1988

Comparative Psychology and the Recursive Structure of Filter Explanations

Nicholas S. Thompson

The International Journal of Comparative Psychology, Vol. 1, No. 4, Summer 1988

FOCAL ARTICLE

COMPARATIVE PSYCHOLOGY AND THE RECURSIVE STRUCTURE OF FILTER EXPLANATIONS

Peter Lipton
Williams College

N.S. Thompson
Clark University

ABSTRACT: Over the last three decades many critics have worried that the theory of adaptation by natural selection is a circular explanation and has no empirical content. We argue, however, that the theory is not circular but recursive. In a recursive explanation the explanandum serves to specify which of a class of potential explainers is being appealed to. Many explanations thought to be circular are in fact recursive, including the classical exemplar of a circular explanation, Moliere's *virtuus dormitiva*. Natural selection theory is a recursive explanation in which the thing to be explained serves to specify a particular Darwinian filter. Unlike circular explanations, recursive explanations can have considerable explanatory value. That value depends somewhat independently on the two parts of a recursive explanation, its explanatory frame and its recursion. The frame of a Darwinian explanation is valuable because it tells us that the trait in question is a result of a Darwinian filter. For instance, to say that the polar bear's fur is white because of natural selection is informative. It says that in the past there were bears of various colors and that over the generations only white bears have passed through the filter until now all bears are white. The value of the recursion of a Darwinian explanation depends on the characterization of the explanandum that occurs in the recursion. An overspecified explanandum leads to an ad hoc Darwinian explanation. An underspecified explanandum leads to a vague Darwinian explanation. While both of these are strictly speaking explanatory, they lack the heuristic value that would earn them the respect of working scientists. To have heuristic value, a Darwinian filter explanation requires characterization of the explanandum in terms of intermediate generality. This analysis suggests that the future of Darwinian theory depends on revival of systematic comparative description.

Comparative Psychology may be thought of as the discipline where two bodies of theory potentially interact: learning theory and Darwinian theory of evolution. Many commentators have observed that these two theories have the same general structure (Nozick, 1974, pp. 18-22, 312-317; Bateson, 1979; Thompson, 1981). Both are filter theories. In

Address correspondence to N.S. Thompson, Clark University, Department of Psychology, 950 Main St., Worcester, MA 01610, USA.

both, a current feature of the behavior or structure of organisms is explained by claiming that this feature has been filtered from a range of features produced by the organisms at an earlier time.

Since the two theories have a similar structure, it would also be expected that they would have similar logical problems. One such problem shared by the two theories is that both are frequently attacked for their alleged circularity. The charge that an explanation is circular must be taken seriously because circular explanations do not in fact explain. They merely restate the question they claim to answer. They are particularly pernicious in science because they give a false sense of understanding and prematurely put a stop to scientific inquiry. Because learning theory and evolutionary theory are so central to comparative psychology, the discovery that both are logically circular would be a devastating discovery for theoretical development in the field.

In this essay, we examine the alleged circulatory of one of the two theories, Darwin's theory of evolution by natural selection. Our general conclusion is that while Darwinian explanations are rarely circular even in the most careless hands, they do have a peculiar logical structure that makes them both easy to confuse with circular explanations and vulnerable to other explanatory hazards. We will not treat learning theory in detail. However, because learning theory is also an example of a filter explanation, our comments about the logical vulnerability of Darwinian theory will have relevance to learning theory as well.

INTRODUCTION

In the central theorem of Darwinism an explanandum state, adaptation, is explained by an explanans process, natural selection. The terms, 'explanandum' (= the thing one is trying to explain) and 'explanans' (= the thing that is doing the explaining) are standard philosophical terminology. Because one of the authors (N.S.T.) is forever confusing them, we will back up our usage of these terms with the equivalent expressions, wherever we can do so without making the text impossibly wordy.

Adaptation, the thing that requires explanation, is the suitability of organisms to their circumstances. Organisms, unlike most of the other natural objects of our world, have the property of being more or less well-designed for the worlds in which they live. In the context of a scientific materialism that denies the possibility of intentional creation, this semblance of designedness in organisms is anomalous.

Natural selection, the thing that explains the anomaly, is the differential reproduction of organisms bearing one type of heritable trait by comparison with those that bear its alternative. Natural selection theory reconciles the apparent designedness of organisms with scientific materialism by making an analogy to the manner in which elaborate breeds of pigeons are produced by pigeon fanciers. Just as the circumstances of the pigeon coop assure that those pigeons with the elegant feathers, flight patterns or courtship routines have the most offspring, so the circumstances of nature assure that in general those organisms which are better adapted to their circumstances will have more offspring than those less well adapted. And if such a reproductive bias continues, either in the pigeon coop or in nature, then it assures in time that pigeons will have elegant feathers, or flight patterns or courtship routines and that organisms in nature will be better designed for their circumstances. Thus adaptation exists for the same reasons that elegance in pigeons exists: because it corresponds to an aesthetic bias in the life environment of the organism (Rindos, 1984).

The putative circularity arises in the following way. The two concepts linked by Darwinian theory, natural selection and adaptation, have very different strengths and weaknesses. While there is a wide agreement that organisms possess the property of adaptation, there is no widely accepted way to define that property. It is, in fact, an exceedingly complex property. To say that an organism is well-designed for a set of circumstances, is to say that being like this organism facilitates the solution of some characteristic "problem" posed by these sorts of circumstances (Dunbar, 1982). Revealing this property of organisms is the province of comparative anatomy and bioengineering (Thompson, 1981; Lewontin, 1978; Williams, 1966; Sommerhoff, 1950). We can grasp it intuitively because it is very like the manner by which we judge the functions of ordinary objects and the intentions of people and animals. Because we can grasp adaptation intuitively, we can easily point to good examples of it in nature.

But to say that something can be intuitively grasped and readily pointed to is a long way from saying that it is sufficiently well-defined for scientific purposes. While intuition is sufficient to lead us to the good examples of a phenomenon, it is rarely sufficient to sort out the borderline cases in a manner consistent with modern scientific practice. Scientific concepts must be reliable and explicit in their application, and unelaborated intuitive criteria cannot pass this test. Moreover, adaptation as it is currently understood is a mathematically intractable concept. No measures have been agreed upon by which we can say how well or how poorly adapted an organism is to particular circumstances. Consequently, of adaptation, we can say that it is readily exemplified, but poorly defined and mathematically intractable.

Natural selection has many of the complementary virtues and vices. Unequivocal instances of natural selection are difficult to point to. It is rarely shown that a particular well-designed property of an organism has come about through a history of differential reproduction of better designed variants (Hailman, 1982). But natural selection is readily defined and its mathematics has been thoroughly developed. Consequently, of natural selection we can say that it is difficult to exemplify, but that it is well-defined and possesses a tractable and powerful mathematics.

Under these circumstances it is understandable why modern science, with its fondness for precise definition and mathematical modeling should be tempted to define adaptation in terms of natural selection. Such a redefinition occurs when the better-adapted organism is defined as the one that has more than its share of offspring. Surprising as it may seem, adaptation is frequently redefined in this way in the literature. In fact, the redefinition is deeply entrenched in classical literature of population biology, for which the central term, "fitness" refers not to some property of appropriateness or suitability to the environment but simply to the relative number of offspring of an organism.

The redefinition has the advantage that it supplies each of the key evolutionary concepts, natural selection and adaptation, with something it lacks. To natural selection it supplies a method of pointing to cases of natural selection at work: each of the obvious cases of adaptation in nature becomes an unequivocal case of a product of natural selection. To adaptation, it supplies a clear definition and access to a sophisticated mathematics: each unequivocal case of adaptation can be precisely defined and analyzed using natural selection mathematics.

The redefinition has the disadvantage that it opens evolutionary theory to accusations of circularity. If the adapted animal is defined as the one that produces the most offspring, then the claim that all animals that are adapted are favorably selected is analytically true and beyond empirical validation. To search for cases to validate the statement that all animals that are selected are adapted under these circumstances would be like conducting a door-to-door survey of bachelors to discover if they are unmarried. The hypothesis will always be validated.

Over the last two decades, a vigorous argument has been conducted over whether this allegation of circularity is well founded, whether the alleged circularity is pernicious and whether, if pernicious, it is avoidable. Some have argued that it is not pernicious. Karl Popper once (1972, as cited—and recanted!—in Popper, 1978) argued, for instance, that the circulatory is not pernicious because Darwinian theory is not an explanation but a "metaphysical research programme" (p. 344) which should be evaluated not for its truth or falsity but for its capacity to help the scientist generate propositions whose truth can be tested. This response, while safe enough, seems monstrously to underestimate the accomplishments of Darwin. The fact is that Darwin's theory constituted an alternative and better explanation of adaptation than the theories of its day. To relegate the modern version of that theory to a heuristic world view is to set the growth of knowledge back a century.

Some have argued that the putative circularity is the result of a confusion between what adaptation actually is and how it can be known. Rosenberg (1980b), for instance, makes a very persuasive argument for the logical distinctness of adaptation and natural selection. But he insists that adaptation is beyond any reasonable possibility of direct scientific treatment and that we must therefore use differential reproduction as the "only practicable general means" of assessing the degree of adaptation of an organism. This argument seems like the scientific equivalent of using the reticence of the suspect as a rationale for torture. It suggests that the greater our practical need to reason circularly, the greater is our justification in doing so.

Some have argued that the circularity is pernicious, but avoidable. Thompson (1981), for instance, points out that the comparative method, a widely respected descriptive method in biology, could form the basis for the specification of adaptation necessary to evolutionary theory and that biologists should work harder at elaborating and formalizing the comparative method as a means of defining and quantifying adaptation (see also Lewontin, 1978).

We now believe that all of these earlier accounts are either incorrect or irrelevant because natural selection explanations are not strictly speaking circular, neither essentially nor as they are practiced. They are, in fact, recursive. The difference is crucial because while circularity in an explanation is a terminal disease, recursion can be healthy and it may even constitute an important preliminary stage in Darwinian explanation. On the other hand, despite the virtues of recursive explanations, they do have their own peculiar vulnerabilities. A theorist who would make use of recursive explanations must understand their limitations or run the risk of making statements that are almost as empty and fatuous as a circular explanations.

THE STRUCTURE OF RECURSIVE EXPLANATIONS

One of the best ways to understand the structure of a recursive explanation is to see how it differs from that of a circular one. In a circular explanation, the explanans is a restatement of the explanandum. For instance, if we answer the question, why are polar bears white? with the reply, "Because they are white!" we have given a circular explanation. Even the reply, "because of their alabaster hue!" would be circular because it is just a dressy way of saying "white." Thus, the explaining concept is simply a restatement of the thing it explains.

In a recursive explanation, the explanans contains a restatement of the explanandum embedded in a larger structure. If the question is, "what makes polar bears white?" then the answer, "Because there is a force in the environment that makes them white!" is recursive, not circular. Even though "makes them white" is a paraphrase of the question, the introduction of the words "force in the environment" embraces the paraphrase in a nonparaphrasing explanation and breaks the logical circle. This recursive explanation does not treat the color of the fur as literally self-explanatory, as does a circulatory explanation; instead it uses the color as a way of specifying something else as the explanation, in this case, the force that brought it about.

Since a recursive explanation does use the explanandum in the explanans, recursive explanations are easily confused with circular ones. Distinguishing them requires a careful comparison of the explanandum with the explanans, that is, of the question with the answer. If the question above had been, "which is the force which makes polar bears white?", then the answer, "The white-making force!" would be circular because that answer is a restatement of that question. In general terms, if a question is of the form, "why is E?", then an explanation is circular if it is of the form, "because of E!", but merely recursive if it is of the form "Because of the cause of E!". A recursive explanation is more informative than a circular explanation just to the extent that we have independent sources of information about the class of causes referred to in the explanation.

What does it mean to say that an explanation is more informative? One way of measuring the information value of an explanation is by the "exclusion test." Better explanations, according to this view, are explanations that exclude more possible causes. Let's imagine that we are trying to explain the "snow" on the screen of our computer. Two explanations are offered us: one, that the snow is caused by a "malfunction in the computer" and two, that the snow is caused by a "glitch in the monitor." Both explanations exclude many causes, for instance, problems in radio interference from outside the computer. Of the two, the second explanation is better because it excludes all the alternative causes that the first explanation excludes and more. For instance, if the "glitch in the monitor" account is correct, we know that we need not look in the central processing unit for the source of our problem.

Circular explanations fail the exclusion test. They are uninformative in that they exclude no alternatives. The claim that the bear is white because it is white is compatible with every possible cause of the bear's fur color. Recursive explanations, by contrast, usually do exclude some alternatives and, by the exclusion test, vary in quality from those that exclude very few alternatives to those that exclude many.

The weakest recursive explanation is one of the form, "X because of the cause of X." This explanation excludes no causes. At most it excludes the possibility that X was uncaused. Nearly as weak is the explanation deployed by Moliere's Doctor in *Le Malade Imaginaire* to explain the sleep-inducing effects of opium. "Why," the Doctor is asked, "Does opium put people to sleep?" "Because," he replies, "it possesses a *virtuus dormativa* (a sleep-inducing power)." Often represented as the archetypal circular explanation, the Doctor's response is not circular but a weak recursive explanation. Even though the *virtuus dormativa* is very nearly a restatement of the question, the *virtuus dormativa* is not the whole explanation. The whole explanation includes the words "it possesses" and this frame gives us some little hint about where to look or where not to look for opium's sleep-inducing effects. For instance, the effects do not have their source in the social environment associated with the drug. They are not hypnotic effects. Nor do they come from the act of smoking, per se. They are not, for instance, the result of carbon monoxide the smoker inhales on account of the incomplete combustion in the opium pipe. But much as it excludes, the explanation is a fairly weak one because we are told nothing of the sleep-making mechanism, only that it is located in the drug.

Intention explanations are also weak recursive explanations. When we answer the question "Why did he do X?" with "Because he intended to do X!", the intention frames the paraphrase of the question. While not therefore strictly speaking circular, intention explanations suffer from difficulties both with their frames and with their recursions. The framing relation, intention, is notoriously difficult to specify (Rosenberg, 1980a).

Thus far we have been providing examples of recursive explanations that are relatively weak. Not all recursive explanations are as weak as intention explanations or *virtuus dormativa* explanations. Filter explanations can be quite informative, even though recursive. Consider the explanation, "The oil is clean because it has passed through a clean oil filter." Such a filter explanation answers a question of the form, "why are all the X's Y?" with the reply, "Because they have passed through a Y-filter!" A filter explanation accounts for the characteristics of an aggregate as a result of a process in which every individual in an earlier aggregate was examined for its possession or nonpossession of some criterional attribute; those possessing the attribute were permitted to pass, those not possessing it were not, so that every individual in the post-filter aggregate meets the criterion of the filter. Filter explanations are informative because, no matter how ignorant we may be about the criterion, the explanation tells us that:

(1) Somewhere, sometime, there was an initial aggregate of X's which were both Y and not Y.
(2) There is a process which distinguishes X's that are Y from X's which are not Y and selects the former.
(3) The present aggregate is the result of the application of the process in (2) to the initial aggregate in (1).

This information excludes at least three possibilities: that the present aggregate is a chance occurrence, or that there never were anything but Y individuals or that all the not-Y individuals mutated into Y's.

Because filter explanations are relatively informative, they have a respectability as explanations that circular explanations and weaker recursive explanations don't have.

EVAULATING DARWINIAN EXPLANATIONS

Darwinian explanations are filter explanations. They answer questions like, "Why are polar bears white?" The sort of answer they give is, "Because, in the history of the species, those bears that matched their surroundings had more offspring than those that didn't with the result that white fur came to characterize the species!" The explanation attributes to the environment a "white bear filter" which limited access to the successive generations to paler bears until all bears were white. No matter how loose is our grasp of the details of the filter mechanism, we know two things about any Darwinian filter explanations. We know that it is not circular because it is recursive. And we know that it has a relatively high value among recursive explanations because it is a filter explanation.

Everything we have said so far concerning Darwinian filter explanations applies to all members of the class. We are left, therefore, with the question, "How do we tell the difference between good Darwinian explanations and not so good ones? By what standard do we evaluate Darwinian filters?" A standard commonly used to measure theories is their generality. This standard is particularly important to scientists, who are always trying to extend the range of application of their concepts. For such purposes, a theory is valuable just to the extent that it can be generalized to many situations. A theory which explains the snow on our computer screen in terms of a glitch in the monitor is useful for the operator of this computer, but it is too specific to give guidance about how this sort of fault comes about.

A better theory explains the snow by reference to a fault in the machine that etched the circuit board in the monitor. Not only does it explain the same fault in other monitors coming from the same factory, it may also explain different faults in the same computer or even faults in other electronic equipment employing the same etching technique. It is this sort of theory that will be most attractive to engineers interested in computer design. For scientific purposes a theory is most valuable when it offers a principle that will explain large classes of particulars. It is, in this sense, a more general theory.

But generality is not the simple good that exclusivity is. Theories can be too general. The theory that I have snow on the screen of my computer because computers are prone to snow-formation under some circumstances is a very general theory but it is so vague that it provides little guidance either to me or to a computer engineer. Thus, for scientific purposes, the best theories seem to display a moderate degree of generality: general enough to cover many instances, specific enough to avoid vagueness (Garfinkel, 1981, pp. 30-1; Hull, 1974, p. 68). Just where the "golden mean" for an explanation will lie depends on many factors, including the interests of the investigator.

On first consideration, it might appear that the generality and exclusivity tests conflict with one another, but in fact the two criteria work together. Exclusivity refers to the number of causes an explanation excludes; generality to the number of cases in which the cause is excluded. Over-general explanations are bad because they exclude few causes; over-specific explanations are bad because they exclude causes in very few cases.

Like other explanations, Darwinian filter explanations are better if they have a moderate degree of generality. If the characterization of the filter is too specific, the resulting explanation will be ad hoc and too difficult to integrate with accounts of other filter mechanisms. If, on the other hand, the characterization of the filter is too broad, the only way to specify a unitary mechanism behind it is to make that mechanism so vague as to be uninformative.

With respect to generality, there is one critical difference between ordinary explanations and recursive explanations. In nonrecursive explanations, the generality of the explanation is independent of the generality of the phenomenon which is being explained. The highly specific question concerning why a particular ball fell at a particular time may be answered by appeal to the very general theory, Newton's theory of gravitation. But with Darwinian filter explanations, because of their recursive structure, the generality of the explanatory principle is determined by the generality of the terms by which the thing to be explained is described. In short, the generality of the explanans is determined by the generality of the explanandum, because the latter is incorporated in the former. For example, if the question is asked at the level of the polar bear's white fur, the recursive answer—that there is a white fur filter—it must be made at the same level. Thus, the requirement that explaining principles be moderately general places a special burden on recursive explanations. For recursive explanations, this requirement entails that the thing to be explained must also be characterized at a moderate level of generality because the characterization of the thing to be explained appears in the explanation. Thus, what is in general an obligation of scientists to give the right sort of answer becomes, for Darwinian scientists, the requirement that we ask the right sort of question.

This analysis suggests that if research is to progress in the explanation of behavior through natural selection, then close attention must be paid to the descriptions we give phenomena and the questions we ask of them. On the one hand we must avoid questions that are excessively general, such as, "Why have contemporary organisms survived?" The recursive answer, that they have passed through a survival filter, while not circular, is uninformative because it is compatible with any and all mechanisms of survival and provides no guidance to which different sorts of mechanisms underly the perseveration of different species. At the other extreme are questions which request explanations concerning the behavior and morphology of single species. The simple question concerning polar bear fur which we have been using throughout this essay, while useful for illustrative purposes, is actually an example of an overspecified question. By asking the question "Why are polar bears white?" we are led by the recursive structure of our explanation to posit a white bear filter. This explanation is correct as far as it goes and can be backed by a specification of the mechanism that constitutes a white bear filter, but it requires us to develop an entirely new explanation whenever we encounter another kind of mammal in another kind of environment.

IMPLICATIONS FOR EVOLUTIONISTS

We began this essay with a brief recounting of the ambivalence that philosophers and others have expressed concerning the Darwinian theory of evolution. Is it a good theory abused by its practitioners or an overused trivial truism? Is it a theory at all, or merely a meta-theoretical commitment? We are now in a position not only to answer these questions but to make clear the sources of confusion from which they arise and to make suggestions for how problems may be avoided in the future.

Darwinian theory is not circular. It is an example of a recursive theory, a kind of theory in which the thing to be explained is not identical to the thing which explains, as in a circular explanation, but is included within the thing which explains. Darwinian recursive theories are filter theories and they gain some considerable explanatory value from the notion of a filter, whether or not the criterion of the filter is specified. As a consequence, no Darwinian theory is a truism nor is it just a meta-theoretical commitment. Darwinian theories are real theories.

Given the similarity of recursive and circular explanations, however, it's easy to see how the two might get confused. The traditional method for identifying a circular explanation is to discover the explanandum lurking in the explanans. As we have seen, this criterion, while necessary to circularity, is not sufficient for it. To be circular, an explanation must *only* repeat the explanandum in the explanans, and this error is avoided by recursive explanations. Of course, recursive explanations vary in

quality. Some recursive explanations, notably Darwinian filter explanations, do provide considerable information. But most recursive explanations, such as the *virtuus dormitiva* of Moliere, provide relatively little additional information and are therefore difficult to distinguish from genuinely circular explanations. Thus, if philosophers have failed to notice that not all explanations are circular which include the thing to be explained within them, then practicing scientists may be forgiven for making the same oversight.

Having said that the informativeness of Darwinian filter theory protects us from making genuinely circular explanations, we must quickly add that it does not protect us from making fools of ourselves. Darwinians can be as pretentious as Moliere's Doctor if they fail to make a distinction between appealing to Darwinian filter explanations as a class and appealing to a well-specified Darwinian filter explanation. The distinction is not all that easy to make and may be determined by such subtle factors as the assumptions of the audience to which the Darwinian expert is speaking. Before a convention of creationists, the assertion that Darwinian filters explain the biological world is extremely informative, even though it includes no specification of how the world has turned out and therefore only the vaguest sort of specification of the filter. Similarly, an overly narrow filter explanation, such as claiming that the polar bear is white because it has been selected for its whiteness, might be appropriately delivered over the loudspeaker system of an arctic tour boat, even though it gives little guidance for the characteristics of organisms other than polar bears.

But the same statements would be worthless made before a convention of biologists, comparative psychologists, or ethologists. Such explanations should be treated with contempt by such an expert audience because they add nothing to the assumptions that Darwinian scientists already bring to the occasion. The frequency of such misappropriations of Darwinian logic may arise in part from the enormous traffic that has recently occurred in evolutionary behavioral biology between writing for the lay reader and writing for professionals. Having been rewarded both financially and intellectually for offering unelaborated Darwinian explanations to lay audiences, Darwinians are tempted to try them on their colleagues. This temptation is enhanced by the fact that so fragmented have the biological sciences become that each of us seems a lay person to each of the others. This trend has gone so far that the titles of professional papers seem to market their authors and ideas as baldly as advertisers do their products. The temptation to offer over- or underspecific Darwinian explanations to professional audiences should of course be resisted. In that context, such explanations are foolish at best and at worst a form of scientific demagoguery.

In order to avoid this sort of error, scientists need to be aware of the distinction between two components in the value of an explanation, the intrinsic and the heuristic components. Intrinsic value is the power of the explanation to answer inquiries; heuristic value is its power to prompt additional inquiries. Even the barest of Darwinian filter explanation has intrinsic explanatory value, if only because it commits the explainer to Darwinian selection as an explanation, rather than, say, to some form of creation theory. But as we have already observed, such intrinsic value is of no use to Darwinians. For Darwinian audiences, Darwinian filter explanations must provide an additional kind of value, a heuristic value. Thus, for questioners seeking a more specific causal account of the phenomenon, recursive explanations serve a directive function. They serve as a transitional step in our understanding of a phenomenon that permits us to talk about an explanation before we have it fully fleshed out with a causal story which connects the circumstances of the organism with the output of the filter.

The power of this heuristic component can be illustrated by an example familiar to all comparative psychologists; the classical research of Tinbergen et al. (1962) on the gulls' habit of removing eggshells from their nests. At a very early stage in his understanding of that behavior, Tinbergen relied heavily on his faith that the behavior existed because it was selected for. The force of this preliminary explanation is that there exists in the environment some sort of "eggshell removal" filter and the effect of making that explanation was to send Tinbergen out into the gulls' environment looking for the factors which might constitute the filter. Through his comparative analysis and experiments, Tinbergen rejected the idea that risk of damage or contamination to the chicks from the eggshells constituted the filter and settled instead on the idea that the filter was the visual discrimination of the nest sites by carrion crows and other visual predators. The bright, uncamouflaged inner surface of the egg helped carrion crows to locate the nests and resulted in nests with broken eggshells loosing more chicks to predators than nests without. Thus, when Tinbergen's fundamental belief in the efficacy of Darwinian filters was brought to bear on his extensive knowledge of gull-nesting biology, the result was a specification of the mechanism that constituted the filter for eggshell removal.

The Tinbergen example illustrates the special role that a Darwinian filter explanation plays in guiding the research of a dedicated evolutionist. A filter explanation serves to focus the investigator's search for the causal mechanism that constitutes the filter. This heuristic value of a filter explanation depends on the degree to which making the explanation leads to a further specification of the way the processes of nature come to act as a filter. And the degree of specification of the filter depends, as we have seen, upon the degree of specification of the thing we are trying to explain. Perhaps Tinbergen's greatest talent as a scientist lay in his ability to ask questions at an appropriate level of generality. Had he asked only narrow questions, such as "Why does this species of gull remove its eggshells?" or broader questions, such as "Why are gulls so successful?" he might never have been led to the crucial comparison with cliff nesting gulls which do not suffer from such severe egg predation and have not, as a consequence, been selected for eggshell removal.

For a Darwinian scientist speaking with other Darwinian scientists the primary value of filter explanations is a heuristic value. Since heuristic value is maximized by descriptions at a moderate level of generality, serious Darwinian scholars of animal behavior must focus their attention on what sorts of questions are likely to be the most productive. And since filter explanations are recursive and dependent for their quality on the quality of the question to which they are the answer, Darwinians cannot afford to ask questions at any level of generality and only worry about the generality of their answers. They must ask the right questions.

This principle may, upon sober consideration, cast doubt on some of the hottest research issues of the day. To ask questions like, "Why is behavior optimal?" or "Why is it altruistic?" may specify filters too broad to provide much guidance. It's hard to imagine scientists searching profitably for an optimality filter or an altruism filter. At the other extreme, asking questions like "Which variable is this species optimizing with its feeding behavior?" may lead to fragmented principleless accumulation of data. Creative use of Darwinian filter explanations requires asking questions that lead to the kind of comparative description that lay at the core of Tinbergen's most creative work and which, unfortunately, has largely gone out of style in recent years.

The ideal question consists of a search for an explanation of general principles that characterize the traits of organisms in relation to their circumstances. As a prelude to that search, we must of course define those principles. The candidate descriptive principle that we recommend most highly is the principle of adaptation, that which was closest, we feel, to Darwin's original intent. The best question for comparative psychologists to ask is why organisms are adapted to their environments. By "adapted" we do not mean of course, reproductively successful, because that would turn the explanation into the vague question cited above. We do mean by "adapted," more or less well designed for the circumstances of their existence. Consequently, asking the question entails the kind of exhaustive descriptive analysis of the concept of adaptation so often recommended (Sommerhoff, 1950; Thompson, 1986b, 1987; Williams, 1966; Curio, 1973) but rarely embarked on. Because of the recursive structure of Darwinian explanations, this descriptive analysis is crucial to the future of all Darwinian sciences—including comparative psychology—and it cannot be further delayed.

In considering the proposal that they take up the descriptive study of adaptation, psychologists should remember that behavioral adaptation occurs at two levels of analysis central to their field: the level of the species and the level of the individual. We can ask, for instance, why the behavior of species members is on the whole adapted to their ecological circumstances. This is the appropriate question for natural selection theory to answer. We can also ask, however, how individuals' behaviors come to be adapted to the flux of conditions in their moment-by-moment and day-to-day lives. This, of course, is a question of the motivational design of behavior, and is, we think, the appropriate question to be approached by a modern and ecologically sophisticated version of the other filter so central to comparative psychology: learning theory. Together, we think these two questions, and the relationship between them, deserve to be the proper domain of a newly constituted field of animal behavior, ethology, behavioral biology, or what-have-you (Thompson, 1986b, 1987). This new animal behavior field will be more theoretical than classical ethology but more descriptive than classical experimental animal psychology and contemporary sociobiology. And because it is attentive to the logical structure of its two central explanations, it will avoid the metatheoretical wrangling that has characterized the study of behavioral evolution in both its recent and more distant past.

ACKNOWLEDGEMENTS

This paper arose from discussions held within the framework of the Clark University Philosophy Colloquium. We are also grateful for lengthy discussions with Jon Bateman, D. Lipton, R. Mitchell, C. Thompson, and P. C. Thompson which helped to develop the ideas contained here.

REFERENCES

Bateson, G. (1979). *Mind and nature: A necessary unity* New York: Dutton.

Curio, E. (1973). Towards a methodology of teleonomy. *Experientia, 29*, 1045-1058.

Friedman, M. (1974). Explanation and scientific understanding. *The Journal of Philosophy, 71*, 5-19.

Dunbar, R. I. M. (1982). Adaptation, fitness and the evolutionary tautology. Kings College Sociobiology Group (Ed.), *Current problems in sociobiology*. Cambridge: Cambridge University Press.

Garfinkel, A. (1981). *Form of explanation* New Haven: Yale University Press.

Hailman, J. (1982). Evolution and behavior: An iconoclastic view. In H. C. Plotkin, *Learning, development and culture* (pp. 205-254). New York: Wiley.

Hull, D. (1974). *Philosophy of biological science*. Englewood Cliffs, N.J.: Prentice-Hall.

Lewontin, R.C. (1978). Adaptation. *Scientific American, 239*, 157-169.

Nozick, R. (1974). *Anarchy, state, and utopia* New York: Basic Books.

Popper, K. (1972). *Objective knowledge* Oxford: Oxford University Press.

Popper, K. (1978). Natural selection and the emergence of mind *Dialectica, 32*, 339-355.

Rindos, D. (1984). *The origin of agriculture, an evolutionary perspective* New York: Academic Press.

Rosenberg, A. (1980a). *Sociobiology and the preemption of social science*. Baltimore: Johns Hopkins.

Rosenberg, A. (1980b). Fitness. *Journal of Philosophy, 80*, 457-473.

Sommerhoff, G. (1950). *Analytical biology*. Oxford: Oxford University Press.

Tinbergen, N., Broekhuysen, G. J., Feekes, F., Houghton, J. C. W., Kruuk, H., & Szulc, E. (1982). Eggshell removal by the black-headed gull, *Larus ridibundus L.*: A behaviour component of camouflage. *Behaviour, 19*, 74-118.

Thompson, N. S. (1976). My descent from the monkey. In P. P. G. Bateson & P. H. Klopfer (Eds.), *Perspectives in ethology* (pp. 221-230). New York: Plenum.

Thompson, N.S. (1981). Toward a falsifiable theory of evolution. In P.P.G. Bateson & P.H. Klopfer (Eds.), Perspectives in ethology (pp. 51-73) New York: Plenum.

Thompson, N. S. (1986a). Deception and the concept of natural design. In R. W. Mitchell & N. S. Thompson (Eds.), *Deception* (pp. 53-66). Albany, N.Y.: SUNY Press.

Thompson, N.S. (1986b). Ethology and the birth of comparative teleonomy. In R. Campan & R. Zayan, (Eds.), *Relevance of models and theories in ethology* (pp. 13-23). Toulouse, Privat, I.E.C.

Thompson, N. S. (1987). The misappropriation of teleonomy. In P. P. G. Bateson & P. H. Klopfer (Eds.). *Perspectives in ethology*. New York: Plenum.

Thompson, N. S. (1987). Natural design and the future of comparative psychology. *Journal of Comparative Psychology, 101*, 282-286.

Williams, G. C. (1966). *Natural selection and adaptation*. Princeton: Princeton University Press.

The International Journal of Comparative Psychology, Vol. 1, No. 4, Summer 1988

RESPONSE

WHY DR. TINBERGEN IS MORE SOUND THAN DR. PANGLOSS

Peter Lipton N.S. Thompson

We know that the editors have chosen their peer-commentators well because we find our position assaulted with equal skill and vigor from opposite directions. On the one hand, Saunders and Ho (1988) take us to task for having so much faith in the neo-Darwinian synthesis as to fall prey to Panglossian adaptionism. On the other hand, Schull (1988) takes us to task for having so little faith in the neo-Darwinian synthesis as to place it on a par with its poor cousin, reinforcement theory.

We suspect that the concept of natural design lies at the core of the difference between ourselves and our critics. Central to our argument is the idea that the behavior and morphology of organisms is characterized by two design properties, adaptation and purpose. Following Ryle (1949), Somerhoff (1950), Tolman (1951), and Hofstadter (1941), we think of these design properties as higher order characteristics of organism, similar to dispositional properties such as brittleness. These properties are not explanatory in any sense. To say that an animal's morphology is adapted or that its behavior is purposive is no more to explain it than saying that a lemon is yellow explains the lemon.

Both critics seem to doubt our assumption that adaptation or purposiveness can be recognized independently of the concepts that are deployed to explain them, natural selection and reinforcement. Saunders and Ho doubt that adaptation can be recognized except in reproductive success. Although he doesn't explicitly say so, Schull seems to doubt that we could ever make good on our commitment to identify the things toward which animals strive independent of their capacity to reinforce instrumental behaviors.

Identifying these properties in organisms in nature independent of their explainers is not easy but it is surely not impossible. The key is the comparative method. Just as it is possible to identify a property of adaptation by comparative analysis of species in relation to their environments, so it is possible to identify a property of purposiveness in the

behavior of individual animals by comparing the individual animal's behavior in relation to its circumstances. And just as you don't need to know how many eggs a woodpecker has to know that its beak is adapted to hammering wood, you don't have to know how many times a rat has been reinforced to know that its present behavior is directed toward discovering food. Like adaptation, purposiveness is a property observable in the behavior of animals.

Having begun with these general comments, let us proceed to consider some particular points.

(1) We are not—repeat *NOT*—apologists for Panglossian adaptionism. In its strongest form, we take the Panglossian position to be that Everything is Adapted and that All Adaptions are Selected. We disagree with both statements. We do, however, believe that some things are adapted and that some of these are selected. Adaptation is a property frequently observed in organisms and natural selection is often a reasonable explanation of that property. One of us (Thompson, 1981, 1987) has urged scientists to pay attention to two crucial problems of our day. When do organisms display the properties of adaptedness and when not? And when is natural selection an appropriate explanation for adaptedness and when not? These questions envisage all sorts of phenomena not envisaged by Panglossian adaptionism. They envisage adaptations not caused by natural selection and they envisage natural selection causing properties of organism other than adaptedness. So, unless Saunders and Ho take the position that adaptedness is a property rarely observed in organisms or that when observed, it is rarely attributable to natural selection, then they cannot pick a fight with us over Panglossian adaptionism.

Schull, on the other hand, is correct to pick a fight with us. He clearly believes that natural selection is the only reasonable explanation for adaptation and that adaptation is what natural selection must necessarily produce. This belief causes him to drive a wedge between what he says is the strong explanation of Darwinism and the weak explanation of reinforcement theory. But because we believe that natural selection is only one of a handful of factors that explain adaptations and because we believe that natural selection produces aberrations as well as adaptations, we don't see the difference that he sees between reinforcement theory and Darwinian natural selection theory.

In particular we don't see the difference in predictive power of the two theories. If we saw a distinct pattern of directedness in the behavior of an animal, and we saw that the performance of some behavior would transport the animal to the goal conditions designated by that pattern of directedness, and finally we saw that the animal was capable of performing and remembering that behavior, we would be surprised if that animal did not learn the behavior. Schull is quite correct to observe that animals don't learn all useful behaviors. But as Saunders and Ho would be quick to point out, species don't produce all useful adaptations. And the reasons seem to be the same: the species cannot evolve what it cannot first produce by recombination or mutation and then subsequently retain through inheritance. Similarly, the individual cannot learn what it cannot first perform through innovation and then subsequently retain through memory. But given that the starting conditions of variation and retention are met, *both* natural selection and learning seem to guarantee systematic change as a matter of logical necessity. Both theories predict that the changes would adapt the organism to its circumstances [or bring it closer to its goals] and these predictions have empirical content and can be falsified in both cases.

But even if learning theory is less well-founded than natural selection, that would not show—as Schull seems to suppose—that the two theories don't have the same explanatory structure. Schull's supposition is a non sequitur. The explanatory structure of a theory is not determined by its evidentiary status. Whatever their foundation, both theories are recursive filter explanations.

(2) What was for us the main point of the target article seems to have been missed by both commentators. We set out to explain for ourselves why it was that the problem of circularity in evolutionary explanation refused to go away. The literature is littered with statements of this problem and with purported solutions of it. Yet it keeps coming back. If circularity is such an obvious flaw in the theory, why do brilliant people seem to still be prey to it? And if it's not a flaw, why do other brilliant people seem to bemoan it?

We think that the target article solves the problem of the recurring circularity problem. The solution we offer is that both schools of thought are correct. The people who say that the theory is not circular (or tautological) are correct because strictly speaking it is not. On the other hand, those who say that the theory lays terrible traps for the unwary are also correct. In other words, just because evolutionary explanations are not circular doesn't mean that they can't be empty and silly. We hope that the target article will stifle forever the useless discussion about whether they are circular or not and begin a more useful discussion about when they are useful and sound and when they are empty and silly. It was on this controversial point that we hoped our commentators would focus.

When is an evolutionary explanation empty and silly? It is empty and silly whenever those that use them abuse the special structure of such explanations. Darwinian explanations are recursive and as such, they have two parts: a frame which specifies what kind of a recursive explanation we are employing and the contents of the frame, which specifies which particular form of that kind of recursive explanation we are employing. For instance, when Tinbergen said that he thought that gulls removed eggshells from their nests because the surface nesting environment selected such behavior he was making two points: 1. The gulls' behavior and morphology is the result of a filter; 2. The particular filter is an "eggshell removal" filter.

Now, the first statement is clearly not circular. No mention is made of the explanandum in the statement. Nor is it empty. To say that something has been filtered is, as our article points out in detail, to say quite a bit about it. That is why we feel we have laid to rest, hopefully forever, the question of the putative circularity of evolutionary explanations.

The second statement does contain the explanandum, but it is not the whole explanation and consequently the whole explanation is not circular. To underline this difference, we called the whole explanation recursive because, while it is not a restatement of the explanandum, it does refer back to it, and that 'back-reference' can be the source of difficulties.

Our goal was to begin a discussion about when such back references were troublesome and when they were benign. We thought Tinbergen's use of such statements was a good exemplar because he managed to do good science while using them and we wondered why he had been so fortunate. We reasoned that because the answer (an eggshell removal filter) to Tinbergen's question, "Why do gulls remove eggshells?" contained the question, there must have been something particularly fortunate about Tinbergen's understanding of his own question that kept him out of trouble. To put the matter briefly, if all evolutionary answers contain their questions and some evolutionary answers are better than others, then perhaps the quality of the explanation has something to do with the quality of the question embedded in it. We suspected that Tinbergen has been particularly happy in his choice of *questions*.

What was the outstanding characteristic of Tinbergen's questions? Tinbergen was first and foremost a comparative ethologist. Instead of asking the question, "Why is this animal this way in this place?" as a freestanding question, he always asked it in the context of a whole series of additional questions: "Why are those other animals those other ways in those other places?" In other words, Tinbergen's questions were more inclusive than they appeared to be. Although he appeared to be asking a question about the nesting habits of a particular gull under a particular set of circumstances, he actually was asking a somewhat broader question. He was asking, how does it come about that the eggshell disposition habits of gulls are *adapted?* In asking a question at this intermediate level of generality he was avoiding two faults of much contemporary adaptionist explanation. He was avoiding the fault of ad hoc examination of the traits of individual species. And he was also avoiding the fault of characterizing his problem in such broad terms. (e.g., optimality) as to prohibit any useful examination of it.

In this respect, Tinbergen was profoundly different from Pangloss. Pangloss' fault was that he had no concept of "the good." Consequently, what ever happened became the good, simply by virtue of its having happened. If his theory was that God arranges that the good always happens, then he should be able to specify the good in advance, and that Pangloss could not do. Tinbergen had a leg up on Pangloss. Because of his grasp of the comparative method and his vast knowledge of natural history, he had independent knowledge of the good. Unlike Pangloss, Tinbergen approached his phenomena with systematic and falsifiable expectations.

If there is any lesson from all of this, it is that evolutionary explanation requires comparative description. Why? Because evolutionary explanations are recursive and therefore contain within them a description of the thing to be explained. As such, the quality of the explanation relies on the quality of the description. Vague descriptions lead to vague explanations, ad hoc descriptions lead to ad hoc explanations. Thus healthy evolutionary explanation relies on a base of well-articulated and general comparative description.

Despite Schull's doubts, we think that the same general rules of thumb can be applied to recursive reinforcement theory. Just as rigorous evolutionary explanation begins with careful comparative description, so rigorous explanation through reinforcement theories requires careful description of purposiveness in behavior. A well-described explanandum is the key to a useful recursive explanans.

(3) Biological and psychological points aside, our critics seem to have made several mistakes that are more purely philosophical. Two of these seem particularly interesting. The first is that there is nothing special about recursive explanations; the second is that a theory cannot explain its own evidence.

According to Saunders and Ho, recursive explanations are not circular, but they are not different from normal (i.e., nonrecursive) causal explanations. Their argument is that, in answer to the question "Why is E," the normal answer "Because of X" can trivially be converted to "Because of X which is the cause of E." The only peculiarity of recursive explanations is that they contain a redundant clause. We are pleased that Saunders and Ho clearly recognize that recursive explanations are not circular, one of the main points of our article, but they should have given more credit to the distinction between a recursive explanation and a normal one. In a recursive explanation, the cause is only specified by reference to its effect. This is not normally the case. If we explain the Chicago Fire by blaming Mrs. O'Leary's cow, we have a normal explanation, since the cause is specified without any reference to its effect. You don't need to know about the fire to know the cow. By contrast, if we explain the presence of a given behavior by saying that there is a filter that selected for that very behavior, we only get a grip on a cause by reference to its effect. The relevant filter is only specified by reference to the resulting behavior.

Recursive explanations cannot be automatically converted into normal ones. The reason is simple: we sometimes do not have an independent characterization of the cause to replace the recursive answer. Saunders and Ho seem to deny this. They argue that it is trivial to convert the recursive "Polar bears are white because there is a force in the environment that makes them white" into the nonrecursive "Polar bears are white because of a force." This is a confusion. What is the answer they are proposing to the question of why polar bears are white? If it is simply that there exists an unspecified force in the environment, this is no explanation at all, because it neither specifies the cause nor tells us anything about the kind of cause it is. There are many forces in the environment, most of which have nothing to do with bears. What is more likely is that they intend "a force" to mean "a force that makes polar bears white." This is a cause,but it is obviously recursively specified. We have no objection to the shorthand "They are white because of a force," except that it may camouflage the recursive structure of the explanation and so mislead us into thinking that there is nothing special about recursive explanations. There really are three types of explanations: normal, circular and recursive. To determine which type you have, first look at the answer, then compare it to the question. If the answer specifies a cause without any reference, explicit or implicit, to the question, then the explanation is normal. If it fails to specify a cause at all, but only restates the question, it is circular. Finally, if the answer does specify a cause, but only in terms of the question, it is recursive.

The second important mistake concerns the relations between explanation and evidence. In analyzing the quality of explanations, it often helps to distinguish two questions: "Is there good reason to think that the explanation is true?" and "Would it be a good explanation if it were true?" The answers to these questions may diverge. In the case of Moliere's explanation that opium puts people to sleep because of its dormative powers, the answer to the first question is yes and to the second, no: the explanation is almost certainly true, but it is a terrible explanation. There are also cases where we have very little reason to believe the explanation, but where it would be a very good explanation if it were true.

To avoid unnecessary offense, examples of these are left as an exercise for the reader. Saunders and Ho, and Schull seem not to make this useful distinction between the evidence for an explanation and its explanatory power. This may have led them to suggest that an explanation cannot account for its own evidence. Thus Saunders and Ho argue that looking for a biological sub-problem "may then have heuristic value, but nothing more than that, since it is only by knowing the solution that we can be at all sure what the sub-problem really was." Similarly, Schull seems to claim that if the only evidence for a reinforcer is the behavior to be explained, then the reinforcer cannot be used to explain the behavior. The general view they seem to hold is that an explanation is no good if the phenomenon to be explained itself provides an essential part of the evidence that the explanation is true.

This view, that "self-evidencing" explanations are bad explanations, may appear plausible, since these explanations seem circular. The evidence is supposed to support the theory while the theory explains the evidence. Nevertheless, there are perfectly good self-evidencing explanations (Hempel, 1965, 372-73). I may explain the distinctive tracks in the snow before me by saying that a person on snowshoes has passed this way, even if my only reason for believing this are the very tracks I am explaining. An astronomer may explain the red-shift in a star by telling us the star's speed of recession, even if she depends on the red shift itself to determine the speed. And self-evidencing explanations are virtually unavoidable when the explanation appeals to unobservable entities, such as subatomic particles and their ilk. Here it seems that the only evidence we have for the theory is the evidence it explains. Indeed there is a popular theory of inductive inference according to which almost all explanations turn out to be self-evidencing. According to this account, known as *Inference to the Best Explanation*, we infer from our evidence, to what would be the best explanation for that evidence (Harman, 1965; Lipton, in press). Of course a particular self-evidencing explanation may be very poorly supported by the available evidence, but you cannot show this simply by pointing out that it depends on the evidence of the phenomenon it is supposed to explain. As the examples above show, the fact that an explanation is self-evidencing entails neither that it is insufficiently supported by the available evidence, nor that it lacks explanatory power.

REFERENCES

Harman, G. (1965). The Inference to the best explanation. *The Philosophical Review, 74*, 88-95.

Hempel, C. (1965). *Aspects of scientific explanation.* Englewood Cliffs: Prentice-Hall.

Hofstadter, A. (1941). Objective teleology. *The Journal of Philosophy, 38*, 29-56.

Lipton, P. (in press). *Inference to the best explanation.* London: Routledge and Keagan.

Ryle, G. (1949). *The concept of mind.* New York: Hutchinson's University Library.

Saunders, P.T., & Ho, M.W. (1988). Why bring Dr. Pangloss into comparative psychology? *International Journal of Comparative Psychology, 1*, 233-237.

Schull, J. (1988). The difference between a circle and a spiral may be smaller in learning than in evolution. *International Journal of Comparative Psychology, 1*, 230-232.

Sommerhoff, G. (1950). *Analytical biology.* Oxford: Oxford University Press.

Thompson, N.S. (1981). Toward a falsifiable theory of evolution. In P.P.G. Bateson & P.H. Klopfer (Eds.), *Perspectives in ethology* (pp. 51-73). New York: Plenum.

Thompson, N.S. (1987). The misappropriation of teleonomy. In P.P.G. Bateson & P.H. Klopfer (Eds.), *Perspectives in ethology* (pp. 259-274). New York: Plenum.

Tolman, E.C. (1951). *Collected papers in psychology.* Berkeley, CA: University of California Press.