer this question by metaphorically comparing natural selection to the designer. Thus one might say that the valves in the veins were put there by natural selection. This point of view is common in biology and biological philosophy, where an adaptation or function is usually defined as the product of past natural selection. For example, in his recent book on *Darwin and Design* Michael Ruse argues vigorously for this point of view: 'The metaphor of design, with the organism as artifact', he states, 'is at the heart of Darwinian evolutionary biology' (2003, 266). I agree – but while Ruse embraces this metaphor, I see it as a serious problem.

Metaphors may have heuristic value in science, but are dangerous if taken literally. This is especially true of the metaphor of intentional

design when used for natural systems. There is nothing in nature that corresponds to the designer of artificial systems; nature has no intentions. Instead, the teleology of nature corresponds only to the *conditional teleology* involved in an intentional act. There is no intention or design, but given a result (e.g., survival), the conditions necessary for that result clearly happened to occur, and can be used to explain that result. Legitimate (conditional) teleological analysis in biology consists of the following: we see some result, and we try to understand that result by the conditions that make it possible. When we see a living organism; we try to understand its life by the conditions that make that life possible. One of these conditions, perhaps, is the presence of valves in the veins. These valves (or rather, their effects) can thus be understood as conditions for the existence of the organism.

The Design Metaphor, Natural Selection, and Teleological Determinism

In discussing the modes of explanation above I pointed out that a false, teleological determinism results when one explains a result based on deterministic laws, without being able to derive the result from those laws. I believe that the metaphor of design by natural selection has fostered such a false determinism in biology. While natural selection is indeed a mechanism of evolution, one can't just ascribe achieved results of the evolutionary process to natural selection. If one does so, natural selection becomes by implication a representational teleological agent in Darwinian evolutionary theory, aiming toward the achieved results.

Why has this been permitted to occur? To understand this, we must look at the other side of the classical *argument of design*. Not only can one explain organismal features *from* God's *design*, or intention (the heart was put there to pump blood); one can argue from the order, end-directedness, adaptedness to function, and beauty of the world *to design*, and thus to a designer. Order, end-directedness and adaptedness appear to be empirical features of the world, and as such, to call for some explanation. The customary explanation for these organismal properties is of course a historical process of natural selection, viewed as in some way analogous to intentional design. It has been the assumed necessity for a historical explanation for these features that has given natural selection its teleological role in evolutionary theory.

However, the argument to design fails, even without natural selection. This is because the problem that the argument to design – and Darwin – set out to solve does not really exist. In particular, the argument ignores the fundamental difference between artifacts and organisms. The teleology of artifacts like a clock is externally imposed, and relative to the goal of some assumed agent. Given the goal, conditional teleology can explain how that goal is realized, but it will not get you the goal itself. On the other hand, the conditional teleology of organisms is intrinsic, in that the only 'goal' in evidence is the continued existence of the system (in a sense examined more closely in Reiss 2005). When the only end we posit is continued existence, or survival, the order, end-directedness, and adaptedness to function of the world can be explained not as the product of design, but as a chance result that is merely a condition for its continued existence. This has been recognized by anti-teleologists since Empedocles and Epicurus. What is most surprising about the historical path that led to the mess we are currently in is that Darwin's 'natural selection' stood in for the creationists' 'Designer', rather than the anti-teleologists' 'chance' subject to the conditions for existence.

The Conditions for Existence and Teleological Explanation in Biology

How, then, is Cuvier's principle – developed in a non-evolutionary context – relevant to present day biology, which is evolutionary? Spencer's passage quoted above gives us a clue. If Cuvier was wrong in thinking that evolution was impossible because the functional integration of the organism was such that any change in one part would render it unable to satisfy its conditions for existence, he was right that any evolutionary change must be subject to these conditions. The conditions for existence of organisms are in fact part of the range of boundary conditions within which the evolutionary process must occur. The unadapted can't satisfy their conditions for existence and die, the adapted satisfy their conditions for existence and live. Adaptedness from this perspective is not a product of evolution, it is a condition for evolution. In this broad sense, then, the principle of natural selection may be considered identical with the principle of the conditions for existence.

However, lest we carry this identification too far, it is important to note another, narrower sense of the principle of natural selection, which has to do with the conditions for the continued existence of

traits in populations. This sense fundamentally concerns the genetic physiology of populations; the basic postulate is that for a trait to continue to exist, organisms characterized by that trait (in Darwin's view) or the genes underlying the trait (in modern views) must survive and reproduce at a higher rate than alternatives and must do so because of its effects (must be 'naturally selected' in a third and even narrower sense). It is this aspect of Darwin's and Wallace's principle that goes beyond anything in Cuvier's, and provides the basis for much of modern evolutionary theory, from evolutionary stable strategies to sex ratios. Nevertheless, recognizing this narrower sense of the principle does not give us a reason to compare natural selection to a designer. Rather, it invites us to provide conditional teleological explanations for the existence of specific traits. This is in fact what most successful evolutionary explanations have been.

Returning to the question of functional explanation, the principle of the conditions for existence helps to explain why we feel that functions should explain the existence of parts of organisms - they do, given that the system itself continues to exist, and the function helps explain that continued existence (cf. Asma 1996; McLaughlin 2001). The beating of the heart in fact explains its own continued existence in two ways: 1) because if it stopped beating, the organism would die and the heart would too, and 2) because if organisms with beating hearts died more frequently than those without, hearts would disappear from the population (given certain assumptions about their genetic basis). However, this is not the same as saying that organisms can be assimilated to artifacts, where the design function can explain the existence of the part, because the causal basis is entirely different.

In organisms, where only conditional teleology applies, organs and behaviors have functions only to the extent that they do actually contribute to the continued existence of the organism. Many people have felt that functional ascriptions should be normative - that the statement that 'the function of the heart is to pump blood' should imply, for example, that a heart that is not pumping blood still has that function. There is no reason to accept such normativity (see Davies 2001). Normative concepts are derived from intentional teleology, and the goal that was aimed at. Nature doesn't aim at anything, and thus can't miss. Everyone agrees that it makes sense to say that the heart is malfunctioning when a heart attack occurs and the individual dies. However, an arrhythmia is likewise only a malfunction if it kills you - normative concepts based on the average or common condition are not compatible with evolution. Thus for Cuvier normative concepts made sense, for us they don't. Functional

explanation in biology must always be relative to the continued existence of the organism (or some other evolutionary entity - see Reiss 2005).

Nevertheless, there is an important distinction to be made among systems that are amenable to conditional teleological explanation. What makes functional explanations in biology testable is that the systems to which we apply functional explanations are members of classes of similar systems. If we have many organisms with beating hearts, we can stop the heart from beating in one and examine the effect, then apply the knowledge gained to other similar organisms. With less lethal manipulations, the same organism at different times may serve as one's replicate. However, if we had only one organism, we would be severely limited in the methods we could use to establish the functions of parts. This is the situation we are in with respect to the overall evolutionary process, which happens only once. Ayala (1999) distinguished between determinate and indeterminate natural teleology, contrasting the individual development of a chicken with the evolutionary development of a species. On the one hand, these processes can be distinguished as the contrast between a representational system and a non-representational one. However, if we remove considerations of representational teleology (developmental programs, etc.), what distinguishes these processes is primarily that we have multiple examples of chicken development to study, with similar initial conditions, but only one example of the evolutionary process.

Relation to the Literature

In the previous sections I have promoted the distinction between representational and conditional teleology and use of the principle of the conditions for existence as a way to solve some of the persistent problems with teleology in biology. To place my views in context, I here examine some of the most influential previous papers on the subject.

The beginning of the modern debate may be taken as 1953, with the publication of Ernst Nagel's article on 'Teleological explanation and teleological systems'. Nagel, working within the deductive-nomological framework, attempts to show that 'teleological (or functional) explanations are equivalent to non-teleological ones, so that the former can be replaced by the latter without loss in asserted content' (Nagel 1953, 541). He takes as his example 'The function of chlorophyll in plants is to enable plants to perform photo-synthesis'. and considers this

equivalent to the 'non-teleological' statement 'A necessary condition for the performance of photosynthesis in plants is the presence of chlorophyll'. This is rather close to the view based on conditional teleology. The fundamental problem with this example is of course that he does not include the further stipulation that photosynthesis is a necessary condition for the continued existence of the plant itself – he lacked an explicit welfare component to ground the functional ascription.

Nagel argues that the difference between a teleological and non-teleological explanation is that a teleological explanation deals with *consequences* for a system, whereas a non-teleological one deals with '*conditions* under which the system persists in its characteristic organization and activities'. I agree with Nagel that the difference between the two types of explanation he describes 'is one of selective attentions, rather than of asserted content' (541). However, I would argue that the two types of explanations contrasted by Nagel are both conditional teleological explanations, because the 'consequences' of a function are by definition satisfaction of conditions for existence, and the 'characteristic organization and activities' of the organism are likewise simply living, or existence. Nagel fails to see this because he doesn't sufficiently recognize the principle of the conditions for existence itself, and fails to distinguish between absolute and hypothetical necessity.

This failure is seen clearly in Nagel's interesting attempt to put Boyle's law teleologically: 'Gases at constant temperature under variable pressure alter their volumes in order to keep the product of pressure and volume constant.' The biggest problem for this statement is that the situation is deterministic, not conditional; it involves absolute necessity. This is why teleological explanation sounds strange here. As noted above, in a deterministic system operating according to known laws, there is nothing left for the end to explain that has not already been explained by the boundary conditions and the relevant laws. Moreover, as Nagel himself later recognizes, there must be some reference point for conditional teleological explanations; they must be grounded in the real existence of some particular system. Only if in nature Boyle's law only sometimes held, and when it did it allowed some system to exist that otherwise would not, might it be reasonable to ascribe a function, if not a purpose, to the varying volume of the gas.

Nagel, on the other hand, sees two ways that his translation of Boyle's law is not that different from explanations in the physical sciences. First, he has recourse to the use in physics of 'isoperimetric' or 'variational' forms of physical laws, such as the principle of least action, and argues that these are 'akin' to teleological explanations. In

spite of their apparent reliance on future causation, however, these are completely deterministic explanations and thus not equivalent to teleological ones.

Nagel's second point is of more interest. He notes that 'the physical sciences unlike biology are in general not concerned with a relatively special class of organized bodies, and they do not investigate the conditions making for the persistence of some selected physical system rather than of others' (543). He here hits upon the principle of the conditions for existence: it is only because we are interested in the persistence, or continued existence, of organisms that we apply functional teleological explanations to them. Of course, the reason that we as scientists are interested in their persistence is that nature is too – the continued existence of life is a real phenomenon.

The remainder of Nagel's article is taken up with goal-directed or purposive systems; he attempts to ground functional ascriptions in the fact that organisms are such purposive systems. As noted above, any such attempt confuses the purposive or representational teleology which may be displayed by actions of the system itself with the conditional teleological (functional) explanation of those actions in terms of the continued existence of the organism.

An equally influential contribution to the debate appeared in 1959 with the publication of Hempel's article on 'The logic of functional analysis' (republished with some alterations in his 1965 book). Hempel is concerned exclusively with natural functions, not artifactual or intentional ones. For Hempel,

> the kind of phenomenon that a functional analysis is invoked to explain is typically some recurrent activity or some behavior pattern in an individual or a group, such as a physiological mechanism, a neurotic trait, a culture pattern or a social institution. And the principal objective of the analysis is to exhibit the contribution which the behavior pattern makes to the preservation or the development of the individual or the group in which it occurs. Thus, functional analysis seeks to understand a behavior pattern or a sociocultural institution by determining the role it plays in keeping the given system in proper working order or maintaining it as a going concern (Hempel 1965, 305).

The emphasis on the role of the part in 'preservation', 'development', and 'maintaining as a going concern' the overall system is a clear recognition of conditional teleology and the principles of the conditions for existence. The notion of 'proper working order', on the other hand, appears dangerously normative, and thus purposive. After all, if what is proper is not intended, but is grounded in continued existence itself, then 'proper working order' can mean only that

continued existence, and does not differ from 'maintaining it as a going concern'.

Hempel next introduces his classic example (examined in almost all subsequent literature): 'The heartbeat in vertebrates has the function of circulating blood through the organism' (305). He considers whether function might just mean 'effect', but rejects this because we would then have to accept 'The heartbeat has the function of producing heart sounds.' He objects that

> a proponent of functional analysis would refuse to assert [this], on the ground that heart sounds are an effect of the heartbeat which is of no importance to the functioning of the organism; whereas the circulation of the blood effects the transportation of nutriment to, and the removal of waste from, various parts of the organism – a process that is indispensable if the organism is to remain in proper working order, and indeed if it is to stay alive.

Based on this concept, the import of the functional statement to 'a proponent of functional analysis' might be understood as: 'The heartbeat has the effect of circulating the blood, and this ensures the satisfaction of certain conditions (supply of nutriment and removal of waste) which are necessary for the proper working of the organism.' If in place of 'proper working' we substitute 'survival' this is in fact an expression of the viewpoint of conditional necessity and the conditions for existence.

Hempel goes on to critique this viewpoint, however. He notes that many other conditions are required for survival, including both those of organisms and environment. Our inability to measure relevant variables and to construct relevant laws means that the system is not deterministic and lawlike (bad for Hempel, working within the deductive-nomological framework). He also objects that we need criteria of 'proper working' and 'normal functioning', though he later notes that his own criterion of survival will in fact serve here (323).

Is functional analysis explanatory? If so, it what sense? Hempel is particularly interested in the assertion that functional ascriptions explain the *presence* of the *part* in the system. He takes as representative the following (bad) argument (here applied to his example):

1) the organism exists (in a setting of a certain specified kind),
2) the organism can exist only if its blood is circulating,
3) the heart, if present, circulates blood
4) therefore the heart must be present in the organism.

As Hempel notes, this explanation is subject to the counterexample of functional equivalents – things other than a heart can circulate blood. Hempel thus concludes that functional explanations are weak at best. I agree with Hempel that this argument is invalid, even though many examples of such thinking can be found in the literature. The problem is that Hempel (and the authors he is criticizing) is here trying to make a functional explanation, which is necessarily a conditional teleological one, into a deterministic one. As we have seen, only in the case of artifacts can one argue directly from the (design) function to the existence of the part, because when the function is an intended one, the part is indeed put there for its intended function.

It may be helpful to recall how functions are established. We examine an organism and note the activities or effects of some part. We observe that the heart is beating and ejecting a certain volume of blood at a certain pressure with each beat. We then try to understand in what way this effect might be a necessary condition for the continued existence of the organism (as it actually exists). This is a functional hypothesis. We can test this hypothesis in several ways, for example by trying to eliminate that effect, while retaining all others.

Given this empirical background to functional ascriptions, what sort of argument is a functional explanation? We have already examined this above; it is a conditional teleological argument. One might put this argument in the following way for Hempel's example:

1) the organism existed (from time t_0 to t_1)
2) blood circulation was a necessary condition for the continued existence of the organism (from time t_0 to t_1)
3) therefore, since the organism existed, something must have been circulating the blood
4) the heart was (empirically) the thing that was circulating the blood.

The existence of the part is not explained by its function, but rather the observed existence of the whole is explained as contingent on the function of the part. Thus the possible existence of functional equivalents is not an issue; if something else replaces this activity of the heart, then one simply has a distinct explanation. While such an explanation does not give a deterministic explanation for the presence of the heart, it does make it clear how the existence of the heart makes the existence of the organism possible, and thus makes possible its own existence.

Hempel in fact considers such a meaning to functional explanation, but regards it as weak or trivial: functional analysis leads 'only to the

conclusion that in some way or other, the blood is being kept circulating properly ... at *t* – hardly a very illuminating result' (314). But the illuminating result is in fact the necessity for blood circulation, and the conditional chain that links blood circulation by the heart to the continued existence of the organism. This is what functional explanation is all about.

For Hempel, however, a deterministic explanation is preferred, and this leads him to consider purposive (what he calls 'self-regulating') systems. Interestingly, he considers these relatively unproblematic cases in which a lawlike explanation is possible; he does not see that these explanations, too, must be conditional if the underlying laws of the system are unknown. Ultimately, for Hempel, like Nagel, the key scientific feature of functional explanations is that they are invoked for self-regulating (or purposive) systems, and the hope is that empirical generalizations may eventually lead to 'a more general theory of self-regulating systems'. Thus while Hempel rather accurately understands that functional explanation can be a mode of conditional teleological explanation, he rejects this mode as unsatisfactory because it does not fit his deductive-nomological model of how science should work: he wants to explain the world on the basis of deterministic (or at least probabilistic) laws, not conditional relations.

In 1973 Larry Wright published his influential 'etiological' or 'selected effects' (Neander 1991) view of functional explanation. Wright explicitly wants to say (as did Hempel) 'X is there *because* it does Z' (Wright 1973, 157). However, unlike Hempel, Wright considers this explanation valid, both for artifacts and parts of artifacts, and for parts of organisms. Essentially this is because he takes intentional design (external representational explanation) as the model for all functional explanation: 'conscious and natural functions are functions in the same sense' (143).[6] He begins with the distinction between function and accident. For him, function is grounded in conscious design as efficient cause: 'intention is so central here that it allows us to say the function of I is to do C, even when I cannot even *do* C' (146). 'X is there *because* it does Z' means 'at least some effort was made to get X (sweep hand, button on dashboard) where it is precisely because it does Z (whatever)' (158). How can such intentional explanation apply to natural functions? Wright invokes natural selection:

[6] Of course, I agree with this statement if applied to conditional functions of artifacts, but Wright is interested in design functions.

> We can say that the natural function of something – say, an organ in an organism – is the reason the organ is there by invoking natural selection. If an organ has been naturally differentially selected for by virtue of something it does, we can say that the reason the organ is there is that it does that something. Hence we can say animals have kidneys *because* they eliminate metabolic wastes from the blood-stream; porcupines have quills *because* they protect them from predatory enemies; plants have chlorophyll *because* chlorophyll enables plants to accomplish photosynthesis; the heart beats *because* its beating pumps blood (159).

When does this natural selection occur? In the past. This is necessary for Wright to maintain the distinction, fundamental to intentional teleology, between functions (intended) and accidents (unintended). In noting that mutations are accidental, he claims that 'that only disqualifies an organ from functionhood for the first – or the first few – generations. If it survives by dint of its doing something, then that something becomes its function on this analysis' (165). In other words, useful actions must be baptized by natural selection before they become functions.

The popularity of this view is perhaps due to its apparent simplicity and plausibility; for many it may be intuitively satisfying. However, by comparing natural selection to the designer of artifacts, Wright is rather clearly making the precise mistake noted above: he is ascribing intentions to nature (see Davies 2001). By ascribing achieved results to a selective process he knows only by these results, he is implicitly supplying a goal for the process. Thus whereas he thinks he is providing an etiological (i.e., causal) analysis, he is in fact providing a teleological deterministic one. His enthusiasm for this view even leads him to endorse the parallels between natural and intentional selection (162-164) because both are based on consequences. He fails to note that natural selection is based on real consequences, whereas intentional selection is based on one's conscious concept of consequences within a model of the world. The past history of selection may account for how porcupines got quills in the first place, but it does not explain this for us if we don't know the details of the selective history. Moreover, it doesn't have anything to do with the current function of the quills, except in the sense that the quills probably had pretty much the same function in previous generations of porcupines as they do today.

But though I have been rather hard on Wright for his teleological determinism, his views are in fact compatible with mine given a few alterations. Most importantly, there is no reason to wait several generations, one generation, or even one second in ascribing a

function to something that is being 'naturally selected'. If over the time period from t_0 to t_1 a feature of an organism has a function (satisfies some condition for the continued existence of the organism) then it is being 'naturally selected' (in the broad sense) at that time. Moreover, as noted above, from the point of view of the genetic physiology of populations (of which natural selection in the narrow sense is a part), it is often reasonable to assume that a feature like a beating heart or a porcupine quill would not persist in the population if it did not have a function, i.e. if it did not satisfy some condition for the existence of the organism. Perhaps this is really what Wright was trying to get at. However, it is important to note that here the function comes first, and the continued existence of the trait is explained in terms of that function. In Wright's account the function comes second; it only appears after natural selection has baptized the fortuitous effect generated by mutation.

A very different approach to functions by Robert Cummins appeared in 1975; this has been called the 'causal role' (Amundson and Lauder 1994) or 'systemic capacity' (Davies 2001) theory of functions. Cummins notes that previous analyses of functions had assumed that:

> A. The point of functional characterization in science is to explain the presence of the item (organ, mechanism, process, or whatever) that is functionally characterized.
> B. For something to perform its function is for it to have certain effects on a containing system, which effects contribute to the performance of some activity of, or the maintenance of some condition in, that containing system. (Cummins 1975, 741)

Cummins resolves to question these assumptions. His critique of assumption (A) incorporates many of the same criticisms I have discussed above. In particular, he notes that it is a failure to distinguish 'teleological' (my 'intentional') from 'functional' explanation that makes people assume (A). Only in the case of 'deliberate action' is (A) true:

> chlorophyll and hearts are not 'there' as the result of any deliberate action; hence the essential presupposition of the explanatory move in question is missing. Once this becomes clear, to continue to insist that there must be some sense in which specifying the function of chlorophyll explains its presence is an act of desperation born of thinking there is no other explanatory use of functional characterization in science (747).

Good enough. But Cummins' critique of assumption (B) is obviously more problematic from my perspective. In fact, Cummins explicitly criticizes a form of functional analysis (his form 8) similar to the one supported here:

> (8) The functions of a part or process in an organism are to be identified with those of its effects contributing to activities or conditions of the organism which sustain or increase the organism's capacity to contribute to survival of the species (755).

Aside from the reference to the 'survival of the species' rather than that of the organism or the genetic lineages it carries (Reiss 2005), and the use of 'capacity to contribute' rather than 'contribution' (see below), his form (8) is essentially my position. But although Cummins recognizes that '(8) doubtless does capture a great many uses of functional language in biology', he considers (8) 'seriously misleading and extremely limited in applicability even within biology.' Why? Because when an organism in a new environment carries out an activity that is detrimental, we can still offer a functional analysis of this detrimental activity. He gives the example of pigeons for which flying ceased to contribute to the maintenance of the species. He claims that we could still say the function of these poor pigeons' wings is to fly. This is an interesting example. Let us look at it a bit more closely.

As discussed above, conditional functional explanation must always be relative to an existing organism (or other evolutionary entity). Therefore, we must think of what the actual pigeons are actually doing. Cummins suggests that perhaps 'in the absence of serious predators, with a readily available food supply, and with no need to migrate, flying simply wastes energy' (755). Thus we must imagine pigeons that really are flying around, but would save energy if they didn't.

Now let us imagine a population composed of two types of pigeons, one which flies a lot, and one which always walks. And let us suppose that the walking pigeons survive at a higher rate and have more offspring, because they're not wasting their energy on flight (per Cummins' hypothesis). We can imagine two different versions of this scenario. On the one hand, the flying pigeons might fly by periodically leaping into the air, flapping around, and settling back down again, walking the remainder of the time. If we knew somehow that this behavior had no contribution to the survival or reproductive success of the pigeons, I maintain that we would not (or should not) call flying a function of the wings in these pigeons, because flying is

in no way a necessary condition for their survival (or that of their genes). That this situation seems implausible is due to the fact that we do not expect a maladaptive trait with no function to be maintained in a population.

On the other hand, consider the situation in which one group of pigeons uses their wings to fly from their nest to the feeding grounds, while another simply walks. That such a situation might arise is not implausible. Here, for the pigeons that fly not to fly, when flying is part of their integrated lifestyle, would obviously not be good for them – it would leave them with no way to get from their nest to the feeding grounds. In this situation I would maintain that a function of the wings in the birds that fly is to fly, because they are in fact surviving by means of flight as part of their lifestyle. By contrast, the wings of the non-flying birds do not have the function of flying, because they never fly with them. Cummins' mistake here is to think that a welfare view commits one to establishing functions in a real flying bird relative to a non-flying bird with an alternative behavior, rather than simply to the same bird lacking the behavior. If functional explanations are relative to specific existing organisms, not real or hypothetical alternatives, the problem never arises.[7]

Cummins' other criticisms of (8) are even weaker. For example, he notes that (8) doesn't address the style of functional explanation, but 'simply identifies effects which, as it happens, are typically explained in that style' (756). This is true, but delimiting the effects which can be called functions is a critical part of functional analysis. This of course depends on an objective standard of value. Cummins denies that such a standard exists, whereas I have argued that the continued existence of the organism is such a standard.

For Cummins, however, it is indeed the style that is distinctive of functional explanations. He views functional explanations as examples of an 'analytical strategy', in which a disposition, or overall 'capacity', is broken up into a number of subsidiary 'capacities'. He gives the example of the parts of an assembly line in relation to the overall product of the assembly line, and schematic diagrams in electronics, which indicate parts by their capacities. He concludes that 'Functional analysis in biology is essentially similar. The biologically significant capacities of an entire organism are explained by analyzing the

[7] However, when we speak of the degree to which the behaviour is *adapted* to the external and internal environment, we are necessarily talking of success relative to specified alternatives, thus a feature can have a function without being the best adapted possibility that one sees or can conceive.

organism into a number of 'systems' – the circulatory system, the digestive system, the nervous system, etc. each of which has its characteristic capacities. These capacities are in turn analyzed into capacities of component organs and structures. Ideally, this strategy is pressed until physiology takes over - i.e., until the analyzing capacities are amenable to the instantiation strategy' (760-761).

What Cummins is here focusing on is what I have called conditional (functional) teleological explanation, where parts are explained as conditions for the whole. However, Cummins does not clearly separate conditional from deterministic explanation. Moreover, for reasons that are not clear to me, Cummins introduces dispositions and capacities into the account of such conditional explanation. He believes that capacities are equivalent to dispositions, and dispositions are law-like properties amenable to standard scientific analysis.

Unfortunately, the equation of dispositions and capacities is problematic. Cummins' 'analytical strategy' is to break up an overall disposition into subsidiary dispositions, such that 'programmed manifestations of the [subsidiary dispositions] results in or amounts to a manifestation of the [overall disposition]' (759). However, this describes a deterministic, reductionistic explanation of the overall disposition. If such an account is available we are unlikely to apply functional analysis at all. And if such an account is not available, if we can't specify the conditions under which the dispositions will manifest themselves, then an account in terms of dispositions involves a false, teleological determinism.

On the other hand, to break up an overall capacity into subsidiary capacities does not have this problem, because when we speak of capacities we implicitly mean something that need not occur (as far as we can specify conditions), but is merely able to occur (cf. Cummins 760, note 17). Capacities are thus compatible with conditional teleological explanation, and this is perhaps why Cummins introduces them. However, it remains unclear to me that this extra layer of capacities adds anything to our explanation. To say that the heart having the capacity to beat explains the capacity of the organism to exist rather than that its beating explains the organism's existence seems at best to involve a lot of extra verbiage. Moreover, though if the heart beats it clearly has the capacity to beat, if it has the capacity to beat it doesn't necessarily beat, and what we really want to explain is not the capacity, but the actuality. Cummins' account thus fails in two ways: 1) it doesn't recognize the real existence of organisms as a valid overall end in nature, by which we can ground our function statements, and 2) it only explains an organism's capacity to exist, not

its actual existence. Nevertheless, Cummins' views (see also Davies 2001) have been quite useful to the ongoing debate as a reminder that the 'etiological' view has its problems.

Although there have been many analyses of biological teleology in the years since Cummins' paper (Allen *et al.* 1998), most have followed in the footsteps of Wright or Cummins (for some notable exceptions, see Ghiselin 1974; Achinstein 1983; Nissen 1997; Asma 1996; McLaughlin 2001). Wright's views have been the most influential, and perhaps may be considered the 'mainstream' viewpoint today. For example, Michael Ruse (2003) has recently summarized his own previous work in a book length treatment of *Darwin and Design*. For Ruse, the argument to design is a valid one, though we now know that natural selection, not God, is the designer of organisms. However, natural selection is only a designer in a metaphorical sense, because it is of course a non-intentional mechanism of evolution. In organisms we are faced with 'organized complexity' and

> this complexity allows for and indeed calls for understanding in human terms of intentionality, of purpose, of design. Organized complexity is artifactual. That was the whole point of natural theology – the argument to design – and the moment and place where Darwin took over. Whether or not organisms really are designed, thanks to natural selection they ... seem as if designed (for the ends of survival and reproduction) (265). The metaphor of design, with the organism as artifact, is at the heart of Darwinian evolutionary biology (266).

What Ruse doesn't tell us is how a metaphor can be explanatory in a scientific sense. If we accept the metaphor as real, then we must attribute an intention (or at least a purpose) to natural selection, because in human design it is the intention that acts as one of the efficient causes. On the other hand, if it is only a metaphor then what can it explain? Ruse claims that the similarities between intentional design and natural selection 'stimulate you to see the similarities, and perhaps even more than there were at first. So a new truth emerges, one that is metaphorical at first but moves toward the literal' (265). However, as noted by Nissen (1997) with reference to an earlier statement of Ruse's views, 'If teleological language in biology were metaphorical, as Ruse asserts, then *it could not also be explanatory*, which he also asserts' (104). Or to put this in more positive terms: only to the extent that our teleological explanations of organismal structure and function do *not* depend on the explicit or implicit existence of representations of goal states in the world can they be truly explanatory.

Conclusions

In this paper I have tried to show that the distinction between representational and conditional teleological explanation is fundamental, and can help resolve many of the difficulties in the literature over the nature of teleological explanation. Most critically, I have argued that to conceive of past selection as the efficient cause of organismal traits leads necessarily to a teleological (and implicitly representational) determinism; there is simply no basis for the analogy between natural selection and an intentional designer. In biology, continued existence, survival, or simply life is the phenomenon we are trying to explain; we can explain it only through a conditional teleology, since determinism fails us and we reject intentional design (external representational teleology). Thus functions of parts of organisms should be understood as Cuvier understood them, as conditions for continued existence of those organisms. I can only hope that I have indeed helped clarify matters, and not just introduced a new source of confusion into the ongoing debate.

Acknowledgments

This paper derives from intellectual influences too numerous to list in detail, though I would like to mention Steve Gould and Leon Croizat, who first awakened me to the problem of biological teleology, and E.S. Russell, who introduced me to Cuvier. I would also like to thank Humboldt State University, which provided a sabbatical leave allowing me to indulge in reading and thinking, and my wife, Karen Reiss, who has put up with me and my obsession for many years. Helpful comments on the manuscript were provided by Peter McLaughlin, Michael Ghiselin, Karen Reiss, and several anonymous reviewers.

References

Achinstein P., 1983, *The Nature of Explanation,* New York: Oxford University Press.

Allen C., Bekoff M., Lauder G. (eds), 1998, *Nature's Purposes: Analyses of Function and Design in Biology*, Cambridge, MA: MIT Press.

Amundson R., Lauder G., 1994, 'Function Without Purpose: The Uses of Causal Role Function in Evolutionary Biology', *Biology and Philosophy*, 9: 443-469.

Appel T.A., 1987, *The Cuvier-Geoffroy Debate: French Biology in the Decades Before Darwin,* New York: Oxford University Press.

Asma, S.T., 1996, *Following Form and Function: A Philosophical Archaeology of Life Science*, Evanston, ILL: Northwestern University Press.
Ayala F.J., 1999, 'Adaptation and Novelty: Teleological Explanations in Evolutionary Biology', *History and Philosophy of the Life Sciences*, 21: 3-33.
Beckner M., 1967, 'Teleology'. In: Edwards P (ed.), *The Encyclopedia of Philosophy*, New York: Macmillan, 8: 88-91.
Boylan M., 1986, 'Monadic and Systemic Teleology'. In: Rescher N. (ed.), *Current Issues in Teleology*, Lanham, MD: University Press of America, 15-25.
Coleman W., 1964, *Georges Cuvier: Zoologist,* Cambridge, MA: Harvard University Press.
Corsi P., 1988, *The Age of Lamarck: Evolutionary Theories in France 1790-1830,* Berkeley: University of California Press.
Cummins R., 1975, 'Functional Analysis', *Journal of Philosophy*, 72: 741-765.
Cuvier G., 1800, *Leçons d'anatomie comparée,* Paris: Baudouin.
Cuvier G., 1817, *Le règne animal distribué d'après son organisation, pour servir de base à l'histoire naturelle des animaux et d'introduction à l'anatomie comparée,* Paris: Deterville.
Davies P.S., 2001, *Norms of Nature: Naturalism and the Nature of Functions*, Cambridge, MA: MIT Press.
Ghiselin M.T., 1974, *The Economy of Nature and the Evolution of Sex*, Berkeley: University of California Press.
Hempel C.G., 1965, 'The Logic of Functional Analysis'. In: Hempel C.G. (ed.), *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, New York: Free Press, 297-330.
Kant I., 1981, *Universal Natural History and Theory of the Heavens*, Edinburgh: Scottish Academic Press.
Lee R., 1833, *Memoirs of Baron Cuvier,* New York: J. & J. Harper.
Letteney M.J., 1999, *Georges Cuvier, Transcendental Naturalist: A Study of Teleological Explanation in Biology*. Ph.D. Dissertation, University of Notre Dame.
McClellan C., 2001, 'The Legacy of Georges Cuvier in Auguste Comte's Natural Philosophy', *Studies in the History and Philosophy of Science,* 32: 1-29.
McLaughlin P., 2001, *What Functions Explain: Functional Explanation and Self-Reproducing Systems,* Cambridge: Cambridge University Press.
Nagel E., 1953, 'Teleological Explanation and Teleological Systems'. In: Feigl H. and Brodbeck M. (eds), *Readings in the Philosophy of Science*, New York: Appleton-Century-Crofts, 537-558.
Neander K., 1991, 'Functions as Selected Effects: The Conceptual Analyst's Defense', *Philosophy of Science*, 58: 168-184.
Nissen L., 1997, *Teleological Language in the Life Sciences,* Lanham, MD: Rowman and Littlefield.
Outram D., 1984, *Georges Cuvier: Vocation, Science and Authority in Post-Revolutionary France,* Manchester: Manchester University Press.
Randall J.H., Jr., 1960, *Aristotle,* New York: Columbia University Press.
Reiss J.O., 2005, 'Relative Fitness, Teleology, and the Adaptive Landscape', *Evolutionary Biology,* 34, (in press).
Rosen R., 1985, *Anticipatory Systems: Philosophical, Mathematical and Methodological Foundations*, Oxford: Pergamon.
Ruse M., 2003, *Darwin and Design: Does Evolution Have a Purpose?* Cambridge, MA: Harvard University Press.

Russell E.S., 1916, *Form and Function: A Contribution to the History of Animal Morphology,* London: John Murray.
Spencer H., 1864, *The Principles of Biology,* New York: D. Appleton and Co.
Woodfield A., 1998, 'Teleology', In: Craig E. (ed.), *Routledge Encyclopedia of Philosophy*, London: Routledge, 9: 295-297.
Wright L., 1973, 'Functions', *Philosophical Review*, 82: 139-168.