University of Alberta

Patterns in Nature: Metaphor in Biological Explanation

By

Thomas Kenneth Bond ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Arts

Department of Philosophy

Edmonton, Alberta

Fall 2004

Your file Votre référence
ISBN: 0-612-95622-9
Our file Notre référence
ISBN: 0-612-95622-9

Table of Contents

List of Tables

List of Figures

Chapter One

Metaphor and Science: The Case of Biology

...science, like the humanities, like literature, is an affair of the imagination.

(Max Black 1962)

1.1 Introduction

Explanations in evolutionary biology provide putative answers to questions about why the living world is as it is. Genuine biological explanations, it might be said, make the properties of and relations between living things intelligible, giving us a picture of the way the world is. It might also be said that this intelligibility is in large part a function of the degree to which explanations can support successful predictions and describe the causal interactions between things in the living and non-living world. If asked to explain why it is that we can give such explanations, we might say that our ability is the result of our explanatory terms and relations connecting with the causal structure of the world. It might be surprising, then, to find that successful explanations in evolutionary biology frequently utilize metaphor. In so far as metaphor is a product of our creation, its presence implies that our understanding of the world through science is, as Black says above, an affair of the imagination, a view at odds with our realist intuitions regarding genuine explanations. How can metaphors be part of genuine biological explanations?

1.2 Darwin's Metaphors

The history of evolutionary biology is suggestive of the link between imagination, metaphor, and understanding. Many scholars have noted that Charles Darwin, the father of evolutionary theory, successfully captured the idea of natural selection by an analogy with artificial selection and economic theory. As a well-respected biography of Darwin puts it,

> Darwin believed that similar imperceptible variations held the key to Nature's own Malthusian selection. Weak, ill-adapted variants were discarded by Nature, as they were by the fancier. The good ones thrived and over the generations particular trends were encouraged. Adaptive features were drawn out, as if by an invisible breeder. 'Artificial selection' showed the craftsman sculpting nature; Nature's own 'selecting' hand was infinitely superior (Desmond and Moore 1991, 428).

Other scholars have pointed to the importance of Darwin's analogy to the development and reception of his theory by his contemporaries (Young 1985) and the importance it had for Darwin making his case for the power of natural selection (Sterrett 2002).

Despite great care in its use, Darwin's phrase "natural selection" was amenable to all sorts of readings and modifications, allowing Darwin's adversaries and critics to turn his own language against him, employing it to invoke the power of a creative, active Deity (Young, 99-101). These anthropomorphic overtones, that is, the attribution of human qualities to the non-human world, are a fundamentally metaphysical position (Young, 92). The presence of "anthropomorphic, voluntarist descriptions of natural selection" are remarkable when set against the "official paradigm" of the time which banished talk of purposes, intentions, and anthropomorphic expressions from scientific explanations and insisted on descriptions in terms of matter, motion and number (93). Young gives numerous examples from the *Origin* that display this "rank anthropomorphism" of nature "acting," "selecting," "watching," and being a "powerful agent" (94). And, we find the following in the "Recapitulation and Conclusion" of Darwin's *The Origin of Species*:

> If then we have under nature variability and a powerful agent always ready to act and select, why should we doubt that variations in any way useful to beings, under their excessively complex relations of life, would be preserved, accumulated, and inherited? Why, if man can by patience select variations most useful to himself, should nature fail in selecting variations useful under changing conditions of life, to her living products? ([1859] 1968, 443)

By the third edition of the *Origin*, Darwin was explicit in insisting that his anthropomorphic language was simply a metaphor for brevity's sake, but that "everyone knows what is meant and is implied by such metaphorical expressions" (Young, 96). Darwin's need to talk in this metaphorical way may have been partly a result of his education, but it was also because "he was providing an abstract account at a general level of how favorable variations might be preserved," rather than an account of a mechanism of how species come to be (98). Darwin himself looked toward a time when his biological terminology would cease to be metaphorical:

> The terms used by naturalists of affinity, relationship, community of type, paternity, morphology, adaptive characters, rudimentary and aborted organs, &c., will cease to be metaphorical, and will have a plain signification. When we no longer look at an organic being as a savage looks at a ship, as at something wholly beyond his comprehension; when we regard every production of nature as one which has had a history; when we contemplate every complex structure and instinct as the summing up of many contrivances, each useful to the possessor...([1859] 1968, 456)

Herbert (1971) argues that, previous to reading the work of Malthus, Darwin did not hold a sufficiently determinate notion of artificial selection to anticipate finding a similar process at work in nature (212). It was Darwin's reading of the work of Malthus that "impelled Darwin to apply what he knew about the struggle at the species level to the individual level, seeing that survival at the species level was the record of evolution and survival at the individual level its propulsion." (217) Thus, the metaphorical transfer of Malthus' description of nature's checks on human populations was required not just for brevity or rhetorical effect, but also to focus Darwin's attention on the competitive aspects of nature. It has been almost 150 years since the *Origin of Species* was first published. Has Darwin's hopeful prediction of an evolutionary biology shorn of metaphor finally come to pass?

1.3 Biological Language Since the Modern Synthesis

Michael Ruse (2000) argues that the Modern Synthesis, the reconciliation of Mendelian genetics with Darwinian evolution that took place in the 1930s and 1940s, has provided us with an entirely new set of metaphors such as Sewall Wright's "adaptive landscape", and more recent work has created still others, for example, John Maynard Smith's "evolutionary stable strategies." Ruse argues that in more formal biology, like that of Wright's mathematical formulation of adaptive landscapes, metaphors add valuable cognitive elements to the mathematical representation. For example, the formal expression of Wright's "adaptive landscape" does not suggest the kind of continuous fitness surface that is expressed by the metaphor (605). Giving up metaphor in evolutionary biology would involve not only "significant pruning," but a concomitant loss of predictive fertility. Further, the ability to use the ideas of evolutionary theory to push into new areas of inquiry and to propose new solutions would be dramatically reduced (607). For Ruse, "[M]etaphors have an absolutely vital heuristic or predictive role in evolutionary biology. Without metaphors...one is stuck simply with a fancy description of what there is." (607) Biologists peer into the unknown through their metaphors (608).

In his opening to *The Triple Helix,* Harvard geneticist Richard Lewontin affirms the necessarily metaphorical nature of the activity of scientific explanation, including the biological sciences:

> It is not possible to do the work of science without using a language filled with metaphors. Virtually the entire body of modern science is an attempt to explain phenomena that cannot be experienced directly by human beings, by reference to forces and processes that we cannot perceive directly because they are too small, like molecules, or too vast, like the entire known universe, or the result of forces that our senses cannot detect, like electromagnetism, or the outcome of complex interactions, like the coming into being of an individual organism from its conception as a fertilized egg. Such explanations, if they are to be not merely formal propositions, framed in an invented technical language, but are to appeal to the understanding of the world that we have gained through ordinary experience, must necessarily involve the use of metaphorical language. Physicists speak of "waves" and "particles" even though there is no medium in which those "waves" move and no solidity to those "particles." Biologists speak of genes as "blueprints" and DNA as "information." (2000, 3)

Though metaphors may indeed be necessary for understanding, they carry with them the danger of being mistaken for the very thing they were enlisted to help explain. This worry is evident in Lewontin's discussion of the implications of the metaphorical attributions of 'development' and the way in which these attributions structure the kinds of questions biologists ask and the hypotheses they advance concerning the similarities and differences between individuals.

Lewontin's assertion of the necessity of metaphor and the implication that it may play a strong cognitive role in scientific theorizing and explanation may strike some as exaggerated, a scientific corollary of the Romantic belief that "language is vitally metaphorical" (Richards [1936]1965, 49) or an implication of the more radical claim that our conceptual system itself is fundamentally metaphorical in nature (Johnson and Lakoff 1980, 3). Perhaps a more traditional and more common attitude toward the role of metaphor in science would be that "...in strict scientific statement metaphor has no place, and that in a less strict scientific discussion metaphor is optional or additional." (Brooks and Penn Warren 1950, 374) The language of science, even informally, need not employ metaphor in its investigations or explanations of phenomena because it uses terms that somehow refer directly to objects that exist independently of our conceptions of them.

But, consider the following ways in which major ideas in evolutionary biology are described:

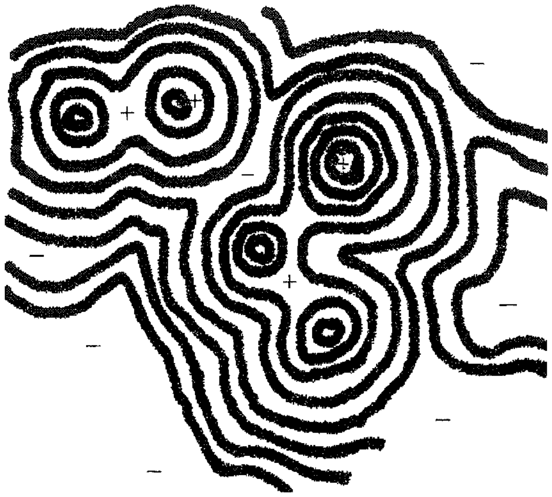
> ...the living world is not a formless mass of randomly combining genes and traits, but a great array of families of related gene combinations, which are clustered on a large but finite number of adaptive peaks. Each living species may be thought of as occupying one of the available peaks in the field of gene combinations. The adaptive valleys are deserted and empty.

> Adjacent adaptive peaks are arranged in groups, which may be likened to mountain ranges in which the separate pinnacles are divided by relatively shallow notches. (Dobzhansky 1951, 9-10)

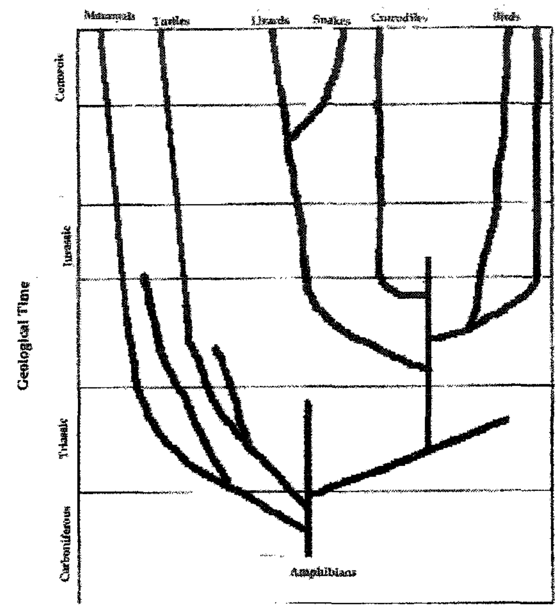
> To infer the history of evolution of a group of organisms is to infer the phylogenetic relationships among the species – the pattern of branching – and to describe, for each branch of the tree, the rate and pattern of anagenetic change in characteristics that interest us. (Futuyma 1986, 286)

These two metaphors describe respectively, a mechanism of evolutionary change, and the history of evolutionary change. The first is a product of the evolutionary synthesis and the second, though formulated long before the synthesis, is used equally well after, and continues to capture the essential shape of evolutionary change in competing schools of thought in systematics.

The metaphor of the adaptive landscape is a dramatic representation of the relationship between an organism's fitness, its genotype and its ecological niche (Dobzhansky 1951, 10). This idea is often represented as a topographic map as below (Figure 1.1). The contour lines symbolize the adaptive value (Darwinian fitness) of the genotypes (Dobzhansky, 9). Because fitness is highest on peaks and selection can only increase fitness, no population could move from one peak to another by natural selection alone; for to do so, selection would have to allow an organism to move downhill - to decrease in average fitness (Futuyma, 172).



The Adaptive Landscape
Fig. 1.1



The Phylogenetic Tree
Fig. 1.2
(Based on Futuyma 1993)

The second metaphor, the "tree of life" or phylogenetic tree, illustrates the relation of species to one another and to ancestral species from which they have descended (Figure 1.2). The tree shows the processes of anagenesis (change within a lineage) and cladogenesis (divergence of one group from another). Among other things, the tree metaphor helps us to formulate expectations about the presence of ancestral or derived characteristics that groups might share (Futuyma, 290). Phylogenetic inference, the determination of the pattern of descent, also uses the concepts of the rooted tree and the unrooted tree. These are phylogenetic trees with the time dimension removed and omitting the "deepest common ancestor" allowing for the construction of several alternative descent patterns (Ridley 1996, 463).

Questions concerning the role of metaphor in scientific explanation and description strike at the heart of the scientific enterprise. As Ruse (1999) explains, though scientists may be guided in their work by epistemic norms such as predictive power, and coherence, consistency and consilience with other areas of accepted scientific theory, the contribution of cultural values, such as social progress, were important in the development of evolutionary theory, though they may have decreased importance now. Ruse observes that there is still something fundamentally cultural about the mature form of evolutionary biology (239). The cultural aspect that he observes is not that of cultural values but of cultural metaphors. Natural selection, trees of life, evolutionary stable strategies, selfish genes, "Through language, the ideas, the pictures, the models, above all the metaphors that evolutionary biology uses, culture comes rushing right back in...an idea from one domain, that of culture, is taken and applied to another domain, that of organisms." (239)

The admission of the presence of metaphors in biological explanation challenges even modest claims of objectivity and realism. If the language scientists use to express their findings is inevitable drawn from their culture, it seems that scientists' descriptions of biological phenomena are more a reflection of their society than of the real world (Ruse, 240). And with the introduction of cultural metaphor comes the spectre of value-laden theorizing, epistemic relativism and antirealism. Thus, the metaphors we use to describe our activities and the entities that we discover, theorize about, and with which we populate our biological/conceptual world, are difficult to ignore and raise deep philosophical issues. This thesis is an attempt to make a contribution to the discussion of the nature and function of metaphor in explanations in evolutionary biology by examining the metaphors in two works of one of the pre-eminent popularizers of evolutionary biology, Richard Dawkins.

1.4 Master of Metaphor

Richard Dawkins' "gene's eye view" of evolution has found wide appeal in both professional and popular biology. The attractiveness of his writing is undoubtedly due in part to his presentation of well-considered arguments that encompass a large field of investigation including molecular biology, paleontology and ethology. Dawkins' first book, *The Selfish Gene* ([1976]1989), is well-known not only for its particular view of the genetic basis of individual and group behaviour - of genes as programmers of individual behaviour - and for its contribution to the debate over the units of selection - that genes are the unit ultimately "selected for" - but also for its vivid use of metaphor. And though his second book, *The Extended Phenotype* ([1982]1999), was aimed more specifically at his professional colleagues, it still contains many of the metaphorical aspects of its more popular precursor. The cries of "Reductionism!" "Misplaced metaphor!" and "Panglossianism!" that these books provoke are reason enough to motivate a more detailed evaluation of the function of Dawkins' metaphors. More reactionary critics aside, if citations are any indication of the influence of an idea or view, then *The Selfish Gene* and *The Extended Phenotype* appear more influential now than when first published, even if only as something to knock down. My own search of the Science Citation Index shows that of a total 837 citations to *The Selfish Gene* in English language articles, 152 citations fell within the first five years after its publication (1976-1981), while 390 citations have been made to *The Selfish Gene* in the last five years (1999-2004). Likewise, for *The Extended Phenotype*, of 195 total citations in English language articles, 70 fell within the first five years after its publication (1982-1987), while 90 have been made within the last five years (1999-2004). Though these numbers are hardly breathtaking, I think that the relative increase in the number of academic references to books that are considered popularizations is suggestive of the power of Dawkins' view to spur reflection on the nature and implications of evolutionary theory (even if these last five years are part of a downward trend). There is little doubt that the metaphorical nature of his writing has had a hand in influencing others.

Yet another reason to re-examine Dawkins' early work is the mounting empirical evidence that calls his gene-centred view into question. His claim for the "textbook orthodoxy" of the gene-centred view was, even in 1989, difficult to support. Like many movements or methodological views, however, with the passage of time it is easier to place Dawkins' ideas in the context within which he was writing and, in doing so, to tease apart the positive and negative effects that they have had.

In a professional piece of scientific writing, prominent metaphor might seem out of place, but the use of metaphor is, for Dawkins, one of the ways in which the popularization of science can actually work to further scientific investigation and theorizing. Metaphors are a way of recapturing the essential ideas of an often difficult and abstruse technical language. They play a didactic role, allowing the uninitiated to appreciate a theory or range of facts previously unknown to them.

The beneficial aspects of metaphor are not limited to this didactic role, however. Dawkins suggests that the metaphorical re-description of technical language and concepts might allow even professional scientists to "see" facts and evidence in new ways. In providing scientists with a novel perspective, these re-descriptions might facilitate the formation of different conceptions of how well-known, but disparate facts, might be coherently put together, thus helping to generate new theories or, more dramatically, to reveal new facts. It is this heuristic and cognitive function of metaphor that makes the prevalence of metaphor in Dawkins' writing philosophically interesting.

Dawkins frequently distinguishes between his metaphorical or non-metaphorical use of a term or concept in an attempt to prevent any misunderstandings. Despite these distinctions, it is often difficult to determine the criteria upon which he is basing his categorization of a term or concept as metaphorical or non-metaphorical. Behind Dawkins' adroit use of metaphor is a particular view of the nature and function of metaphor in language and thought and, more importantly and more specifically, in scientific (i.e., biological) writing, investigation and theory. Dawkins himself seems to work on the assumption that metaphors in scientific writing are always amenable to translation into the non-metaphorical language of the specialist. Further, some of the most prevalent metaphors in his writing are ones that go unacknowledged.

A survey of the available literature on Dawkins' use of metaphor reveals that philosophers, though aware of both the insightful nature of Dawkins' metaphors and some of the difficulties his metaphors raise, have provided few extended discussions of the function of his metaphors in his explanations. Critical responses to Dawkins' work rarely take place within a well-defined philosophical theory of metaphor. Further, critics of Dawkins' writing often fail to draw out the implications of their criticisms for the use of metaphor in explanations in evolutionary biology more generally.

1.5 Outline of Argument

It is my contention that Dawkins' view is based on an attenuated and traditional view of metaphor - a view that denies serious cognitive function to metaphor. The literal/metaphorical distinction that Dawkins employs needs to be replaced with a more robust philosophical theory of metaphor. Once this replacement has been made, some of his cherished literal concepts can be seen as metaphors. When these concepts are viewed as metaphors, we can more easily identify both the assumptions Dawkins makes and his metaphysical biases. The elaboration of the positive and negative functions of metaphor in Dawkins' writing leads to a more general view of the function of metaphor in explanations in evolutionary biology: Some metaphors may be stylistic or didactic devices, but others are essential cognitive instruments that enable us to discriminate and articulate real patterns in the world. Our recognition of these patterns allows for successful prediction and explanation.

My argument will be divided into two parts. I will begin Chapter Two by outlining what I take to be a predominant view of explanation in evolutionary biology. I will then locate my investigation of the function of metaphor within a general discussion of metaphors, analogies and models in scientific theorizing and briefly outline the most relevant metaphysical and epistemological issues that accompany discussions of metaphor.

In Chapter Three, I will give a more detailed description of the worries that metaphor raises primarily for the scientific realist and contrast these with Janet Kourany's criticism of the realism/antirealism debate. Kourany argues that we must set these debates aside in order to engage in a more socially responsible philosophy of science. In light of her criticisms and because investigations of metaphor seem to inescapably raise metaphysical and epistemological questions, I suggest that we ought to look for a philosophical position that makes minimal commitments in these areas. Arthur Fine's "Natural Ontological Attitude" appears to be an attractive position for this goal, but I will argue that instrumentalism is a more adequate position.

In Chapter Four, I will introduce three views of metaphor. The first two I will discuss, dubbed the substitution view and the comparison view, are traditional views of metaphor. Though each describes some aspects of metaphor, they prove to be insufficient as general accounts of metaphor. The third view, Max Black's interaction view of metaphor, describes metaphor as the interaction of two isomorphic systems. This view accords metaphor great importance, accommodating the view that metaphor has cognitive significance. I conclude the chapter with a discussion of some possible difficulties for Black's view that metaphors are "cognitive instruments" and attempt to resolve them in a way consonant with the instrumentalist position adopted at the end of Chapter Three.

Chapter Five is a summary of Richard Dawkins' gene's-eye view of evolution as it is developed in *The Selfish Gene* and *The Extended Phenotype*. I end the chapter with a brief description of the view of metaphor that I think Dawkins holds. Chapter Six is an extended examination of three metaphors that pervade Dawkins' writing, namely, the gene as a *replicator* in the processes of heredity and selection; the gene as a *ruler* of development and behaviour, and the gene as a *reasoner*. It is my intention in Chapter Six to show that Dawkins' use of these three metaphors confuses the metaphorical description with the phenomena being described. The interaction of the subjects of the metaphor provides insight, but they equally serve to limit artificially the phenomena being investigated. The successful role these metaphors play in prediction and explanation suggest that they are more than *mere* metaphors.

I will attempt to draw together the results of Chapters Two through Six together in Chapter Seven to argue that if metaphors are to be predictive instruments, then they must somehow connect with successful induction. I argue that they do this by making salient patterns in the observable world, patterns that by their nature are 'there' in the world. This argument requires that we modify our provisional instrumentalist position in a way that brings us closer to realism. I end by exploring some of the challenges this view might face and possible directions for further inquiry.

Chapter Two

Exploring the Function of Metaphor in Evolutionary Biology

2.1 Explanation in Evolutionary Biology

Any exploration of the function of metaphor in explanation requires a description of what explanation is taken to be. What kinds of explanation does evolutionary theory purport to give us? Broadly speaking, evolutionary biology draws together and unifies subjects as diverse as ecology, paleontology and genetics in an attempt to work out the consequences and details of two ideas, the tree of life and natural selection (Sober 2000, 7). Spelling out the details of the tree of life - to say which species are descended from which others and which characteristics evolved in a lineage - is one goal of evolutionary explanation. This kind of explanation aims to describe the changes that have taken place throughout the vast history of life on earth and is constructed on the basis of evidence such as the fossil record and inferences from present patterns of change and the present traits of organisms (Futuyma 1986, 13). Although an historical event may be inferred with confidence, the causal mechanism may remain unknown. For example, the origin of flowering plants in gymnosperms is known, but the exact mechanism that gave rise to that diversification (for example, pollination, seed covering or vegetative architecture) is difficult to discern (14). So, these historical explanations tell us the possible paths that evolution might have taken, but not the causes for these paths being taken.

The other goal of explanation in evolutionary biology, then, is to say why the course of evolution has taken the path that it has, to explain the genetic and environmental mechanisms of evolutionary change, the predominant one being natural selection. For example, the possibility that very similar traits can arise from similar selection pressures can help us to make sense of the fact that two species that do not share a recent common ancestor display comparable traits. The determination of these mechanisms is a result of finding correlations between repeated events and the conditions within which they occur, and through the formulation of models (either verbal or mathematical) that are tested against observations and experiments (14).

Lennox (1992) states that Darwin's description of evolution by natural selection gives us an explanatory pattern - an abstract description of the kinds of entities and processes that would be included in any adequate evolutionary explanation (271). Darwin's theory, he says, specifies five fundamental attributes of organic life: (i) populations have an ancestor-dependent history; (ii) members of populations inherit traits and pass them on to descendants; (iii) members vary with respect to heritable traits; (iv) there is competition between members for resources because of a geometric increase in population; (v) the environment is complex and changing (271). By giving us this explanatory pattern, Darwin has given us a model of a mechanism for a specific kind of change - evolution by natural selection.

Lennox tells us that there have been attempts to show that evolutionary change follows deductively from the above five attributes (271). If such an attempt were successful then explanations in evolutionary biology might be shown to exhibit the explanatory structure of the deductive-nomological model of explanation in which the phenomena in question can be derived from a general law and a statement of the relevant conditions (Kosso 1992, 54). But were such a formulation to succeed, Lennox says, it would miss the point of Darwin's theory because Darwin's theory is an attempt to state the causal processes that interact to mechanically produce a certain effect (Lennox, 271). While evolutionary theory postulates other causal mechanisms to explain why things are the way they are, natural selection is at the core.

If we want to avail ourselves of this model in order to explain a change, such as the shape of a horse's hoof from its ancestors, then our explanation must be able to specify the mechanisms suggested by our model and incorporate the relevant facts of natural history. Though evolutionary explanations may be invoked to make sense of a wide variety of phenomena, all these explanations will have the concepts of natural selection, fitness and adaptation playing important roles. Evolutionary explanations may also incorporate assumptions about the workings of nature. One common assumption is the "adaptationist assumption": a change in a feature or process is usually in response to a change in the environment, so the trait under consideration is assumed to be adaptive. A second assumption is the "engineering assumption": the trait in question is efficiently or even optimally designed for the role it plays (284). There is a tendency amongst evolutionary biologists to invoke models that embody these two assumptions even in the absence of empirical support for either the range of possible phenotypes among which nature must have selected or for the heritability of the trait in question and with little support for the idea that the trait is a feature that is actually optimized (286). This tendency has earned this approach the sobriquet "Panglossianism," in honour of Voltaire's Dr. Pangloss, a man who thought that even the most horrendous tragedies were for the best purpose (Gould and Lewontin 1993, 76). As we will see, both of these assumptions are conspicuous features of Richard Dawkins' explanations.

2.2 Teleological Explanations

One reason why the adaptationist and engineering assumptions may be so successful and attractive is that talk of functions and purposes is a necessary part of constructing genuine causal explanations. Francisco Ayala (1995) argues that the biological sciences, especially evolutionary biology, *must* frame explanations in teleological terms. Genuine evolutionary explanations necessarily make reference to the function or purpose of a structure or a process within more complex systems. The purpose or function of a given structure, for example, a human hand, is explained by reference to its ability to increase the reproductive success of the organism that has it. Explanations of the structure of non-biological entities need not be framed in teleological terms because their structure is not a result of natural selection; the shape of a mountain came about by geological processes, but its shape did not come about to serve some end, such as, the greater reproductive capacities of mountains (273). That teleological explanations seem to have explanatory force, even when operating as proto-explanations raises the problem of separating adaptive "just so" stories that explain the presence of a trait from real causal explanations. Just as in the non-teleological explanations of the physical sciences, a particular cause-effect relationship must be maintained between the *explanandum* and the *explanans* for a teleological explanation to be successful. So, the existence of seasons is explained by the motion of the earth around the sun not vice versa, but the sharpness of a knife can be explained by referring to the fact that the knife is made for the purpose of cutting (274). "Teleological explanations account for the existence of a certain feature in a system by demonstrating the feature's contribution to a specific property or state of the system…the contribution *must be the reason why the feature or behavior exists at all.*" (274 original emphasis)

Ayala suggests that we view teleological explanations in biology as having an internal component and an indeterminate component to distinguish them from teleological explanations that would more appropriately apply to the goal-oriented behaviour of humans. The teleological features explained by biologists are the result of a natural process devoid of conscious design - the design of a bird's wings for flying, for example - the internal component, and the result is one of several possible alternatives that imply that there is no pre-determined endstate - the indeterminate component (1995, 277). Ayala's main point is that strictly causal explanations of structures and processes cannot tell us the importance of these structures and processes. "Although teleological explanations are compatible with causal accounts, they cannot be reformulated in non-teleological form without loss of explanatory content." (269) If such explanations are necessary, this may point to an ineliminable role for metaphor in biological explanation. Two questions present themselves here: Firstly, are teleological descriptions and explanations necessarily metaphorical? Secondly, are they really necessary?

Allen, Bekoff and Lauder (1998) suggest that, though the concept of purpose is perhaps the most important foundational issue in the philosophy of biology, claims about function and design in explanations are not to be considered as literal or as metaphorical, but as referring "more or less circuitously to various natural processes or properties." (2) Teleological explanations may appear to set the patterns of explanation in biology apart from other sciences, but it is not agreed that they are required. Explanations for why organisms have the traits they do might be reformulated in statistical terms, obviating the need for teleological explanation (1).

Though adaptive explanations might provide evolutionary explanations with their "guts" (Lennox, 296), there is no full agreement regarding their explanatory status. What does seem agreed upon, however, is that explanations in evolutionary biology must satisfy a functional and historical dimension. That is, they must explain what x does and how x got there.

2.3 Metaphors and Analogies

In providing possible causal explanations and in making predictions, evolutionary biologists make use of a variety of metaphors, analogies, and models. The sizable number of books and articles on the nature of metaphor and its role in scientific theorizing and discovery suggest that, despite the acknowledgement of the cognitive and scientific role of metaphor by philosophers of science, the debate concerning just how significant the role of metaphor is remains unsettled and the compatibility of metaphor with scientific realism is still under debate (Marcos 1997). The unsettled debate about metaphor resembles two others, namely, debates about the nature and function of analogies and models. The use of analogies and models in scientific explanation may have a more respectable aura than the use of metaphor, but their use raises many of the same questions about the influence of culture and the possibility of maintaining a realist view of our understanding of the biological world. I do not focus specifically on models and analogies in this thesis, but metaphors, analogies, and models are closely connected. I want to say something, even if somewhat briefly, about these connections.

The idea that metaphor and analogy are closely related is an ancient one. As part of his description of metaphor in the *Poetics*, Aristotle says that one basis for the formation of metaphor is the analogy. "...[W]henever there are four terms so related that the second (B) is to the first (A), as the fourth (D) is to the third (C)...one may then metaphorically put D in lieu of B , and B in lieu of D." (1941, 1457b) Aristotle also gives an example of such an analogy: "As old age (D) is to life (C), so is evening (B) to day."(1457b) Based on this analogy, one might describe evening as "the old age of the day" or old age as "the evening of life."(1457b) According to such a description, an analogy is a kind of "highly selective similarity," by which people focus on certain commonalities between things and ignore others (Genter and Jeziorski 1989, 448).

It is usually assumed that we can divide analogies into two distinct kinds. One kind, a qualitative or material analogy (Hesse 1966, 68) is one where there is an observable qualitative relationship between two things. For example, drawing an analogy between a cell and a factory ignores the materials from which each is constructed, but focuses our attention on the fact that both cells and factories take in resources to maintain their operation and to manufacture their products (Genter and Jeziorski, 448). Comparing a familiar thing with an unfamiliar thing in this way might allow for useful insight or prediction about observable properties of the unfamiliar object or process (Maynard Smith 2000a, 178; Hesse 1966, 68). The second kind of analogy is the isomorphic or formal analogy. In a formal analogy there is assumed to be a one to one correspondence between the structural properties of the familiar object or process in the analogy with those of the unfamiliar object or process. Further, the causal relationships between the terms on each side of the analogy are thought to be identical. This structural and causal correspondence makes both sides of the analogy not only expressible in the same formal language such as mathematics, but also by the same expression (Maynard Smith 2000a, 178; Hesse 1966, 68). For example, John Maynard Smith (2000a) holds that there is an isomorphic analogy between the genetic code (the full sequence of DNA in a genome) and Morse code, one that is so close that it does not even require justification (183). A second example is the analogy between light and sound waves - both are expressible in terms of the same mathematical expression, a wave function (Hesse 1966, 28). The relationship between metaphors and analogies then is that qualitative or structural similarities between two things are perceived and expressed either explicitly in the form of an analogy or implicitly in the form of a metaphor. Metaphors are compressed analogies.

2.4 Metaphors and Models

As Ruse's remarks in Chapter One showed, models too are thought by some to be closely connected to metaphor. "A model itself is a metaphor, *but one that has been carefully pruned.*" (Midgley 1979, 447 original emphasis). By thinking of one thing in terms of another - that is, metaphorically - one perceives similarities. By carefully distinguishing between positive similarities (the ways in which two things are alike) and negative similarities (the ways in which two things are different) and retaining only the positive ones, one constructs a legitimate model (Midgley, 447). Max Black (1962) distinguishes between two kinds of model: scale models and analogue models. Scale models are "likenesses of material objects, systems, or processes, whether real or imaginary that preserve relative proportion"(220). Analogue models are models of "some material object, system, or process designed to reproduce as faithfully as possible in some new medium the *structure* or web of relationships in an original." (222 original emphasis) The analogue model is based on a formal or isomorphic analogy and so it is assumed that the analogy is justified because the model shares with its original the same *structure* (Black 1962, 223). A formal representation of the system captures this structural relationship. Based on the above, there are two ways we might view the relationships between metaphors, analogies and models: One view is that we explore the possible similarities between two things by creating metaphors and, by selecting and pruning those similarities that we feel are most relevant, we construct a model. The second view is that we construct a model on the basis of recognized analogies and, if we so desire, we can express these analogies in a more compact way as a metaphor.

Depending on which of the three is taken to be most basic - metaphors, analogies or models – our assessment of the nature and function of any one of these will determine the nature of the other two. If analogy is seen as the basis for metaphor and models, then determining the nature and function of analogy in theorizing will, to a large degree, determine the nature and function of metaphors and models. For the scientific realist, there are three main debates about the employment of metaphors, analogies and models in biological theorizing: (1) Are they necessary or additional to theorizing? (2) Are they constructive or merely descriptive? (3) Are they efficacious or pernicious?

2.5 Metaphors, Analogies and Models: Necessary or Additional?

Philosopher Mary Hesse begins her *Models and Analogies in Science* (1966) by contrasting two schools of thought regarding the use of models in science. For one school of thought, the "Duhemian"(named after the French physicist and philosopher Pierre Duhem), models and analogies might play an important role in suggesting theories, but this need not lead us to the conclusion that they are an indispensable part of theorizing. Models are useful psychological aids, valuable for didactic or heuristic purposes; they are suggestive, not strictly descriptive and ultimately unnecessary. There is no implication of truth in a model and many things can be aids in this sense, including dreams. The essence of a theory on this view is that it is a deductive system unencumbered by analogy, like geometry (4). The main strength of a model is its ability to concisely express the relevant aspects of a theory, but the terms of any theory can always be clearly separated from any particular model being used. The terms of the theory can be directly translated into the language of observable phenomena for any particular application. The Duhemian believes that because two models may contradict each other in their observable details, yet share the same formal expression, models are additional to the logic of theory (52).

The contrasting "Campbellian" school (named after the English physicist N.R. Campbell) holds that an essential aspect of any theory is that it is dynamic and so always capable of elaboration to account for new phenomena (4). The analogies contained within a model allow us to make novel predictions. Without models, the Campbellian believes, the extension of a theory to cover new phenomena would proceed arbitrarily and prediction in new domains would be impossible. To see a model as descriptive only ignores its "growing points", the very places that can be further elaborated (10). The Campbellian believes that, even if theoretical terms refer to so-called observables, these terms still need interpretation, and models allow us to do this (21). Formal considerations cannot by themselves supply an interpretation of a theory that allows the theory to be extended. Models are a necessary part of any activity worthy of the name of scientific theorizing. "Pure mathematics provides the *form* of an explanation, by showing what *kinds* of function would approximately fit the known data. But causal explanations must be sought elsewhere."(Black 1962, 225 original emphasis)

Many biologists express the view that metaphors, analogies, and models are useful, but ultimately replaceable. Others adopt the contrasting view that they are somehow an integral part of theories. As we saw in 2.2, the debate over whether or not teleological language is essential to evolutionary theory is one instance of the debate about whether metaphor is necessary or additional.

2.6 Metaphors, Analogies and Models: Description or Invention?

One aspect of the descriptive role of metaphor is that it helps scientists to understand phenomena by "mapping" a system of correspondences or descriptions from one system to another. The basis for this mapping is an isomorphism between a well-known set of relations and a set of relatively unknown relations. Metaphor "...is a way of aligning and focusing on relational commonalities independently of the objects in which those relations are embedded" (Getner and Jeziorski, 449). Scientific metaphors have their ground in structural relationships that are independent of us. These relationships constrain the metaphors that can be used successfully. This view sees metaphor as an elliptical isomorphic analogy. Metaphors, analogies, and models describe the way things are because of a 1:1 correspondence between the structural relationships within the two subjects of a metaphor.

A contrasting view sees the metaphorical description of phenomena as equivalent to invention or construction. This is a negative construal of a general view about the "theory-ladeness" of observation. What we say we perceive reflects prior theoretical assumptions. "Observation of *x* is shaped by prior knowledge of *x*." (Hanson 1958, 19) So, it is thought that metaphors lurk in the background of all our so-called observation descriptions, organizing the data that we assume to "read off" the face of nature. Some of these assumptions are implicit in the language we use to express our perceptions or observations and it is argued that a close inspection of scientific discourse should reveal any prior theoretical commitments being brought to bear on observations and theory-building. There is no descriptive statement that does not go beyond the "given" (Hesse 1966, 15).

Metaphor is perhaps one way in which the pictorial nature of vision and the linguistic nature of propositional knowledge are merged. Metaphors help us to *see* a digger wasp *as* having made a cost/benefit analysis of digging a new burrow or searching for an empty one. This "seeing as" allows for us to *see that* a digger wasp might do something different (invade a burrow) if the population density changes. However, the theory-ladenness of observation implies that neither "seeing as" nor "seeing that" in one particular way is inevitable or even necessary. This is not to deny that there are genuine cases of seeing something unknown or without metaphor. But what is seen in such cases is most likely incoherent with our background of established knowledge (Hanson, 20). Contrary to the Duhemian view, a scientist does not search for a deductive system that his interpretations of data fit. The interpretation of data is involved in the process of making a deductive system, the result of attempting to make data fit intelligibly with better-known data (Hanson 1958, 72).

2.7 Metaphor, Analogies and Models: Efficacious or Pernicious?

The presence of metaphor in scientific explanations can be viewed as efficacious or pernicious, but these views need not be alternatives, both may be held simultaneously. Writing within the context of social policy and technological innovation and drawing on Max Black's (1955) work on metaphor, Donald Schön (1993) agrees that the creator of a metaphor somehow "feels" the similarity between two domains. The construction of a useful metaphor is an intuitive and pre-analytic process that unites relevant similarities between two phenomena that cannot be immediately articulated. Only later, when formulated as an analogy, are the similarities stated explicitly. So, a synthetic paintbrush is improved by being "seen" as a pump and then re-described in terms of the relations between the parts of a pump. Schön highlights the fact that metaphor is an important heuristic device that enables scientists to organize, emphasize and suppress various features of phenomena under investigation in their attempt to achieve insight and to make novel discoveries. Schön's view offers us a third way of characterizing the relationship between metaphors, analogies and models. Metaphors are constructed prior to the articulation of the analogies upon which we base our models. Schön's view is closely related to one that sees the use of models and analogies as a convenient way of thinking about theories – they avoid the complications and difficulties of thinking explicitly about the language or symbolism by which the theory is represented (Braithwaite 1953, 92).

As was suggested in section 2.6, observation descriptions ("red-here-now"), the putative neutral language of scientific descriptions, are not written on the face of nature to be transferred directly into language. These descriptions are already interpretations, interpretations that depend on a framework of assumptions. Evelyn Fox Keller (1992) takes such a view of metaphor and language to imply that the metaphorical structuring of theories actually works to inhibit further constructive work, working insidiously to maintain the status quo:

> General mechanism[s] by which the particular conventions of language employed by a scientific community not only can permit a tacit incorporation of ideology into scientific theory but also can protect participants from recognition of such ideological influences and thus effectively secure the theoretical structure from substantive critical revision. (142)

These general mechanisms and conventions exclude from biological theory any phenomena that do not "fit" or threaten to undermine the ideological commitments that are "...built into science by the language we use in both constructing and applying our theories." (Keller 1992, 142) The putative literal nature of scientific writing is deceptive; there are underlying metaphors that surreptitiously guide our theorizing.

Worries about the perniciousness of metaphor motivate demands by some feminist philosophers for a detailed examination of the language of scientific theorizing: "We need a critical theory of metaphor in science in order to expose the metaphors by which we learn to view the world scientifically, not because these metaphors are necessarily "wrong," but because they are so powerful." (Stepan cited in Keller & Longino 1996, 7)

The fact that metaphors, analogies, and models make salient previously unsuspected features of phenomena is closely connected to the fact that they can be misleading, suppressing features of that same phenomena and making it possible for scientists to mistake the model or metaphor for the thing being modeled. There is also a danger that the characteristics of the objects in a model will be confused with the theoretical concepts and that the attributes that belong to the model-objects will also be attributed to the theoretical concepts. For example, Richard Dawkins asserts that he is not speaking metaphorically when he describes chains of DNA as "sets of instructions for how to develop"(1986, 136) or that the "information technology of the genes is digital."(137) We might not ordinarily think of chemical structures as "programs" or "instructions" in the same way that we have programs for our computer or in the way that we give instructions to a friend; these descriptions are simply metaphorical extensions from computational and semantic domains to the domain of biological investigation.

To suggest that DNA really is a program would be, for Colin Turbayne (1962), evidence of being victimized by metaphor. The victim of metaphor, Turbayne says, "accepts one way of sorting or bundling or allocating the facts as the only way to sort, bundle, or allocate them...he confuses a special view of the world with the world. He is...unknowingly, a metaphysician" (217). Turbayne argues that the geometrical metaphor used with enormous success by Descartes and Newton in their descriptions of optics and the nature of perception, was imported metaphysically into science because of its apparent success in explanation and prediction. Philosophers and scientists came to believe that the world really was limited to, not by, the description of the world given by the elaboration of their mechanistic metaphors. In contrast, when Charles Darwin asked his fellow scientists to imagine a process of nature that operates as if there were a selecting agent, was anyone (any scientist) really taken in by the implicit assertion of purpose or design in his anthropomorphism? (Young, 121)

Recent work in linguistics and psychology suggests that metaphor is an essential part of our cognition and not limited to pedagogy and heuristic. If this is true, it would seem that the prevalence of metaphor in evolutionary biology, both in professional and less technical writing is inescapable (Ortony, 1993). Basing their work partly on Michael Reddy's investigation of the pervasiveness of the conduit metaphor in our descriptions of communication, Mark Johnson and George Lakoff (1980), have explored the conceptual nature of metaphor as it is revealed to us through our everyday language. On the assumption that our language reflects our structure of thought, and acknowledging the importance of cultural practices in shaping metaphors, they argue that all of our abstract thought is grounded on primitive facts about our embodied experience. The way we experience and conceptualize the world is structured metaphorically.

In summary, evolutionary biology attempts to explain why the living world is as it is and to describe the historical path it followed to get here. The explanations are ideally causal in nature and allow us to make inferences about what the past might have been like and, based on this understanding, to predict what we might expect to find now and in the future. Whether metaphor is necessary in scientific explanation, and to what degree it is descriptive or creative is an issue of continuing debate. What seems in less doubt, however, is that metaphor has a positive, beneficial role to play in evolutionary biology, a role that is won only at the expense of constant attentiveness to its possibly pernicious aspects. As Richard Lewontin perceptively noted, "The price of metaphor is eternal vigilance."(2000, 4 original source misquoted)

Chapter Three

Beyond Realism and Antirealism

3.1 The Realism/Antirealism Debate

Contrasting views about the role of metaphor in explanation generally fall into two metaphysical and epistemological categories, those of realism and antirealism. As one might have guessed from my discussion so far, the use of culturally and technologically based metaphors in explanations help to motivate realist and antirealist debates. As a consequence, any attempt to try to delve into the nature of metaphor in biological explanation has a strong tendency to direct our attention toward realist and antirealist debates. The realist/antirealist debate forms the background for discussions of metaphor because the metaphysical and epistemological position one holds will structure the claims one is willing to countenance regarding the function of metaphors in explanation. For example, writers who are reluctant to allow metaphor a significant cognitive role (especially an irreplaceable one) are usually realists. On the other hand, an antirealist might be unwilling to accept that metaphors play any explanatory role in theories because antirealists generally reject the notion that theories are explanatory. In light of this, to investigate the function of metaphor, we need to commit ourselves to a set of minimal metaphysical and epistemological assumptions.

Debates about the efficacious or pernicious nature of metaphor in evolutionary biology and science in general highlight a metaphysical and an epistemological concern for many philosophers:

> What is the true nature of science? Is it objective? [Is it] something which tells us about the real world, out there? Is science a reflection or epiphenomenon of culture? ...Something which changes as society changes and which tells us less about reality and more about ourselves? Is evolution a social construction?" (Ruse 1999, 236)

For a realist like Richard Boyd, metaphors, despite their imaginative component, must refer to things in the world that exist independently of us. The success of science, including evolutionary biology, cries out for an explanation, and the best such explanation is that science tells a true story about the causal structure of the world (cited in Gross 1990, 241). On Boyd' s (1993) account, the use of metaphor is part of a general function in science of accommodating our language to as yet undiscovered features of the world; metaphors provide a way to introduce theoretical terminology for probable features of the world and are not conceptual make-believe. Boyd's view and other similar views presuppose that "nature has one and only one set of joints to which the evolving terminology of science comes closer and closer with time." (Kuhn 1993, 541)

Contrasting with Boyd's view are the views of theorists such as Thomas Kuhn or Alan Gross both of whom deny Boyd's assumption that scientific terminology somehow gets closer to the so-called 'joints of nature'. Basing his case on a close reading of the style of some classic science texts, Gross argues that, because scientific knowledge is based on consensus, this knowledge is largely the result of persuading others about the truth of one's statements about what the world is like (4). Realism is just one of the tools that the scientist-as-rhetorician uses to achieve this goal.

On Gross' view, scientists are 'motivational realists'. Postulating the existence of mind-independent entities is a regulative principle required for conducting science (200). Scientists do not intend to create realist theories; the belief that such theories are possible is the "psychological anchor that makes [a scientist's] life meaningful." (200) If we accept Gross' characterization of scientific theorizing, we will see that scientists are not united by the variously postulated ontologies of science that they hold. Rather, they are united by a common conviction that their work has the same goal, namely, the discovery of the causal structure of the world (201). The truths of science are the result of a consensus around the coherence of a range of utterances, and not a fit between fact and reality (201).

3.2 Essentially Contested Concepts

One reason why Gross finds defenses of realism unconvincing is that he considers realism an "essentially contested concept", a phrase he attributes to the philosopher W.B. Gallie. In "Essentially Contested Concepts" (1956), Gallie articulated a historicist view of concepts used in aesthetics, social and political philosophy and metaphysics. His development of the notion of "essentially contested concepts" was an attempt to capture what he saw as "the most striking paradox of philosophy" – the seemingly continuous debates over the central concepts of philosophy. Though some philosophical debates may seem endless, Gallie believed that he saw in their persistence a clue to the function of philosophical concepts: they ensure intellectual vitality across the entire range of human inquiry and knowledge (Gallie 1968, 7). By vitality, Gallie seems to have in mind the continual strengthening of evidence and refinement of argument for one's position to win others over to one's view.

Gallie articulated a set of necessary and sufficient conditions that any concept must satisfy to be considered essentially contested. One condition is that the concept must be appraisive, signifying a particular achievement or way of doing something as valuable. A second condition is that the concept must signify a type of complexity that allows the achievement to be variously describable. It is also important for Gallie's view that all competing uses of a contested concept refer to the same exemplar. For example, a particular work of representational painting is picked out as "real" painting and therefore as an example of "a work of art." Various proponents accept the exemplar, but each picks out a different aspect of the painting as the quality that confers the status of "a work of art." This shared exemplar ensures that the groups are engaged in a legitimate dispute because they are talking about the same thing. The concept becomes contested and remains so because there is no quantitative scale to which the competing groups might appeal to measure achievement or success. Without such a scale, it is impossible to articulate a criterion upon which one way of painting or playing might be judged better than another. This can be clearly seen in the case of "work of art," for here there is no quantitative scale against which something can be counted as more a work of art than something else.

Gallie argues that although universal agreement is impossible, the dispute is nonetheless genuine. It is possible to show the rationality of a particular use of a term or change of concept: "...[T]he logic of conversion from one contested use of an essentially contested concept to another is on all-fours with the logic of every unique decision" and this makes conversions and changes justifiable (185).

Gallie's historical analysis of philosophical concepts enlightens us to the shifting definitions that seem to characterize philosophical disputes. If Gallie is correct and his conclusions also apply to metaphysical concepts, his emphasis on the role of persuasion would have radical implications for our understanding of science. Scientific claims would have no exclusive right to revealing the true nature of things, for all scientists would be claiming to do that. The adoption of one theory or practice or description would instead be the result of a successful campaign of persuasion with respect to some instrumental or aesthetic value. Gross thinks that the concept of realism conforms to the patterns of essentially contested concepts given by Gallie: All users of the concept agree that they are discussing the same subject (194); it is an appraisive term, signaling the fact that language describes the real; and finally, no arguments for or against realism are equally convincing, but none are capable of ending debate either (194).

Gross' characterization of scientific knowledge as the result of a campaign of persuasion connects with an antirealist view about metaphor and scientific explanation. Antirealists of this kind might espouse a 'strong creativity thesis' regarding metaphor, the thesis that metaphors actually create similarities rather than simply picking out already existing, but previously unnoticed similarities between two things (Black 1993). This antirealist stance is well expressed by Colin Turbayne (1962) in *The Myth of Metaphor.* When we alter or re-allocate the facts within a different metaphysic, we alter the facts. "The tomato re-allocated to the fruit class changes its taste history." (22) The implication for scientists is that, rather than investigating the world as it "really" is, they are investigating our experience of it, the world as it is independently of us being ultimately beyond our epistemic reach. The emphasis Turbayne places on the ubiquity and creativity of metaphor implies that any metaphor could be a source for theory building. Mary Hesse (1966) argues that, at least in the case of science, this line of thinking ought to be resisted because it leads to the view that theoretical models are irrefutable and this is a consequence that the methods of science themselves refute: " …no model even gets off the ground unless some antecedent similarity or analogies are discovered between it and the explanandum." (163)

Michael Ruse (1999) argues that the persistence with which our culture throws up new metaphors that lead to new scientific approaches and problems need not plunge us into a "crippling subjectivity" that debars valid comparison between theories and practices (249). Culturally produced metaphors, according to Ruse, enabled Charles Darwin to "push his ideas in a fertile manner into new dimensions of human behaviour and evolution."(244) Metaphors, the products of culture, further epistemic goals, such as consilience and predictive scope. Darwin could extend the metaphor of the division of labour to apply to social insects and competition between organisms. However, the epistemic norms of science are indifferent to race, sex, class and cultural heritage, theories are tested in "the fiery pit of experience." (249) The epistemic norms of science and the test of experience provide science with its objectivity. It is against this test of experience that we should judge the rightness or wrongness of metaphors according to their success at achieving the epistemic norms (252). Ruse concludes that science cannot be decisive on the point of metaphysical or epistemological realism or antirealism. The evidence suggests that either position might reasonably be argued for. Whether one is a realist or a antirealist, you can still make the distinction between good and bad science.

Focusing on concerns about the role of metaphor in evolutionary biology, Ruse thinks, tends to confuse two debates: It confuses the traditional philosophical debate over realism and non-realism with the new debate about standards and culture. Ruse's view on this matter is compelling because, not least for the practicing biologist, it allows one to get back to work and ignore what might seem like a rather arcane debate. It also puts the centrepiece of science, its values and methods, rather than an abstruse philosophical vocabulary, back in the spotlight.

As appealing as this approach sounds, however, Ruse has left some questions unasked. If culture can provide us with metaphors that further epistemic norms, how is this possible? What is it about the nature of metaphors that allows us to take a value-laden concept and apply it to a new domain allowing us to broaden the scope of our prediction or provide us with greater predictive accuracy? Why should we be convinced that metaphors are only a part of our experience and not constitutive of it? Though Ruse has hinted at the answers he might give to these questions, it is important to point out that they have not in fact been answered.

3.3 Metaphors and World Hypotheses

If the investigation of metaphor in explanations raises metaphysical and epistemological issues, Pepper's (1942) account of how our systems of explanation are themselves the results of metaphor demonstrates that it is possible to begin with metaphysical inquiry and end up at metaphor. Our knowledge, Pepper says, begins with uncriticized evidence, middle-sized or commonsense fact - our everyday perceptions and thoughts, a bird's song, your feelings for a loved one, the distance from your house to work, belief in space and time, etc. (39). If we attempt to analyze these common-sense facts, uncriticized fact becomes criticized fact, and is transformed as a result. Metaphor plays an important role in the process of analyzing and organizing commonsense fact.

Commonsense facts are items of evidence that ought to be doubted and that require corroboration. There are two types of corroboration: multiplicative and structural. The persuasive force of multiplicative corroboration is primarily social - the agreement of a person with a person about the same fact. The force of structural corroboration, on the other hand, "comes from the massiveness of convergent evidence upon the same point of fact."(49) Structural corroboration requires hypotheses in order to bring about this convergence of the corroborative evidence. The connective hypotheses are metaphorical in nature and the reason for this is as follows,

> A man desiring to understand the world looks about for a clue to its comprehension. He pitches upon some area of common-sense fact and tries if he cannot understand other areas in terms of this one. The original area becomes then his basic analogy or root metaphor. He describes as best he can the characteristics of this area, or if you will, discriminates its structure. A list of its structural characteristics becomes his basic concepts of explanation and description. We call them a set of categories. In terms of these categories he proceeds to study all other areas of fact whether uncriticized or previously criticized. He undertakes to interpret all facts in terms of these categories. ...Some root metaphors prove more fertile than others, have greater powers of expansion and of adjustment. These survive in comparison with the others and generate the relatively adequate world theories. (91)

Different metaphors will draw the lines in different places. According to Pepper, there are only four root metaphors that provide adequate world hypothesis: formism, mechanism, contextualism, and organicism. The world is filled with things that seem to be alike and describing this fact as precisely as possible in terms of the root metaphor of similarity gives rise to imminent formism or Platonic idealism. Mechanism is based on the root metaphor of a machine, the particular machine determining the species of mechanism that is developed. Mechanism is sometimes called materialism or realism.

Contextualism is based on the root metaphor of the dynamic dramatic event and it structures the view that we call pragmatism. The last hypothesis, organicism, is based on the root metaphor of integration and generates the position traditionally characterized as absolute idealism (Pepper, 141). The claims to truth of a given system (like the materialism of the physical sciences) is a result of the working out of the metaphor, not the result of a complete description of common sense fact because such a description is, according to Pepper, never possible. Pepper's account is important because it highlights the central role of metaphor in our explanations. However, it also highlights the use of metaphor in our attempts to understand the world around us, a function illustrated by Darwin's use of artificial selection in making sense of how organisms might become adapted to particular environments and by this process, come to vary.

I am in no place to make pronouncements about whether or not Pepper's four-fold scheme adequately captures all metaphysical systems. But I think it is suggestive that he begins with an examination of the creation of metaphysical systems and arrives at the idea that metaphor provides the foundation for our deepest explanations. One problem with Pepper's view, however, is that he has invoked similarity as itself a metaphor and this presents a difficulty in explaining, in a non-circular way, the process by which we group certain facts together in order to conform with a guiding "root metaphor." It is usually assumed that metaphors are constructed on the basis of similarities between things, whether explicitly recognized or not. Pepper's account also suggests that we have, to some degree, a choice about what metaphor we use to structure the facts of our experience. As I noted in 3.2, this is a contentious notion. If we accept Pepper's view, we also have to accept that our interpretation of metaphor will already presuppose some root metaphor that we use to organize and corroborate facts gathered about the function of metaphor.

3.4 Escaping the Debate

The interminable nature of the realism/antirealism debate has led Janet Kourany (2000) to argue that the motivating questions of the debate and the resultant answers distract us from pursuing a socially responsible philosophy of science. Investigations of the use of metaphor in biology and in science in general take place within philosophies of science that attempt to provide a comprehensive view of scientific aims, experiments, results, claims, theories, etc. But Kourany claims that the way that science affects and is affected by its social context is left out of many of these philosophies (S88). One of the basic problems in attaining a philosophy of science that includes this dimension is the concentration of attention on the realism/antirealism debate (S89).

The realism/antirealism debate tends to direct our thinking about scientific knowledge in either of two unproductive directions: scientific knowledge is either true or it is useful. Both of these directions are discussion stoppers. If scientific knowledge is true, then that is just how things are and we had better get used to it. If scientific knowledge is only useful then it still has value, but this view "buries the question of what the knowledge is useful *for*." (Kourany, S88 original emphasis)

Philosophers involved in realist/antirealist debates typically focus on either the aims or the results of science. The question of whether the aim of science is truth or empirical adequacy is both empirical and normative – aspects of which both sides of the debate tend to ignore. The empirical dimension is ignored because of a failure to consider evidence collected by scientists, researchers' professed aims, their actual research goals, and the source of their research funds (S91). Both realists and antirealists pronounce these aspects of science as peripheral to the actual aims of science. But, Kourany asks, how can these issues be considered peripheral when they so clearly shape the structure of science?

With respect to normative claims, the actual aims of scientists themselves are relevant in judging whether or not they ought to have the aims they do. For example, we might wonder about the political aim of sociologists who research environmental racism (S92). Those philosophers who are concerned with the results of science also typically ignore the aims of scientists, something that crucially shapes the evaluative criteria used to select research projects, experiments and theories. For example, the evaluative criteria frequently used by feminist scientists emphasize novelty, ontological heterogeneity (equal consideration of different entities) and applicability to current human needs (Longino cited in Kourany, S93). The selection of this criteria is motivated in part by the aim of feminist scientists to expose gender bias. Philosophers of science who omit such aims and criteria from their accounts are in a poor place from which to make pronouncements on the truth or empirical adequacy of the results of science (S95).

Kourany concludes that "we should either stop engaging in the realism/antirealism controversy entirely, or else engage in it in a more socially responsible way – by gathering the right kinds of empirical and normative data, and framing more helpful versions of the questions we want to ask."(S98) Ian Hacking's example of how the priority of funding and research for weapons today will radically shape the form of the research programs that are open to us tomorrow typifies the critical philosophy of science that Kourany has in mind (S97). I am not as optimistic as Kourany about the possibility of getting beyond the debate in order to engage in a more socially responsible philosophy of science. The language in which evidence, theories aims and research goals are talked about often contain unacknowledged metaphorical components. Pursuing a critical philosophy of science will still require that we have a clear understanding about how we conceive of metaphor, how we think it functions and the implications this has for making and adjudicating competing scientific claims.

Investigating the function of metaphor and confronting the metaphysical and epistemological concerns that accompany such an investigation is one way to engage in the socially responsible philosophy of science that Kourany advocates. Questions about metaphor help us to focus on the transfer of ideas from one domain to another according to one's aim, and to question why biologists choose one metaphor rather than another if, in fact, choice is an option. I hope that my survey has shown that discussing the role of metaphor brings with it a tendency to direct discussion toward debates about realism/antirealism, debates that Kourany thinks we would do best to avoid. In light of this difficulty, I think that the best strategy is to find a provisional philosophical position that carries a minimal amount of metaphysical and epistemological baggage. Arthur Fine's "Natural Ontological Attitude" has presented itself as just such a position.

3.5 A Middle Ground? Fine's "Natural Ontological Attitude"

According to Arthur Fine (1986), the "Natural Ontological Attitude" [NOA] is a minimalist attitude about the metaphysics and epistemology of scientific claims, a true alternative to realism and antirealism. Fine characterizes realism as any position that makes the following claims: 1) There exists a definite world that contains entities with relations and properties that are largely independent of human acts and agents; 2) It is possible to obtain a reliable and relatively observer-independent account of these entities. The addition of one other claim produces scientific realism: 3) Science aims at and is more or less able to provide an articulation of the structure of this world through its theories and principles. The thesis that the truths of science correspond with features of an observer-independent world connects the third thesis to the first two. Antirealism, Fine says, denies the first three of these theses and truth as correspondence, and replaces the notion of truth-as-correspondence with an account of the aim of science in terms of truth-as-acceptance or in terms of empirical adequacy, thus dividing antirealists into those who are "truthmongers" and those who are not. Both the realist and antirealist positions, Fine argues, are unsound.

Realist arguments of the "ground-level" variety affirm that the particular results of sciences, for example, successful prediction, can only be explained by realism. All ground level varieties of realism make the same assumption that the entities, relations and the properties used in explanatory hypotheses; they are approximately true. These arguments, Fine says, beg the question "as to the significance of explanatory hypothesis by assuming that they carry truth as well as explanatory efficacy." (115)

Methodological arguments for realism contend that the success of particular methods of science can only be explained by realism but fail in their attempt to give an answer to what Fine calls "the problem of the small handful."(117) This is the problem of giving an account of how a small handful of alternative theories, a small handful of seriously considered competitors of those alternatives, and a strong family resemblance between those competitors and previous theories (theoretical conservatism) can produce good successor theories. Realism has no account of why there are only a few alternative theories that cannot be matched equally well by instrumentalism, an antirealist position: it is very difficult to come up with alternative theories that conform to the empirical constraints imposed by already successful theories. When we turn to the reason for serious competitors, any answer cast in terms of "being well confirmed" assumes the truth of the ontology and laws in question. There is no valid inference of the form "T is well confirmed, therefore the objects and laws required by T exist" (117). This is demonstrated by shifts in scientific ontologies, such as those revealed by studies in the history of science. The answer to the third question of why this method produces good successor theories, uses the notion of the transfer of approximate truth from one theory to another. The realist can only explain at best why successor theories cover the same ground, an argument open to the instrumentalist as well. " If we insist on preserving the well-confirmed components of earlier theories in later theories, then, of course, the later ones will do well over the well-confirmed ground."(118) Further, the history of science reveals a more acute problem, for it shows us that scientific tinkering frequently leads to failure. The problem for the realist, then, is how to explain the occasional success of a strategy that usually fails (119).

When we turn to the antirealist arguments, Fine says, we see that they too are unsound. Antirealists recognize the fact that the great failing of the truth component of realism is that there is no way to verify independently a correspondence theory of truth, thus making truth inaccessible. "Truthmongering" antirealists attempt to recast the concept of truth in terms of acceptance. That is, to say that a proposition is true is just to say that there is a set of people who would accept the proposition under a certain set of circumstances (139). Such an account of truth, Fine argues, is an acceptance of a behaviourist assumption that the "working practices of conceptual exchange exhaust the meaning of the exchange, providing it with its significance and its content."(140) Thus, antirealists of this stripe arbitrarily fix the limits of the meaning of conceptual exchange. Fine argues further that casting truth in terms of judgments about what people would accept in certain circumstances results in a regress. To arrive at such a judgment, we would need to extrapolate from what is the case to what would be the case given certain conditions. Establishing what is the case, however, itself requires successful extrapolation (141). Fine even goes so far as to doubt the intelligibility of such accounts, contending that if you think that you understand them, you have probably read into them some other idea of truth, even the idea of truth-as-correspondence. Thus, antirealists might make legitimate points about how people use truth concepts, but that does not mean that they can limit the significance of truth to its use. The only way to stop this interminable debate, says Fine, is to stop thinking of truth as a substantive thing, something that can be "captured" by some theory, picture or account (142).

An alternative to the truthmongering antirealism is constructive empiricism. Constructive empiricism affirms that the aim of science is to produce empirically adequate theories, and that the acceptance of a theory is nothing more than the belief that such a theory is empirically adequate. "Empirical adequacy" refers to the fact that the consequences entailed by the theory about the things we could observe are true (Bird 1998, 290). According to constructive empiricism, the adequacy of a model does not entail that all of its parts correspond to the world. Models are good if they "fit those phenomena to be saved." (Van Frassen 1988, 138) Fine criticizes this view, first, on the grounds that the criterion of empirical adequacy requires an a priori distinction between entities that are observable and those that are unobservable, so "forcing the hand of science where [constructive empiricism] is supposed to follow it."(144) Second, we might wonder why we should choose to restrict our belief to the level of observables, those things we can perceive by our own senses rather than detectables, what we are able to perceive by virtue of sophisticated scientific instruments like cloud chambers or electron microscopes. Any worries that our entities, such as atoms, might be artifacts of our instruments can be waylaid by sufficient caution and rigor in testing. So, Fine holds that our instruments provide us with compelling evidence that we actually do detect atoms. The rejection of belief in unobservables is arbitrary and the same would go for other "information-bearing modalities."(146) Antirealist views that appeal to empirical adequacy or require a rigid distinction between what is observable and what is not ought to be rejected.

All versions of realism and antirealism, Fine concludes, have attempted to provide science with an interpretation, with a context with which to make sense of its aims and results. But why make this assumption? The traditional framing of the realism/antirealism debate has obscured the fact that there exists a viable metaphysical and epistemological position, a "core position", that is added to in various ways to produce either realism or antirealism. This core position amounts to the attitude that "I accept the truth of science in the same way that I trust my senses with respect to everyday, middle-sized objects."(127) This acceptance, Fine argues, is the conclusion of a modest argument: 1) I trust the evidence of my senses about the existence and features of everyday objects; 2) I have similar confidence in the check, double-check, triple check of scientific investigation; 3) If scientists tell me that there are molecules and atoms and quarks with attendant properties and relations, I should believe them. Truth, Fine concludes, is openended, growing with science. The significance of any questions regarding truth or belief are answered in terms of the practices and logic of truth judging, a practice that draws on past judgments, present knowledge and reevaluates itself as it goes along. By accepting these truths of science in the same way as our everyday perceptual truths, Fine means that these truths are taken into one's life and one's belief. One's behaviour changes to accommodate these sets of beliefs without trying to make a distinction between kinds of truth or the status of existence claims. The claims are judged instead in terms of their centrality to our lives and degree of belief (127).

3.6 NOA: Implicit Realism?

Is Fine's position really as neutral as he hopes? I do not believe that it is. Musgrave (1989) argues convincingly that if Fine is committed to accepting the referential semantics of Tarski, then he is committed to realism. To advocate that "a sentence is true just in case the entities referred to stand in the referred to relations"(Fine, 130) and to suggest that this is based on our "ordinary referential semantics" as articulated by Tarski, is to already accept a theory of truth, something that Fine hoped to avoid. Tarskian semantics, Musgrave argues, do not form a part of our ordinary common-sense semantics or operate as any standard rules of usage, though they might explain parts of them.

By itself, Tarski's Convention T ('X' is true if and only if X is true) does not commit one to realism, but a robust realism requires only the conjunction of three theses, one of which is this convention:

(a) take statements about X's at face value for logico-philosophical purposes;

(b) apply Tarski's Convention T to these statements;

(c) accept some of those statements…as true. (Musgrave, 390)

The conjunction of these three theses, however, just is Fine's Natural Ontological Attitude! Fine might reject (a), but he must do so on the basis of a philosophical account of the impossibility of a correspondence relation. Such a worry is either based on a notion of independence that prevents us from causally interacting with things (a silly account according to Musgrave), or on account of the way in which we are somehow trapped within our language and concepts and unable to talk about anything else, a position that has its end in linguistic relativism (393). The only other reasonable option open to Fine is to remain non-committal about statements about X's, regarding all philosophical commitments as equal. If this is Fine's position, Musgrave argues, Fine is in trouble. Musgrave says,

> The NOA [pronounced "knower"] is supposed to 'accept', say 'There is a full moon tonight' and 'Electrons are negatively charged.' But does the NOA know what he has accepted? Remember, he leaves it open whether or not these statements are to be taken at face value, says nothing about how they are to be interpreted, what their ontological commitments are. The unphilosophical NOA does not just know nothing philosophical – he knows nothing at all. (398)

Musgrave's argument provides good reasons for us to be suspicious of Fine's assumed neutrality. My reading of Fine's claim that "It is fundamental to NOA that science has a history rooted in everyday thinking"(149) raises further concerns. On one reading, this claim seems almost trivial, for we had to get science from somewhere and a reasonable explanation is that it grew out of attempting to solve our everyday problems, problems the solutions to which became increasingly complicated. But though science might have its roots in our everyday thinking, it does not follow that it *is* our everyday thinking. So, what is the relation between our everyday thinking and scientific methodologies and theorizing? Why should we accept Fine's plea that we ought to believe in scientific truths *in the same way* as we do our everyday perceptions? This raises the question of whether we accept scientific claims *for the same reasons* that we accept everyday claims. Lastly, I think Musgrave's criticisms of Fine's position help to drive home Pepper's main point concerning world hypotheses: we cannot begin to make sense of our experience, cannot give an explanation, without imposing some organizing scheme that makes substantive assumptions about the nature of our experience.

3.7 NOA: Instrumentalism?

Given Musgrave's and Kourany's criticisms and Fine's construal of what it means to take things as true, I suggest that the best provisional philosophy of science at this stage is instrumentalism. Traditionally, instrumentalism denies that theories can be evaluated in terms of their proximity to truth and emphasizes instead the legitimacy or adequacy of theories and models by reference to their performance in experiment, engineering, and similar contexts (Bird 1998, 125; Keller 2000, 74). The instrumentalist emphasizes the predictive efficacy of scientific theories, but withholds the attribution of truth or falsity to them. The problem with the adoption of empirical adequacy, Fine argued, was that the limit between observables and unobservables seemed arbitrary when confronted with the power of technologies that allow us to extend our senses. However, if Fine wanted a modest argument that connects with our judgments about everyday things, I think that an appeal to empirical adequacy and successful prediction is better than his mild realism.

NOA commits us to saying that, if metaphors are similar to theories, then metaphors, too, are determinately true or false. As we shall see in the next chapter, however, this sort of claim about metaphors brings with it a view of metaphor that I wish to reject. Theories are instruments that help us to make predictions and, as I hope to show, so are metaphors. Rather than seeing this philosophy as undercutting the most prized aspect of scientific theories, namely, their ability to explain, we ought to view it instead as according rather well with the reasons why we trust our senses. We don't trust them because they are true (as NOA might have us believe), we trust them because they work.

Because instrumentalism denies the attribution of truth to theories it is antirealist. One objection to the adoption of this position is that it entails the view that theoretical entities are mere fictions. If one accepts the view that theoretical entities are fictions, an obvious objection is that it does not seem reasonable to suppose that we can explain the behaviour of something by saying that it behaves "as if" x were the case at the level of unobservables (Salmon 1984, 6). We ought to commit ourselves to the fact that if we explain an event by saying that "x caused y," x must exist. As I intimated in Chapters One and Two, "as-if" talk is a central feature of models, analogies and metaphors in science. But in the development of explanations these are frequently modified or rejected, so this objection seems too strong. The "as-if" talk of instrumentalism is consistent with the "as-if" character that seems implicit in metaphors, analogies and models.

Adopting an instrumentalist philosophy of science helps us in several ways. First, we are not committed to claiming truth or falsity regarding theoretical statements in the sense of the realist, so we avoid making unnecessary existence claims prior to our investigation. Second, the problem of the small handful can be explained (that the instrumentalist can do this is part of Fine's argument against realism). Third, it captures the motivational aspects of the minimalist view that Fine desired. That is, on an instrumentalist view, we can explain what it means to accept scientific claims in the same way as ordinary truths: we accept scientific claims because they are capable of making predictions. If Musgrave is correct in thinking that NOA is a robust realist metaphysics rather than the instrumentalism that it at first appears to be, then I think that instrumentalism, as I have construed it, comes closer to our workaday, pre-reflective state than does Fine's NOA. If Fine wanted a truly minimalist philosophy of science that avoids redefining truth, he should have opted for instrumentalism.

Recall that Kourany's criticisms against realist/antirealist debates were directed toward instrumentalism. However, I question her criticism that instrumentalism does not provide us with a critical outlook. On the view that theories are intellectual tools, questions of use and applicability are at the forefront of theory evaluation. Instrumentalism can affirm that the history and practice of science, a chronicle of its problem setting and solving, provides both a rich context for its meaningfulness and an invitation for us to ask pragmatic and socially critical questions: Why is *this* a problem that ought to concern us? Why is it an important problem *now*? What is the evidence for this? Why do you see the problem *that way*? And these are not questions limited to science, but ones applicable to projects in the social sphere in general. Kourany's worry that truth cast in terms of utility somehow allows us to ignore the question "useful for what?" is not a worry about instrumentalism. *That* worry concerns people's willingness to be concerned about the projects that they are implicitly engaged in by virtue of living in a society. That is a worry about people being uncritical in general, not a problem uniquely generated by instrumentalism.

Any other criteria that are supposed to aid the realist in making her case are easily explained on instrumentalist grounds. So, for example, Kosso (1992) argues that giving up realist commitments means giving up what is really interesting about science, namely, understanding what is happening in the realm of unobservables. Why should we be convinced that *this* is the interesting part about science? It is difficult to see how science would be demonstrably different if we adopted the attitude that theories help us to frame problems and develop theoretical tools for solving those problems rather than the attitude that science must decide in exactly what way "The world, both those features that are manifest and those removed from experience, is in fact one way or another" (Kosso, 95) We could think of this approach as being analogous to Pascal's wager: If realism is true, more instrumentally efficacious theories are more likely to be true than not. Focusing on instrumental virtues will most likely produce theories that approximate the truth to a greater and greater degree. If realism is ultimately untenable or false, as the antirealists hold, then we have still produced instrumentally useful theories and nothing has been lost. It looks as though instrumentalism allows us to occupy the middle ground that Fine so desperately wanted to claim.

Framing an explanation in terms of what happens at the unobservable realm helps in prediction, but we do not need to make any existence claims about these entities, nor need we believe in the entities and processes postulated in these explanations in the same way as we do everyday events. Instrumentalists are not bothered by the fact that something that seemed well justified and true (from the realist's point of view) turns out to be false (from the realist's point of view). Instrumentalists cannot be criticized for ignoring other virtues of theories either. Virtues such as fruitfulness (explaining the greatest amount of phenomena with the least theoretical additions), coherence with already accepted beliefs, etc., are instrumental virtues based on our limited cognitive powers and our unwillingness to radically alter entrenched beliefs. Nothing is to be gained for any practical purpose by linking them to Truth.

I have argued that, despite Fine's criticisms, instrumentalism has virtues that still recommend it over Fine's NOA. Before closing this chapter, I will describe what I take to be two of the most serious challenges to instrumentalism. The first challenge for the instrumentalist, and the basis for Fine's rejection of this form of antirealism, is the distinction the instrumentalist makes between 'observables' and 'unobservables'. Fine argued that the instrumentalist fails to provide a clear distinction between those things that are observable and those that are not. By the term 'observable', scientists usually mean the ability to make reliable reports about an entity or process (Bird, 132). On the basis of this definition, scientific reports give us reasonable evidence for our belief in the entities and processes that the instrumentalist wants to characterize as unobservable. This, says Fine, demonstrates that instrumentalists have no reason not to assign truth and falsity to theoretical entities and processes.

The instrumentalist traditionally confines talk of truth and falsity to things that can be discerned with unaided senses (Bird 1998, 131). Scientific theories use terms that do not refer to observable entities and relations, so neither the theoretical terms, nor the theories they are part of are 'truth evaluable' (Bird, 126). Because the instrumentalist is committed to a distinction between observables and unobservables and a corresponding distinction between theoretical statements and non-theoretical statements, there must be a clear distinction between statements about observable things and statements about theoretical things. One problem with the instrumentalist's distinction between observational terms and theoretical terms is that the categories of observable/non-observable and theoretical/non-theoretical are orthogonal dimensions and allow for various combinations. Besides the entities that are observable/non-theoretical (colours) and non-observable/theoretical (neutrinos), there are also entities that qualify as observable/theoretical (gene) and non-observable/non-theoretical (the furthest star) (Bird, 131). On this basis it is argued that the instrumentalist's distinction is vague at best.

My reply to this challenge will have to wait until the final chapter, and it is not one that can resolve this issue. What I will argue, however, is that if this is a serious challenge for the instrumentalist, the scientific realist has an equally serious difficulty making a distinction between metaphorical and non-metaphorical language.

The second criticism of the instrumentalist is that they restrict the role of scientific theories to that of "predictive instruments" but fail to provide any convincing account of how prediction is possible if theories are not true or false. In my view, this is the strongest criticism of the instrumentalist position (Bird 1998, 143-147 also expresses this view). One possible reply that the instrumentalist might give is that predictive success is merely illusory, a function of disciplinary homogeneity with respect to the metaphysical assumptions being made. Such a reply has much in common with Gallie's view of essentially contested concepts that I described earlier. I think that there is an alternative reply to this challenge, but articulating it relies on my examination of Dawkins' use of metaphor and so, it too, must wait until the final chapter.

Having explored the ways in which questions about metaphor in science motivate realist and antirealist debates concerning the nature and status of scientific knowledge, and having looked at some criticisms of this debate, I argued that we ought to adopt the most modest philosophical assumptions we could. This modest, provisional position I have concluded is an instrumentalist one. In my brief discussion so far, metaphor has been variously defined as describing one thing in terms of another, understanding and experiencing one thing in terms of another, and as a process by which new concepts are expressed and suggested. I think that all of these descriptions explain parts of what metaphors do, but no single one of them captures the full range of the function of metaphor. Before we make a decision about how we ought to view the function of metaphors in biological theories, we should be clear about the possible functions that metaphor could have. And so, I turn in the next chapter to an examination of three competing views of metaphor.

Chapter Four

Three Views of Metaphor

4.1 Recognizing Metaphors

The nature and function of metaphor has been a topic of study since at least the time of Aristotle, but it was not until the middle of the last century that philosophers began to consider the topic worthy of sustained critical investigation. Max Black's essay "Metaphor," published in 1955, departed sharply form earlier accounts of metaphor by ascribing to metaphor a cognitive significance that it had, until then, been denied. To help the reader fully appreciate the ways in which Black's account of the nature and function of metaphor differed from earlier accounts and how sharply he departed from the attitudes and views of earlier writers, I will begin with an account of those views of metaphor to which Black was reacting and the reasons why he considered earlier accounts of metaphor inadequate. Following this, I will spend the remainder of the chapter explicating Black's interaction view of metaphor. I will end with a brief discussion of some possible difficulties raised by his view. The interaction account of metaphor is such a view. Though Black's articulation of the interaction view is a realist account, I will argue that it is also compatible with the minimal assumptions of the instrumentalist view adopted at the end of the last chapter.

To begin, I will assume that we can all make a *prima facie* distinction between metaphorical expressions and non-metaphorical expressions. I will further assume that there are features of statements that make them recognizably metaphorical. For the sake of analysis we must be capable of referring to the parts of a metaphorical expression and the expression in its entirety without equivocating on the word "metaphor." Max Black has developed a vocabulary for this purpose and I will adopt his terminology and use it throughout my discussion.

"The chairman ploughed through the discussion" would be recognized by most people as a metaphorical expression - the exercise of moderating a discussion being described in terms of ploughing. In such an expression, Black holds that there is a violation of the semantic and syntactic rules that govern the literal use of words and this violation acts to narrow our attention to one word or phrase on the basis of which we describe the expression as metaphorical (Black 1955, 273). The "violation" consists in the fact that a discussion is not something that we conventionally think of as being ploughable, a characteristic we usually ascribe to fields and an activity associated with farming. Our narrowed attention is the result of nonsense or self-contradiction in a literal reading of the expression and we quickly recognize the word(s) that are the cause of our confusion - "ploughed." In Black's terminology, "ploughed" is the "focus" of the metaphorical expression, while the rest is the "frame".

We know from the expression that the chairman does not physically plough anything because a discussion is not the sort of thing that you can physically plough through. More general statements of metaphor of the form "A is B", for example, "life is a journey," Black terms a *metaphor-theme*. A metaphor-theme is an abstraction from a metaphorical statement and can be used repeatedly, adapted and modified by a variety of speakers or thinkers on any number of occasions (Black 1980, 29). For example, "Life is a journey" might be an abstraction from the metaphorical statement "We've come a long way together." I shall thus use "metaphor" to refer to the metaphorical expression while its constituent parts will be referred to as the "focus" and the "frame."

The view that metaphor is central to the development of our language, and that it provides a way for us to come to understand previously unknown things, was not only uncommon, it was virtually inconceivable before Black's essay. Metaphor was considered a stylistic device used to beautify oratory, to capture the imagination and thus to help the speaker evoke certain emotions in the listener both for pleasure and for persuasion. Or, metaphor was considered a compressed simile, an expedient way of making a comparison that could be fully articulated at the listener's convenience. Both attitudes, dubbed respectively, the substitution view and the comparison view, are semantic views of metaphor. They limit the nature and function of metaphor to a use of language.

4.2 Origins: Aristotle's View of Metaphor

In my brief discussion of the relationship between metaphor and analogy in Chapter Two, I made mention of Aristotle's view of metaphor. There, I pointed out that Aristotle thought that analogy was one basis for the construction of metaphor. I want to describe his view in a little more detail here because both the substitution view and the comparison view find their original inspiration in Aristotle's analysis.

Aristotle's view of metaphor is usually inferred from his treatment of metaphor in the *Poetics*. He says, "Metaphor consists in giving the thing a name that belongs to something else; the transference being either from genus to species, or from species to genus, or from species to species on grounds of analogy."(1457b) An example of the first construction, that of transferring a general name to something more specific, is the expression "Here stands my ship" because the lying of a ship at anchor is the action of standing for a particular kind of thing. Aristotle's analysis of metaphor provided the basis for subsequent accounts of metaphor for the next two thousand years and limited views of metaphor in three ways. First, it made metaphor a function of individual words restricting metaphor to a use of nouns (Poetics 1457b). Second, it characterized metaphor as a deviation from literal language, one that involves a transfer of meaning from a word that properly applies to something to a word that does not. Third, it implied that any metaphor is based on recognized similarities between two things (Johnson 1981, 6).

With respect to the possibility of learning how to use metaphor, Aristotle says, "[T]he greatest thing by far is to be a master of metaphor. It is the one thing that cannot be learnt from others; and it is also a sign of genius, since a good metaphor implies an intuitive perception of the similarity in dissimilars." (Poetics 1459a, 1) Metaphor is not just a deviant use of language, the ability to use it is something either one has or one does not. Further, an ability to use metaphor points to an exceptional mind, a surprising thing to claim given that many, if not all, children seem to use metaphor in the process of acquiring language. As we will see below, the rejection of Aristotle's characterization of metaphor, not long after the beginning of the twentieth century, helped to initiate a new way of looking at how metaphors work and the meaning conveyed by their use.

4.3 The Substitution View

What does it mean to "plough through a discussion?" The substitution view assumes that the speaker could speak more plainly, but wants, for rhetorical effect, to employ a trope (a use of language where what is meant is stated indirectly) (Black 1955, 279). According to this view, metaphor is a substitute for literal expression and the meaning of the metaphorical expression is just the same as that of the literal expression. So, "the chairman ploughed through the discussion" is a more mellifluous way of saying that the chairman dealt with matters expediently, or that he pushed ahead, ignoring what others wanted to say. Determining just what is meant in this case may require paying particular attention to the circumstances within which the metaphor was uttered. Regarding this sort of case, Black says that we would need to determine how *seriously* the focus is to be treated. "Would [the speaker] be just as content to have some rough synonym, or would only *that* word serve?"(Black 1955, 277) Metaphor, on this view, lacks any cognitive or semantic value; it is purely a stylistic device. However, metaphors are often ambiguous and not capable of precise replacement. Indeed, it is just this characteristic that has been held to be one of the chief virtues of metaphor. With respect to this ambiguity or difficulty in paraphrasing, the substitution view gives us few clues about how to proceed other than to look harder and to be more creative in our paraphrasing.

A substitution metaphor is used for two reasons. First, it is used because there is a gap in our literal vocabulary. In such a case we use the substitution metaphor to plug this gap, a function Black terms "catechresis."(Black 1955, 279 note) An example of catachresis would be our use of "orange" to refer to the colour, when it had previously only applied to the fruit or vice versa. When a metaphor serves its purpose very well its metaphorical role disappears and it enjoys a new denotation. This is the common explanation of so-called "dead metaphors". Second, a metaphor is employed for stylistic purposes to provide enjoyment for the hearer as she casts about for the true meaning of the expression. The enjoyment is the thrill of the hunt, so to speak. Though both uses of metaphor are no doubt employed, the power of metaphor does not seem limited to these two aspects. When we say "life is a river" we are not simply replacing one statement with another. Rather, we are drawn to make connections between living and the movement of a river. It is this function of metaphor, its ability to suggest to the hearer inferences to be drawn, that the substitution view fails to capture.

4.4 The Comparison View

The comparison view of metaphor is a species of a more general view of figurative language within which tropes that involve semantic change are a result of some function of a literal expression. Examples of these tropes include irony, where the function is to invert one's meaning, stating the opposite of what one intends ("That was a smart thing to say" as a remark when something silly or stupid has been said), and hyperbole in which one purposely exaggerates an intended meaning ("I haven't gone to class in a million years"). Metaphor, according to the comparison view, acts to transform the literal meaning through the function of similarity or analogy. With the help of the metaphoric frame and contextual clues, the reader inverts the function to ascertain the ground of the intended analogy or similarity and the author's intended meaning. For example, if I say after a particularly good run of luck that "I live a dog's life" the hearer is to recognize that this is not a literal statement, not least for the reason that I do not live in a dog house or eat dog food. The hearer must cast about for some analogy or similarity, perhaps that dogs' lives are thought to be pleasant in some ways, for example, that they do not work and, if they are taken care of, are catered to in ways that makes them enviable. The hearer then understands that I am comparing the pleasantness of my life with that of a dog's for these reasons. (If a dog's life is thought of negatively then the corresponding analogies can likewise be made.)

A comparative metaphor can be replaced by a literal comparison by drawing on resemblances between the focus and the frame. Metaphor, on this account, is a suggested simile. For example, the expression "life is a river" might suggest the fact that the decisions we make create a winding course; that decisions made navigating a river are often difficult as are the decisions that we make in life; that traveling upon a calm river is easy going as are our lives when there are no pressing needs or decisions to be made, and so on. The comparison view seems to explain and allow for the suggestiveness that the substitution view could not. However, the comparison view of metaphor is not without its difficulties either.

Mark Johnson (1981) summarizes five main objections that are thought to show this view to be an insufficient account of the function of metaphor (26-27). The first objection, attributed to Max Black, is that the comparison theory does not tell us how we are to choose from the many similarities that may hold between two subjects. Black points out that similarities admit of degrees and are not "objectively given." Any attempt to achieve such objectivity will generate questions concerning the degrees in which something is more like one thing than another with respect to some property or set of properties. An increasing proportion of these questions reduces the effectiveness of any metaphor. On the strength of this objection Black takes the comparison view to be so vague as to be virtually empty. Further, Donald Davidson (1981) argues that we will forever be uncertain about the exact similarities being referred to (unless we are explicitly told) and that metaphor on this view is trivial because "...everything is like everything, and in endless ways."(1981, 209) Paraphrases that rely on similarities could be endless, an endlessness that "...springs from the fact that [the comparison] attempts to spell out what the metaphor makes us notice, and to this there is no clear end." (220 note)

The second objection, also attributed to Black (1955), is that the comparison view over-emphasizes the role of similarities, ignoring the fact that dissimilarities and disanalogies are often the basis for an effective metaphor, for example, "That was no picnic." "The insight we gain is often…more a result of dissimilarities that force us to imaginatively reconstruct our way of comprehending things." (26) The third, fourth and fifth objections to the comparison view Johnson credits to John Searle. The third objection is that comparison theorists see correctly that similarity sometimes plays a role in our comprehension of metaphor, but they commit the error of making similarity the basis for the *meaning* of all metaphors. Fourth, they make similarity the *sole* basis for the comprehension of metaphor. And, fifth, these two mistakes rest on three false assumptions: (1) Understanding metaphor involves comparing two existing things. Searle argues that this is false because we often use imaginary things as a subject of a metaphor as in the expression "Sally is a dragon." (2) Metaphorical truth depends on the truth of the attributions prior to their use in the metaphor. This is false. The truth of a metaphor is independent of the truth or falsity of the individual characteristics on which a metaphorical inference is based. So, when I say that "Richard is a gorilla" or "man is a wolf" and if I take this to mean that Richard or men are nasty, prone to violence and fiercely competitive, the metaphorical expressions will be remain true even though gorillas and wolves might be gentle and nurturing. What this shows is that our understanding of metaphor may have more to do with relations of beliefs rather than with accepted facts. (3) All metaphors are essentially similies. Perhaps the strongest objection to the comparison theory is that there are metaphors for which no literal similarities are available. Searle's classic example is "Sally is a block of ice." In this metaphor there can be no literal comparison between a block of ice and a human being because there are no relevantly similar properties between the two.

The comparison view has gone further than the substitution view in explaining some of the various characteristics of metaphor; however, as these five objections show, it is still lacking. Though Black does not deny that some metaphors are best seen as substitution metaphors or as comparison metaphors, he has good reason to believe that the comparison theory does not adequately capture either the meaning or our comprehension of metaphor.

4.5 The Interaction View: Background

The first serious challenge to the assumptions of the substitution and comparison views came in 1936 with the appearance of I.A. Richard's *Philosophy of Rhetoric.* In Richards' reflections on metaphor, we can see intimations of all modern accounts of metaphor that depart from the traditional view. In fact, Black's (1955) paper, "Metaphor," perhaps the landmark paper for most modern accounts, finds its inspiration in a brief remark made by Richards. A short summary of the challenges Richards raised will set the tone for a more sustained discussion of Black's work on metaphor.

Richards ([1936]1962) attributes to Aristotle's view of metaphor the "evil presence of three assumptions" that prevented the proper recognition of metaphor. First, the belief that the "eye for resemblances" necessary for the production of metaphor is an ability that some have, while others do not is simply false. Richards objects, "...we all live, and speak, only through our eye for resemblances." (89)

Second, Aristotle's assumption that everything else may be taught except the imaginative production of metaphor does not stand up to the most casual scrutiny. "As individuals we gain our command of metaphor just as we learn whatever else makes us distinctively human...." (90) We learn with and through the use of language, a language that can only aid us because of its metaphorical aspect.

The third, and worst assumption is that metaphor is something special, a deviation from normal language. In contrast, Richards contends that metaphor is "the omnipresent principle of all its free action."(90)

According to Johnson (1981), the dismissive attitudes of traditional views toward metaphor were further entrenched by the rise of logical positivism and its commitment to a verificationist view of meaning. Two of the main tenets of positivism, the separability of the emotive and cognitive functions of language and the belief that scientific knowledge could be expressed in literal and verifiable sentences, impeded the philosophical investigation of metaphor. Metaphor was seen as a use of language that evoked certain attitudes or feelings in the speaker, and metaphorical expressions were viewed as capable of restatement in literal terms without loss of cognitive content (17). In light of these facts, Richards' account is impressive not only because of its rebuttal of traditional views, but also because it appeared at a time when the traditional account was drawing new energies from a philosophy that in turn drew its inspiration from the successes of experimental science.

As an aside, it is worth noting that Richards' characterization of Aristotle's descriptions of metaphor constituting an "evil presence" is by no means uncontroversial. Marcos (1997), for example, argues that it was a narrow and superficial reading of Aristotle's work that led him to be seen as denying any cognitive efficacy to metaphor and, thus, as the father of the traditional view. According to Marcos, a thorough reading of Aristotle's biological treatises shows that Aristotle used metaphor, analogy and models in a way consonant with the interaction view (126). Referring to the prominent and important role of metaphor in some of Aristotle's writing, Marcos goes so far as to claim that "without metaphor, there would be no *De Anima* at all." (127) And that, according to Aristotle himself, the meaning of the central Aristotelian concept "act" can only be grasped by metaphor (127). This more thorough reading reveals that by his biological writing, more than by his few remarks on metaphor, Aristotle revealed metaphor to be an invaluable explanatory resource in biology.

Having rejected Aristotle's assumptions, Richards presents a characterization of metaphor that he believes better captures its nature and function. One of the first things he did that helped to clarify the nature and function of metaphor was to introduce a vocabulary, the tenor and the vehicle, that allowed theorists to talk about the component parts of metaphor without equivocating on the word "metaphor." Richards ([1936]1962) hoped that the introduction of this vocabulary would obviate persistent confusions between the entire metaphoric expression and the parts of the metaphor (96).

Rather than being a special use of language, Richards observed that metaphor permeates our language (92). Though the traditional view characterized metaphor as a function of words, Richards declared that it is "fundamentally…a borrowing between an intercourse of thoughts." (94) Metaphor is essentially cognitive. Metaphor is not just the relating of one thing to another based on similarity (93). Richards surmised that the deep metaphysical and epistemological implications of a considered account of metaphor were partly responsible for the superficial accounts of metaphor that had been produced up until then, an academic version of a child's hope that "if I close my eyes it will go away."

Richards makes three suggestions with respect to the ground for metaphor. First, the working of a figure of speech "has nothing necessarily to do with how any images may be backing it up." A metaphor does not necessarily rely on images in the mind (98). Second, not all metaphors are created based on similarity or resemblance between the tenor (i.e., the focus) and the vehicle (i.e., the frame). Richards gives the examples "giddy brink," "jovial wine," and "daring wound" as instances of metaphor that appeal not to similarities but to the association of one's experience with that object. So, our feelings of giddiness when we stand near the edge of a precipice are associated with the cliff (107). Also, successful metaphors might rely on "disparity action," the dissimilarities between two thoughts (127). Third, "A metaphor may work admirably without our being able with any confidence to say how it works or what is the ground of the shift. " (117) In other words, not all metaphors are capable of being rephrased as a set of similarity statements.

By rejecting the traditional assumptions of metaphor and by locating the basis and functioning of metaphor in thought, Richards provided the intimations of a theory of metaphor that Max Black would use almost twenty years later to sound the call for more serious philosophical investigations of metaphor.

4.6 Black's Interaction View

Black's (1955) examination of metaphor in "Metaphor" was his first attempt to articulate in greater detail I.A. Richard's insight that "...when we use a metaphor we have two thoughts of different things active together and supported by a single word, or phrase, whose meaning is a resultant of their interaction." (Richards [1936]1965, 51) One main puzzle about metaphor, Black contended, was that "...the writer or speaker is employing conventional means to produce a nonstandard effect, while using only the standard syntactic and semantic resources of his speech community." (1993, 23) Their effect may be non-standard, but our need for them is essential. Metaphorical expressions have, Black asserted, distinct capacities and we need them "...when there can be no question as yet of the precision of scientific statement."(1955, 284)

Black hoped that further elaboration of this view could remedy the defects of the two traditional views of metaphor: The requirement that metaphors be paraphraseable by picking out relevant similarities between two objects or processes, and the requirement that metaphors pick out already perceived similarities, thus denying metaphor any cognitive role. Black averred, on the contrary, that metaphor has a definite cognitive content. It is for this reason that he held that a literal comparison lacks the "ambience" and "suggestiveness" of the "imposed" view of the primary subject upon which a metaphor's power to illuminate depends (1993, 31). Black thought that metaphor somehow "imposes" its view because he thought there were instances in which we often say one thing is another

> ...in cases where, prior to the construction of the metaphor, we would have been hard put to it to find any *literal* resemblance between M [the frame] and L [the focus]. It would be more illuminating in some of these cases to say that the metaphor *creates* the similarity than to say that it formulates some similarity antecedently existing. (1955, 284)

Black's suggestion that the creation of the metaphor creates similarities, rather than the metaphor being constructed on the basis of recognized similarities could be called a 'strong creativity thesis', the thesis that metaphors are in someway 'self-certifying' - they create the very similarities required for their production. Black (1993), however, found this view of metaphor "unsettling" and pointed to his hedging remark that "it would be more illuminating" to view metaphor this way. He believed that metaphors reveal connections without making them (35). Though Black avoided committing himself to "worldmaking" metaphors, by rejecting the strong creativity thesis, we are still left with the puzzle of how a metaphor can be constructed without already having an idea of the relevant similarities that might provide the basis for the production of that metaphor. This puzzle brings us to the central feature of Black's interaction account.

4.7 The Implication-Complex

At the heart of Black's interaction view is the thought that the meaning of a metaphor is a result of the interaction of its parts (1955, 285). The frame of a metaphor extends the meaning of the frame onto the focus, thus creating a new meaning by forcing the reader or hearer to connect the two ideas and more dramatically, to reveal the very similarities that the comparison view would require prior to the creation of the metaphor. What makes this creative power of metaphor plausible or even possible? Black's answer lay in the systems of relationships attributed to the focus and those of the frame (1993, 27). In connecting the two subjects of the metaphorical expression (the focus and the frame) the reader makes use of an "implication-complex," a set of freely and readily evoked common beliefs or understandings about the subjects of the metaphor. For example, the expression "man is a wolf" calls to mind a host of certain platitudes about wolves that are commonplace within the speaker's and hearer's speech community, but ones that need not be true of actual wolves to apply. Calling a man metaphorically a wolf evokes a system of associated inferences about wolves - that they are fierce, travel in packs, have lives dominated by hunger and constant struggle, etc. These commonplaces are then applied to man and a new system of implications is developed for man, for example, that the life of a human is one of constant struggle and hunger and that his ferocity goes with such a life. The wolf metaphor brings into sharp relief those aspects of human behaviour that can be talked about in "wolf language" without strain, pushing into the background those aspects that cannot be so talked about. Thus, the metaphor organizes our view of man (1955, 289).

Black described this organizing force of metaphor more explicitly when commenting on the metaphor "a battle is a game of chess."

> Suppose I am set the task of describing a battle in words drawn as largely as possible from the vocabulary of chess. These latter terms determine a system of implications which will proceed to control my description of the battle. The enforced choice of the chess vocabulary will lead to some aspects of the battle to be emphasized, others to be neglected, and all to be organized in a way that would cause much more strain in other modes of description. The chess vocabulary filters and transforms: it not only selects, it brings forward aspects of the battle that might not be seen at all through another medium…to describe a battle as if it were a game of chess is accordingly to exclude, by choice of language, all the more emotionally disturbing aspects of warfare. (289)

Notice here that a metaphor does two things: it *highlights* or selects certain aspects for our attention, e.g., strategy, positions of advantage and threat, etc., and it *excludes* aspects of the phenomena, namely the human cost of war, both its emotional and physical impact. Likewise, to retain the interactive nature of the metaphor, we should look for ways in which our view of chess might be subtly changed as well. Chess becomes a more aggressive game, with the cost of losing pieces now seen as losing "men" in the battle of wits with your opponent. The emotional nature of the game is heightened and perhaps winning and losing become more important.

In a third and even more explicit explanation of the working of an implication-complex, Black asked us to consider the metaphorical expression "Marriage is a zero-sum game" and proceeded to spell out an associated implication-complex. Note that both subjects are seen as systems of properties and relations. This systematic element is central to Black's interaction view. The system of corresponding implications for this metaphor will depend on our interpretation of "contest," "opponents", and "winning." I offer a modification of Black's suggested correspondence as one example (1993, 29) (Table 4.1) :

	Marriage is a Zero-sum Game	
Game	Interaction	**Marriage**
(G1) A "game" is a contest	⟷	(M1) A marriage is a sustained struggle
(G2) Between two opponents	⟷	(M2) Between two contestants
(G3) In which one player can win only at the expense of the other	⟷	(M3) In which the rewards of one contestant are gained at the other's expense

The Implication Complex

Table 4.1

This metaphor makes suggestions and valuations that Black believed necessarily attach themselves to a game theory of marriage and colour our perception of it, a characteristic of metaphor that was earlier referred to as "ambience." In this case we might have a negative view of marriage because competition and struggle have become prominent characteristics of a love relationship. The identification of one game implication-statement with its correlated one for marriage need not be based on strict replacement, but may be based on similarity (G1 and M1) or extension ("win" in G3 which is extended to cover cases in which there may be no clear winner or loser as in M3). If we choose to interpret opponents differently, viewing them instead as the married couple and their society and winning as happiness, M1-M3 would remain unchanged, but the ambience would be very different. There is no need to think that we have this implication-complex already in mind when we bring the focus and frame together. Some metaphors are deeply "resonant" providing for a rich series of implications that are revealed after the construction of the metaphor. An implication-complex can become so detailed and so rich and intricate in its explication that the primary subject is all but forgotten in the expansion of the metaphor. For exactly this reason, Turbayne (1963) says of Plato's *Republic*, "the details specified in his model of the state are so complex and unusual that posterity...misnamed the whole work." (19)

Regardless of our interpretation of the implicative subjects, Black contended that the connection between the two implication-complexes is made by virtue of an isomorphism (strict sameness of structure) that the two systems share. Further, if we ignore the connotations that are projected by the frame or secondary subject we have an analogue-model, that is, some system or process designed to reproduce as faithfully as possible in some new medium the structure or web of relationships in the original (1962, 223). Black described analogue models as dangerous because the ability to model in different media introduces irrelevancies and distortions increasing the chance of fallacious inference. "Every metaphor," said Black, "is the tip of a submerged model," by which he meant that any given metaphor can be used to generate an implication-complex revealing some shared isomorphic relationship that is supported by the frame, thus modeling the ascriptions being made of the frame and not simply transferring attitudes and connotations from focus to frame (1993, 30). The transfer of meaning is not always restricted to associated commonplaces. They may be constructed, the writer establishing novel patterns of implication before using them as frames for her metaphors (Black 1955, 290).

The three main innovations that Black made that set his view apart from the comparison view are, (1) The suggestion that the focus and frame "interact"; (2) That the basis for the interaction is an entire system of implication rather than a discrete set of similarities; (3) That metaphors, by revealing relations, ought to be viewed as cognitively important. Because of the necessity of this system of implications, Black believed that any attempt to transpose an interactive metaphor into a literal equivalent would eliminate the cognitive content of the metaphor; a restatement "will not have the same power to inform and enlighten as the original." (1955, 293) The double subject in an interactive metaphor is incompatible with the substitution view that posited only a single subject, while the lack of any isolable ground for comparison in many cases is incompatible with a comparison view.

Let us summarize what has been said so far by looking at how the three views of metaphor (substitution, comparison, and interaction) compare with one another using a single example, the tree of life from Chapter One (Table 4.2).

	Basis for Metaphor	How it Works
Substitution	Substitution for a literal statement that is formulated prior to the construction of the metaphor.	• All similarities recognized prior to construction of metaphor. • Talk in "tree language" can easily be replaced by non-metaphorical talk of relations of descent, e.g., a description of shared features. • Metaphorical description unnecessary for understanding or learning anything new about relations of descent. • For rhetorical effect only.
Comparison	Perceived similarities between phylogenetic relations and the branching structure of a tree.	• Many individual similarities recognized prior to construction of metaphor. • Talk in "tree language" can easily be replaced by literal talk of relations of descent, e.g., a description of shared features. • Seeing some similarities might lead to noticing other similarities such as the tree being "unrooted". • Once noticed these similarities are describable in literal language, e.g., in terms of topological properties, and are cognitively equivalent to the metaphorical description, or more precise.
Interaction	Initial "felt" similarity, or some perceived similarity between phylogenetic relations and branching structure of tree, possibly initiated by diagram.	• The complex of associations for "tree" and for the relations of descent interact as systems. • Components of the two complexes may be related by similarity, identity, extension or analogy and so are not all related in the same way. E.g., "branching" may be a similarity relation because both a tree's branches and a phylogenetic tree's branches have spatial and temporal aspects. • The "tree complex" organizes our conception or relations of descent, and might also affect our understanding of trees. • Replacing "tree" language with non-metaphorical language changes our understanding. • Highly suggestive and open-ended with respect to showing us how species may be related.

Three Views of Metaphor

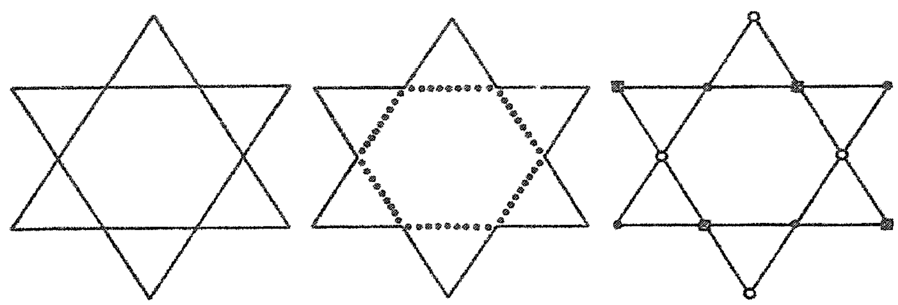
Table 4.2

4.8 Metaphor as a "Cognitive Instrument"

What does it mean to see one thing *as* something else? On the comparison view, one answer to this question might be that it means to "...think by proceeding from the known to the unknown, by extending a familiar term to an unfamiliar fact or situation." (Brooks and Penn Warren 1950, 375) On the interaction view, however, Black suggested that the concepts used in a metaphor change when they are brought together, possibly producing conceptual innovation. The interaction of metaphor relies on more than the extension of terms; it is the result of a reciprocal influence of one system of implications on another. Black contrasted the cognitive function of an interactive metaphor and comparison metaphor by means of an analogy: it is the difference between "looking at a scene with blue spectacles and comparing that scene with something else." (1993, 30) This analogy implies that one set of implications structures our entire perception of another phenomena. It also suggests that there is a process of cognitive structuring that is not only pervasive, changing our entire understanding of the phenomena, but that this change in understanding can be virtually instantaneous, proceeding without conscious or deliberate explication of the implicative complex. The eyeglasses metaphor also suggests that metaphors play an important role in structuring our perceptions and conceptions, a role that is strongly reminiscent of N.R. Hanson's famous claim that our observations are "theory-laden" - our prior assumptions about the world influence our choice of observations and our decisions about what observations are credible (Kosso, 115).

As the phrase "seeing one thing as another" suggests, using a metaphor involves noticing different aspects of something. Though no new perceptual information is imported, our comprehension or understanding of the thing is changed, much like the change in perception that can occur when looking at a double-aspect drawing (old woman/young woman) where we see the drawing first as one thing and then as another. Similarly, Black asserted that the Star of David can be seen in many different ways

(Figure 4.1): It is a star, or two overlapping triangles, or a hexagon surrounded by six small triangles, or three over-lapping parallelograms (1993, 31).



Seeing One Thing As Another
Fig. 4.1
(From Black 1993, 31)

These different aspects are different ways in which we might see the Star of David, the new frame reorganizing our view of it to force us to consider different structural relationships. But again, this does not capture the richness of an implication-complex that Black said is essential. Saying " 'Life is the receipt and transmission of information,' is at least to be thinking of life as the passage of information (but not that merely)."(1993, 31) Not as the passage of information merely because this metaphor has a system of inferences that go with it just as did "marriage is a zero-sum game." And, recalling the ambience of an interaction metaphor, looking at life through the lens of this metaphor will also shape our attitudes and values about life.

4.9 Ontological and Epistemological Productions

Metaphors, Black contended, have a strong creative role in cognition because they can generate new knowledge and insight by changing relationships between the things designated; they somehow reveal relationships without making them (1993, 35). Though Black shied away from accepting the stronger claim that metaphors might create similarities (the "strong creativity thesis"), he did affirm that metaphors reveal similarities and allow for the discovery of analogies of structure (isomorphic relations). This revelatory aspect of metaphor raises questions about the knowledge that metaphors make possible. How do we come to know new things via metaphor? In turn, this raises a question about the existence of the relations and properties revealed by metaphors. If the perceived similarities were not created by the metaphor, why couldn't we discover them in some other way? For example, why not think that inferential arguments are equally efficacious at making new connections and helping us to see things in new ways (Green 1993, 612)? What is the nature of the interaction between focus and frame that allows for this epistemological access that is unavailable in any other way?

To clarify what he meant by metaphor giving us access to the world, Black suggested that we consider how we might answer the question "Did X exist before it was perceived?" What comes closest to what he means by the ability for metaphors to "create similarities" is the answer to the question "Did the slow-motion appearance of a galloping horse exist before the invention of cinematography?"(Black 1993, 37) The motion picture camera mediated the slow-motion appearance of the horse allowing its feet to be seen clearly. Once seen in a film, however, this appearance became part of the world. Black thought that, in a similar way, metaphors make connections between similarities that allow for our perception of those same similarities and they are then present (37).

Did the camera make us see a galloping horse in some radically new way? Before anyone had seen the slow-motion horse people still knew the rhythm of a galloping horse from riding, listening and watching a horse gallop. They knew the order in which the feet hit the ground by watching and examining hoof prints on the ground. We might be tempted to assume that the slow motion picture of the horse galloping is still a horse galloping, only slower. But the point here is that without cinematography the slow-motion appearance of the horse would not be available to us at all. Slow-motion is not something done slower, it is the same motion only temporally stretched to allow us to see things that happen too rapidly for us to be aware of them. This new view reveals properties and relations that were heretofore unknowable. In a similar way, the puzzle of whether or not all four of a horse's feet left the ground at once when galloping could not be solved until, in 1878, a galloping horse had been "frozen" in midair by the photographic techniques of Eadweard Muybridge. Once this puzzle was solved, it changed the way we saw horses galloping by confirming what some thought but could not see to confirm and changing what others had previously believed. Seeing a horse galloping after this puzzle was solved was seeing it differently. Muybridge's photograph "captured aspects of motion whose speed had made them as invisible as the moons of Jupiter before the telescope…"(Markovits, 2003). The moons of Jupiter existed before anyone saw them in a telescope just as a horse's hooves left the ground before this fact was captured in a photograph. Likewise, interaction metaphors are cognitive instruments, epistemologically and ontologically on par with other paradigmatic instruments that reveal new things and allow for the production of new knowledge.

Metaphors play a representational role of showing us "how things are" in much the same way as other cognitive instruments such as photographs, maps, drawings and models. Correspondingly, it would be incorrect to call a metaphor true or false; metaphor is not a use of language that allows for that determination. Metaphors are not substitutes for "bundles of statements of facts."(Black 1993, 39) More correctly we can ask if the metaphor is apt, or correct or incorrect. Black concluded, "The imputed isomorphisms can…be rendered explicit and are then proper subjects for the determination of appropriateness, faithfulness, partiality, superficiality and the like." (1993, 38)

4.10 Interaction, Instruments and Obstacles

Black's interaction view of metaphor is appealing because it opens up the possibility for metaphors to play a role in our understanding and developing knowledge of the world. However, Black's view also faces several difficulties. First, if metaphorical understanding is a result of the interaction of associated implications, then it is not clear how metaphor operates in cases where little if anything at all is known about the properties of the focus, providing little with which to initiate interaction in a meaningful way. One response to this difficulty is to suggest that there is no need for the interaction to be perfectly reciprocal, and in fact it probably rarely is. The better-known system of implications is used to provide a provisional structure that can be modified and revised as investigation proceeds, and this is probably what happens in the elaboration of an interaction metaphor.

A second worry is a consequence of Black's contention that it would be incorrect to call a metaphor true or false. Truth is a traditional prerequisite for knowledge. Having denied the status of truth or falsity to metaphor creates a tension in the idea that metaphors can play a role in knowledge. This problem might be solved in one of two ways: Either we adopt a deflationary notion of truth that places minimal constraints on what it means for a metaphor to "represent" the world truly, or we need a view of knowledge that can allow metaphor to make a contribution independently of its veridicality. Either one would be satisfied by adopting NOA as our default metaphysical and epistemological stance.

The second solution is the one with which I think Black is most closely aligned and his view of metaphor, along with its realist tendencies, is compatible with NOA. However, at the risk of making too much of Black's reliance on the representational role of metaphor and his comparison with maps, I want to suggest that if the knowledge made available by means of metaphor is anything like that made available by maps, our provisional instrumentalist stance makes better sense of the representational role of metaphor.

The comparison of metaphors with maps suggests that metaphors have a prima facie compatibility with both realism and Fine's NOA. On either of these views, to the extent that maps or metaphors are accurate, they represent the world. "[F]or a map to fail to represent is for it to fail as a map."(Sismondo and Chrisman 2001, S41) However, despite this representational aspect, what maps depict is in part a function of their intended use and the resources available to cartographers. Like the features of a map, the implications of a metaphorical system that are made salient and the choice of which implications to elaborate are a function of the intended use of the metaphor and our knowledge of the focus and the frame. On the NOA view, scientific theories might plausibly be construed as maps of the natural world and as such they "don't embody any strong correspondence relation, but good maps, like good theories, brutely represent their subject matter." (S41)

This appeal to brute representation, as a replacement for correspondence, misses the ways in which maps in many cases are not attempts to simply reflect reality. Particular projections are chosen for instrumental reasons. For example, the Mercator projection preserves angles so that that observed bearings are consistent with angles on the map and compass bearings are preserved (S42). This projection, however, distorts ground distance. The Mercator projection is a good one, but only for a particular use, namely, navigation. Contour lines are also useful tools, but they are only representations of where an abstract plane intersects the landscape and as such distort the actual contour of the landscape (S44). Finally, boundaries and features of objects are sharply defined on the map, a sharpness lacking in the natural terrain. We sharpen the boundaries of those features that we are interested in (S45).

The boundaries that metaphors draw select and sharpen the features of phenomena in which we are interested for some purpose. Like maps, metaphors are made to be used. Only in the context of cartographic traditions can one say that there are good and bad maps (S47). Likewise, it may only be in the context of our attempts to make predictions that we can say that a scientific metaphor is good or bad, accurate or inaccurate, useful or useless.

Emphasizing the instrument-like function of metaphors and the role our interests play in dictating the forms of representation that metaphors allow makes the strong creativity thesis an appealing description of the ways in which similarities are brought to our attention - they are created along with the metaphor. Framed this way, metaphorical claims resemble the claims of some sociologists of knowledge who argue that the passivity of nature entails that it is only human activity that makes the facts of nature. The problem of the underdetermination of theory by evidence leads them to the conclusion that disputes about scientific conclusions are resolved on the basis of the interests they serve, rather than truth and good reasons (Wray 2001, 467; Longino 2000, 274) Technological innovations "make possible new thoughts expressible in new languages and so new 'facts.' " (Wilson 1996, 169) I want to resist this line of reasoning, however. Just because theories and metaphors are best seen in a instrumentalist light does not mean that any similarities will work or that there is no recalcitrance in the use of a metaphor.

Some instruments, such as microscopes, can and do enhance normal perceptual experience and create new experiences (Wilson 1996, 171). The view espoused by some sociologists of knowledge neglects the fact that instruments are prosthetic and as such "are causally connected to the world in ways that permit them to deliver information to us, as well as to interact with entities normally beyond our manipulations."(171)

Metaphors, like our paradigmatic instruments, "show us things that we did not expect and cannot ignore but that are inconsistent with pre-existing interests and values capable of altering them."(178) The use of metaphor does have a cultural component (particular metaphors are not available until the basis for the metaphor is), and so the constructionist claim that metaphors help to make facts is correct in recognizing this cultural dimension. Metaphors, like other instruments that contribute to knowledge, also help us to extend our understanding of the world into previously inaccessible areas and to challenge previously held beliefs. Metaphors, too, are "experience-altering and experience-producing technologies."(179) But what we learn via metaphor, like what we learn via instruments, is tested against our experience through the production of hypotheses and on the basis of which we make predictions. It is in this sense that I think metaphors might function in biological investigation and explanation. We can make predictions where before we could not.

To borrow a remark of Wittgenstein's, if we think of words as tools, then metaphors seem to allow us to adapt tools to different tasks, to make do with what we have.

> Think of the tools in a tool box: there is a hammer, pliers, a saw, a screw-driver, a rule, a glue-pot, glue, nails and screws. – The functions of words are as diverse as the functions of these objects. (And in both cases there are similarities) Of course, what confuses us is the uniform appearance of words when we hear them spoken or meet them in script and print. For their application is not presented to us so clearly. (Wittgenstein 1968, par. 11)

I am not sure if Black wanted to restrict the instruments to which the operation of metaphor ought to be compared to representational items such as maps and pictures, and there seems no reason to do so. The map metaphor is suggestive, but on the interaction thesis that Black holds, it might be illuminating to compare the instruments that can be used to describe metaphor. There is always a metaphor for the different functions that metaphors seem to play. They are maps or pictures; they are microscopes or telescopes; they are Rorschach inkblot tests, etc., and each metaphor carries with it different and perhaps incompatible meanings of what metaphors might do.

Emphasizing the making-do aspect of metaphor reminds us of the function of metaphor as catechresis. I have not denied that metaphors do function in this way, I have only tried to emphasize that this description was insufficient to capture all the ways in which metaphors function. Sociologists of knowledge, or "strong programmers", neglect the stubbornness of fit between metaphors and the world, while NOA neglects the instrumental nature of the instruments that "map" the world and the implications this has for our view of metaphor. If we are to adopt a version of metaphor that is compatible with our provisional philosophy of science, it is best if it is this instrumentalist construal of the interaction view.

Chapter Five

The Selfish Gene and The Extended Phenotype

5.1 Dawkins' View From the Gene

In 1976, Richard Dawkins, then a lecturer in zoology at Oxford University, published *The Selfish* Gene, a work of popular science that sought to provide a view of evolution that would show how most organismal and group traits, and especially supposed selfish and altruistic behaviours, could be explained as adaptations for the benefit of a single unit of selection, the gene. In partial response to critics of *The Selfish Gene*, Dawkins published *The Extended Phenotyp*e, a more technical work aimed at his professional colleagues that extended and elaborated his "gene's eye view" of evolution.[1]

Though Dawkins' stated aim is to give a biological account of altruistic and selfish behaviour, the general nature of his argument permits him to unify a broad range of phenomena, allowing them all to be explained in terms of a single causal mechanism –

[1] Both books have been revised and re-issued by Oxford University Press. *The Selfish Gene* was re-issued in 1989, *The Extended Phenotype* in 1999. I base my examination of Dawkins' views upon these revised editions. The second edition of *The Selfish Gene* contains two additional chapters that summarize parts of *The Extended Phenotype*, as well as notes that contain replies to his early critics. Dawkins' arguments in these new editions remain the same as in the first editions.

gene replication. Dawkins argues that this view, more than any other, gets the facts of biology right. Together, Dawkins' two books helped to reformulate central questions about behavioural biology in terms of the genetic theory of natural selection: Most traits are adaptations for the benefit of the genes that helped to produce those traits. These adaptations may be anything from the production of particular sections of DNA, hormones or pigments, to gross phenotypic traits such as long legs or aggressive behaviour, to even larger structures such as a beaver ponds and a bee hives that are detached from and many times the size of the organism that builds them. All the facts of life fall into place, Dawkins says, once we recognize that all evolutionary change is driven by the selection of genes (1989, xi).

In order to articulate such a wide-ranging and comprehensive view, Dawkins appeals to the notions that genes are *replicators*, that organisms and groups are *vehicles*, that there are *levels* of selection, that genes can *recognize* copies of themselves, that they *control* development and *instruct* the manufacture of bodies and that these qualities make them causal agents in a way that other things are not. He makes analogies to *rowing trials* and *Chicago gangsters* and he invokes *strategies, pay-offs* and other economic and game-theoretic concepts in order to explain how behaviour is *programmed*.

Throughout his lengthy and engaging argument Dawkins is quick to draw a line between his use of metaphor, model and analogy and the true, non-metaphorical features of the world. He frequently reminds the reader that brevity is the sole reason for talking about genes as if they were purposeful and about individual organisms as if they performed complicated cost/benefit calculations. Dawkins' pedagogical skill and rhetorical embellishments should not, if we are reading carefully, distract us from the truth: understanding the living world means coming to see evolution and adaptation the way Richard Dawkins does, from the point of view of the gene.

I will begin my examination of Dawkins' writing with a brief description of the background against which his view was formulated and then continue with a summary of Dawkins' view as it is developed in *The Selfish Gene* and *The Extended Replicator*. Dawkins has written hundreds of pages articulating, defending and supplementing his view in books and journal articles. My summary must necessarily be brief, but I will aim to give a fair account of what I take to be the essential features of his position. This summary will set the stage for the topic of the next chapter, the function of Dawkins' metaphors in his explanations.

5.2 Challenging The Received View of Evolution

In his introduction to the second edition of *The Selfish Gene* (1989), Richard Dawkins describes how the central message of this book has, in the intervening years since the book's original publication, become textbook orthodoxy, despite the fact that the argument for this message has come to be seen as extreme (1989, x). An early review of the first edition concludes with the observation that Dawkins' extreme gene-selectionist viewpoint is in stark contrast to the caution population geneticists display in their work. Geneticists, the reviewer says, acknowledge the interweaving of cultural and biological evolution, supporting the claim that when it first appeared the message of *The*

Selfish Gene was not the received view (Gibson 1977, 394). Situated within its historical context, it is easier to make sense of Dawkins' claim that his view is an orthodox one, for his gene's-eye view is in many respects a popularization of views that preceded and were contemporaneous with his writing of *The Selfish Gene*. Though Mary Midgley (1979) goes too far in claiming that the meme, the unit of cultural evolution, is Dawkins' only original idea (456), there is truth in the claim that a great many of his ideas had their original expression in biological writing other than Dawkins'. We should hardly be surprised that Dawkins drew on the work of others in gathering the many pieces that he fits together in writing such a wide-ranging scientific work. In fact, Dawkins singles out G.C. Williams, R.L. Trivers, W.D. Hamilton and J. Maynard Smith as biologists who had "invigorated" ethology by the infusion of their ideas (1989, viii).

But Dawkins' more recent claim that his view is orthodox is far from uncontentious. Indeed, Sterelny and Griffiths (1999) characterize the gene's eye view of evolution as one of several *challenges* to the received view. "The received view conceives of natural selection as the result of competition between individual organisms in a population. Differences among those organisms result in their differing success… ."(38) In contrast, gene-selectionists (those who adopt the gene as the only unit of selection) hold that genes are the only entity that are copied and copied closely enough to form lineages or "copy generations." In his textbook, *Evolutionary Biology,* Futuyma (1986) accepts natural selection as the central principle of evolutionary biology, but in his view the entities that are selected can be individual genes, large parts of the genome, individuals, populations or species, though most of the theory of natural selection centres on the differential survival and reproduction of individual organisms (151). Genic selection is but one aspect of a principle that is wide ranging, and it is generally restricted to cases in which some genes have a representation in the gametes significantly above fifty percent, as do meiotic drive genes, for example (153). It seems that Dawkins proclamation of the "textbook orthodoxy" of his view, though true in some minor ways, exaggerates the extent to which his view was adopted by the biological community.

One way in which Dawkins' view is orthodox is in its adaptationist stance. Gould and Lewontin (1993) claim that adaptationism, which can be traced back to the late nineteenth century to Alfred Wallace and August Weismann, has come to dominate evolutionary thought in England and the United States since the 1930s (581). Adaptationists assume as a general explanatory approach that most of an organism's traits contribute to the survival and reproduction of that organism and that this explains why that organism has that particular feature. In other words, most traits are adaptations. "This program regards natural selection as so powerful and the constraints upon it so few that direct production of adaptations through its operation becomes the primary cause of nearly all organic form, function and behavior." (584) This assumption was one of the main characteristics of biological explanation that I described in section 2.3. Though adaptationism has been criticized as a narrow minded methodology that fails to give consideration to explanations of traits in terms other than natural selection (Gould and Lewontin 1993, 581), Dawkins argues that anyone who objects to this approach really has in mind a caricature of adaptationism, what Stephen Jay Gould has termed "Panglossianism" - an exaggerated view that sees any trait as an optimal solution to an evolutionary problem (1993, 30). Dawkins points out many caveats to the adaptationist approach, the recognition of which make adaptationism sound like a highly successful but modest approach to developing explanations for traits. Sterelny and Griffiths also argue that a commitment to adaptationism is not necessarily conjoined with an "optimality hypothesis."(1999, 44) At least this one aspect of Dawkins' position has some agreement with the received view.

5.3 Historical Background

(i) William Hamilton

In his paper "The Genetical Evolution of Social Behaviour," William Hamilton (1964) developed a mathematical model to demonstrate that organisms will evolve behaviours to maximize what he termed their "inclusive fitness" (the total measure of all replica genes of an individual in its relatives) (1). Whereas the classical model of fitness asks how any given trait affects the fitness of the individual with that trait, Hamilton's model asks us to consider to what degree any given trait confers benefits on that individual's inclusive fitness (8). His model predicts that, with respect to relatives who are closely related to an individual, there will be positive selection for those behaviours that tend to maximize an individual's inclusive fitness, giving the appearance of altruistic behaviours.

> ...[F]or a gene to receive positive selection it is not necessarily enough that it should increase the fitness of its bearer above the average if this tends to be done at the heavy expense of related individuals, because relatives, on account of their common ancestry, tend to carry replicas of the same gene; and conversely that a gene may receive positive selection even though disadvantageous to its bearers if it causes them to confer sufficiently large advantages on relatives. (Hamilton 1964, 17)

From his consideration of the selection for genes that maximize inclusive fitness, Hamilton derives a general principle: "The social behaviour of a species evolves in such a way that in each distinct behaviour-evoking situation the individual will seem to value his neighbour's fitness against his own according to the coefficients of relationship appropriate to that situation." (19) Thus, individuals may display behaviours that appear altruistic if our attention is focused on the individual, but from the point of view of gene-copies, it is not so. The implications of Hamilton's model are extended by him to cover the possible evolution of recognition effects, mutual grooming, warning-calls, the degree of territoriality between distantly related individuals, and reproductive altruism in Hymenoptera (the insect group that includes ants, bees and wasps).

(ii) George C. Williams

G.C. Williams' *Adaptation and Natural Selection* (1966) is in many ways the model for Dawkins' work. Williams' main concern was to purge biology of the idea that group or biotic adaptations exist, arguing instead that adaptations exist for the benefit of the individual gene. This gene-selectionist view, Williams thought, would provide a more disciplined basis for the explanation of adaptations, a concern that also moved Dawkins to write *The Selfish Gene* and *The Extended Phenotype*. It is also a paean for natural selection as the predominant mechanism in evolution. Williams, unlike Dawkins, had a more tempered view of attributing adaptation status to behaviours and structures, though the recipient of the benefits of adaptations are not in doubt:

> Adaptation is a special and onerous concept that should be used only when it is really necessary. When it must be recognized, it should be attributed to no higher a level of organization that is demanded by the evidence. In explaining adaptation, one should assume the adequacy of the simplest form of natural selection, that of alternative alleles in Mendelian populations....(5)

All adaptations, Williams argued, are the result of natural selection not only between competing individual organisms, but among the genes of those organisms. "A gene is selected on one basis only, its average effectiveness in producing individuals able to maximize the gene's representation in future generations." (251) The selection of individuals is reduced to the selection of the genes that, in their expression, are responsible for producing the body that will help to propagate them. By giving answers to some of the most perplexing questions in evolutionary biology, questions such as "What structures and behaviours are adaptations?" and "For what are these adaptations to the benefit of?", Williams hoped to provide a framework for explanation that emphasized parsimony in appeal to levels of selection.

Williams argued that the conditions required for natural selection, primarily the stability and longevity of the unit of selection, discount selection at levels above the gene: Only the gene is stable enough to be effectively selected (109). "If there is an ultimate indivisible fragment it is, by definition, "the gene" that is treated in the abstract discussions of population genetics."(24) At the time of writing, Williams did not see the current theory of adaptation and natural selection as having been formulated rigorously enough to allow for the realization of its full implications and applications. It was part of Williams' project to help move things in this direction and, once this had been done, the theory might be able to generate more questions for empirical research. In generating and answering these questions, the details of the theory of natural selection might require revision, but the main structure erected by the mathematical work of Fisher, Hamilton and Haldane would remain.

> "Perhaps today's theory of natural selection, which is essentially that provided more than thirty years ago by Fisher, Haldane, Wright, is somewhat like Dalton's atomic theory. It may not, in any absolute or permanent sense, represent the truth, but I am convinced that it is the light and the way."(273)

Drawing on Hamilton's work, Williams argued that the formation of groups is the statistical summation of individually adaptive behaviours (215). Explaining group behaviour in a way that is consistent with the logic of natural selection requires one to shift one's attention from the group to the individual and to ask how that behaviour might benefit the individual. And, as noted above, Williams advocated shifting one's attention again to the gene in order to explain individual behaviour. The explanation for all adaptive behaviours, whether they are displayed by an individual or a group, are ultimately explained in terms of the propagation of genes.

(iii) Robert L. Trivers

In "The Evolution of Reciprocal Alruism," Robert Trivers (1971) expanded on Hamilton's model of the evolution of social behaviours between closely related individuals. His model showed that instances of altruistic behaviours between distantly related organisms (cases in which kin-selection can be ruled out) can also be explained by natural selection. Trivers' model suggested that we view these behaviours as benefitting the organism that acts altruistically, albeit in the long run. By recasting altruistic behaviours in terms of how they benefit the organism performing the altruistic act, Trivers' model was consistent with Hamilton's explanation. And like Hamilton's, Trivers' explanation of altruism was "designed to take the altruism out of altruism."(1971, 35) His model of "reciprocal altruism" explained three behaviours that are often held up as paradigm altruistic behaviours: cleaning symbioses among different species of fish, warning cries in birds, and human altruism.

Trivers calculated the costs and benefits of altruistic behaviours in terms of the chances of the relevant alleles propagating themselves in the population (36). For any population, altruist genes will come to predominate as long as the net benefit to the altruist exceeds that to the non-altruist, and altruistic behaviours are heritable traits. The benefit to the non-altruist must be kept small and this happens when the cost of altruistic acts is kept low. The cost of altruistic acts will be kept low if the altruist is "grudging," that is, if the altruist refuses any more altruistic acts to a non-altruist and restricts altruistic acts to those who are also altruists. Describing exchanges between individuals in terms of costs and benefits meant that Trivers' model, unlike Hamilton's, could be applied to different species.

Selection favours altruism when there are (1) many chances for altruism, (2) low dispersal rates in the population, so repeated interactions between the same individuals, and (3) individuals that are mutually dependent (36-7). Dispersal rate and degree of dependence are especially salient in determining the chances that any individual will interact with the same individual.

On Trivers' model, cleaning behaviours can be explained in terms of the benefit that a fish derives from being able to return easily and consistently to the same cleaner rather than from the perspective of the benefit to the cleaner (43). Triver's model also turned the explanation of bird warning calls on its head. Instead of asking how a warning call aids the group or postulating a complex system of functionality and non-functionality of calls in relation to breeding season, it asked how the warning call benefits the bird giving the call. "It does not matter that in giving a warning call the caller is helping its non-calling neighbours more than it is helping itself. What counts is that it outcompetes conspecifics from areas in which no one is giving warning calls." (45)

Trivers introduced the idea that a scenario where two individuals are in repeated symmetrical situations of conflict is analogous to the "Prisoner's Dilemma" and that reciprocal altruism can be reformulated in these terms (38). Dawkins makes extensive use of the iterated Prisoner's Dilemma in his discussion of the evolution of altruistic behaviours such as blood-sharing in vampire bats (though this is a theme that he develops in the second edition of *The Selfish Gene*). Trivers' discussion of the selection of capacities for cheating detection and the selection of new forms of cheating to counter these discriminatory capabilities (1971, 50-1), as well as his discussion of the selection of ever subtler mimicry and deception by the offspring (1974, 257), foreshadow Dawkins' elaboration of "arms races" in *The Extended Phenotype*.

In Trivers' 1971 paper we encounter the terms 'cheaters' being used to refer to individuals who refuse to reciprocate altruistic actions, a term that Mary Midgley (1978) roundly criticizes Dawkins for using, apparently unaware of this earlier use (440). Trivers' discussion of these behaviours, also introduced the concept of "time lag": the individual that first performs an altruistic act must wait a significant period of time before she is helped in return (38). Time lag is crucial for the evolution of reciprocally altruistic behaviours because it places constraints on the circumstances in which the altruist has some guarantee that the "help" she gives returns to benefit her in the form of "help" from another, and in which the "cheater" is selected against (39).

Trivers (1974) challenged the received view that parent-offspring conflict can be explained as the parents' attempt to reach an equilibrium between increasing chances of offspring survival and the ability to invest in other offspring. He argued that we ought to view offspring as agents actively working towards securing parental investment, thus making parent-offspring conflict an inevitable part of sexual reproduction as offspring attempt to elicit ever-greater amounts of parental investment in order to maximize their reproductive success from the earliest possible moment (249). Trivers not only initiated a shift in view (a precursor to Dawkins' admonition to consider whose gene's any particular behaviour is for), but also adopted economic language to explain the evolution of behaviours, for example, his use of cost/benefit ratios to predict behaviour - a conspicuous element of Dawkins' own writing. Further, Trivers introduced the idea that the parent-offspring conflict implies that offspring employ psychological weapons in their attempt to upset the investment strategies of their parents (257), an idea that Dawkins makes use of to great effect in *The Extended Phenotype.*

5.4 The Selfish Gene

Dawkins was writing to popularize the most current ideas in evolutionary biology, to draw together the many themes addressed by Hamilton, Trivers and Williams, and to make accessible responses to many of the most persistent challenges to orthodox Darwinism. Partly because he hoped to reach out to a wide audience (*The Selfish Gene* was written with three readers in mind: the layman, the expert and the student) and because the truth of his view was based on some technically sophisticated work in population genetics and game theory, Dawkins combined colourful turns of phrase and instructive metaphors with simplified models, analogies and thought experiments to make his ideas accessible. Through these devices Dawkins hoped to convince his readers of an astonishing truth: " We are survival machines – robot vehicles blindly programmed to preserve the selfish molecules known as genes."(1989, vii)

Ordinarily, we might interpret this as a dramatic metaphor for a more conservative claim, but Dawkins is sincere in claiming that this is no metaphor; his claim is about the nature of living things and our place among them.

5.5 The Problem of Altruism

A selfish individual is one whose welfare is increased at the expense of another. An altruistic individual is one whose behaviour reduces its own welfare while increasing that of another. Welfare in both cases is to be understood in terms of the number of offspring produced. Putative cases of selfishness are the cannibalism of the female praying mantis and the reluctance of emperor penguins to be first into the water. The female praying mantis bites off her mate's head before coitus is complete, thus obtaining a good meal while having her eggs fertilized (Dawkins 1989, 5). Emperor penguins are vulnerable to attack by seals in the water and so any individual would benefit by having another test the water for the presence of seals before they dive in (5). Examples of altruistic acts are the "kamikaze" behaviour of stinging honeybees, bird alarm calls and, the most obvious of all, parental care of offspring. For a biologist to explain how any one of these behaviours of an organism evolved, it is necessary to think of how that behaviour enhanced the reproductive success of the individual that displays that behaviour.

Altruistic behaviours pose a particular challenge to biologists because altruistic individuals seem purposely to reduce their reproductive fitness relative to other individuals who do not display these behaviours. The orthodox, individual view of selection has difficulty explaining the evolution of putative altruistic behaviours because individuals who do not bear the cost of altruistic behaviours prosper, allowing selfish behaviours to sweep through the population and eventually wiping out altruistic individuals. This difficulty motivates explanations of altruistic behaviours in terms of group selection; those groups that possess individuals that are altruistic prosper over those that do not.

Dawkins argues that attempts to explain altruistic behaviour by appealing to the good of the group not only get the facts of evolution just plain wrong, but result in absurdity. There is no logical reason that altruistic behaviour amongst species would not extend beyond the borders of the species. If group selection is true, then we should reasonably expect that antagonistic species, like lions and gazelles, ought to cooperate for the good of all mammals. Even more damaging to group selection theory is the fact that it is impossible to give an account of how such altruistic groups might actually evolve because they are always susceptible to "treachery from within" - altruistic groups can be exploited by individuals that adopt a more self-serving strategy.

In order to explain altruism, we must be able to settle on a unit of selection that receives the benefits of altruistic behaviour, allowing it to prosper relative to selfish individuals. Any serious science of adaptation must be able to isolate the unit of selection - the entity that any trait or behavior is for the benefit of - because attributing adaptations to different kinds of entities, for example, the gene or the individual or the group, will produce quite different explanations (1999, 81). Dawkins contends that the real explanation of altruistic behaviour lies not at the group level, but at the orthodox level of the individual. However, it is not the individual organism that is being ultimately selected; rather, the fundamental unit of selection is the gene.

5.6 Replicators and Vehicles

To understand why genes are the true units of selection, the individuals that benefit from adaptations, Dawkins (1989) begins with a speculative explanation of the origin of life. Darwin's dictum the "survival of the fittest" is actually a special case of a more general principle Dawkins calls the "survival of the stable" (12). A stable thing is a collection of atoms that is permanent enough or common enough that it deserves a name, for example, a mountain or rain drops.

In the chemical soup that constituted the early conditions on earth, a molecule was thrown together that had the unique property of acting as a template or mold for making more of itself. Dawkins calls this copying molecule a *replicator* (15). Replicators had three attributes that made them capable of taking part in the first process of natural selection: they had relative longevity, they varied in their rates of replication (fecundity) and they had a near-perfect rate of copying fidelity. This last attribute was crucial for evolution. There must have been the possibility of cumulative variation so that subtle variations in the behaviour of molecules allowed for cumulative complexity. The errors that accrued might have been the ability for one kind of molecule to dismantle other molecules and to make use of those liberated building blocks to make more copies of themselves. These replicating molecules might also have hit on the ability to construct protective coatings or shells. Dawkins suggests that this is how the first cells came about. The most important aspect of this story, however, is that these replicating molecules began to construct containers or vehicles for their continued existence. Dawkins terms these replicator containers "survival machines." He continues,

> Was there to be any end to the gradual improvement in the techniques and artifices used by the replicators to ensure their own improvement? …Now they swarm in huge colonies, safe inside gigantic lumbering robots, sealed off from the outside world, communicating with it by tortuous indirect routes, manipulating it by remote control. They are in you and in me; they created us, body and mind; and their preservation is the ultimate rationale for our existence. They have come a long way, those replicators. Now they go by the name of genes, and we are their survival machines. (20)

On Dawkins' view, we are not survival machines metaphorically, but literally: "A monkey is a machine that preserves genes up trees, a fish is a machine that preserves genes in water…."(21) Dawkins' description of replicators as the active agents in evolution, in this case our genes, suggests that as vehicles we are merely passive containers for our genes. Hull (1981) criticizes exactly this point in Dawkins' first formulation of the selfish gene concept. Dawkins description of the replicator, Hull says, runs together two powers: the power to reproduce one's structure and the power to do so differentially (33). Though Hull is making specific reference to a paper written by Dawkins two years after the publication of *The Selfish Gene,* Dawkins' original formulation of the replicator concept does run these two processes together. Replicators have the ability to make copies of themselves and they make copies of themselves by gathering the resources they need from their surroundings such that they outnumber other replicators; "There was a struggle for existence among replicator varieties." (1989, 19) And, in *The Extended Phenotype,* Dawkins seems to make a similar error; "Genes manipulate the world and shape it to assist their replication."(1989, 5) Hull contends that these two processes are distinct enough to warrant a terminological distinction. He suggests retaining "replication" for the process of copying and adding the term "interaction" to denote the causal contact that replicators must have with their environment if their distribution in later generations is to change. The entities that correspond to these two processes are replicators and interactors. Replicators pass on their structure through replication; interactors produce differential replication by interacting as cohesive wholes with their environment (Hull, 33).

The tendency for writers to fail to make this distinction, Hull believes, has led to needless confusion about the unit of selection. If this distinction is kept firmly in mind, one can see that though genes may be the only unit of replication, there may be many levels of interaction and thus more than one interactor. Can entities more inclusive than genes be replicators? According to Hull, whenever the transmission of structure exhibits a certain level of "directness" then the entity involved can be considered a replicator. So, in the case of similar, homozygous organisms that sexually reproduce, it is not the genome that is the replicator as Dawkins might argue, but the entire organism (34). The important point about Hull's argument, however, is that both replication and interaction must be present for evolution by natural selection to occur. The gene alone cannot be the unit of selection.

In response to Hull's criticisms, Dawkins (1989) revised his vocabulary and made a distinction between replicators and vehicles (formerly survival machines). The main difference between replicators and vehicles is that vehicles do not replicate themselves and function only to propagate replicators. In contrast, vehicles do something that replicators cannot do; they interact directly with the world through their behaviour (1989, 254).

Dawkins still argues that genes are the only entities that deserve the title of "replicator" and unit of selection. Though organisms and groups are of great functional importance they are "vehicles," entities in which replicators travel about and whose attributes are affected by the replicators inside it (112). Thus, Dawkins recognizes the existence of two different entities, but he does not acknowledge two *processes* required for evolution by natural selection.

Dawkins' solution to the puzzle of altruism begins with his assumption that one can make reasonable guesses about the attributes of our genes by knowing something about the environment in which they have succeeded. He explains this approach by means of an analogy:

> If we were told that a man had lived a long and prosperous life in the world of Chicago gangsters, we would be entitled to make some guesses as to the sort of man he was. We might expect that he would have qualities such as toughness, a quick trigger finger, and the ability to attract loyal friends.

> These would not be infallible deductions but you can make some inferences about a man's character if you know something about the conditions in which he has survived and prospered ... Like successful Chicago gangsters, our genes have survived, in some cases for millions of years, in a highly competitive world. This entitles us to expect certain qualities in our genes. I shall argue that a predominant quality to be expected in a successful gene is ruthless selfishness. (2)

If we combine the view that organisms are vehicles with the selfishness suggested by the Chicago gangster analogy, we should not be surprised if organismal behaviour is selfish. Genes are responsible for a body's traits that make it more likely to survive and reproduce than would genes for alternative traits. Genes are competing directly with their alleles for places on a chromosome. "Any gene that behaves in such a way as to increase its own survival chances in the gene pool at the expense of its alleles will, by definition, tend to survive." (36) If genes are selfish, Dawkins says, then we ought to expect them to program their vehicles to act selfishly also. "This gene selfishness will usually give rise to selfishness in individual behaviour."(1989, 2) It would seem, however, that selfishness of individual genes makes the development of a complex organism unlikely since each gene will be constructing its own machine for its own benefit. How would these selfish individual entities ever cooperate in the complex process of organismal development?

5.7 Genes, Programs and Development

DNA has two properties that Dawkins thinks are essential. The first is that it is a *replicator*. The second is that provides a *code* or *instructions* for the manufacture of proteins, the building blocks for bodies. Some small stretches of DNA code for particular protein chains and this fact has led some biologists to define a gene as that particular stretch that codes for an amino acid (a "cistron"). The possibility that cistrons might be altered by crossing-over, by point mutation, or by deletion during meiosis (cellular division that results in the formation of gametes) means that their longevity might be compromised. Dawkins wants a definition that maximizes the longevity of the genetic unit. As he notes, the longer the genetic unit the more likely it is to be subject to the vagaries of meiotic division, so he defines a gene as "... any portion of chromosomal material that potentially lasts for enough generations to serve as a unit of natural selection."(28) Individuals and groups fail to be units of selection because they are too temporary in terms of their stability across evolutionary time. Groups, individuals, and even chromosomes in sexually reproducing species are too fleeting. "Genetically speaking, individuals and groups are like clouds in the sky or dust-storms in the desert."(34) Only the gene is long lived enough to be the unit of natural selection, all other traditional units of selection are in reality vehicles for genes.

Genes have priority over other causal factors in biological explanation because genes control embryonic development in a way that other things do not. DNA supervises the manufacturing of proteins (23). DNA is instructions for making a body; they are like architect's plans (22). Even though genes have this enormous power, Dawkins admits that making a body is a complicated venture, one so complicated that the interaction of genes and their multiple effects makes it "...almost impossible to disentangle the contribution of one gene from that of another." (24) Now this complexity seems at odds with Dawkins' claim of the particulate nature of genes and their selfishness. Though development is controlled by an interlocking web of relationships so complex that "no one factor, genetic or environmental, can be considered as any single 'cause' of any part of a baby," Dawkins thinks that we can still talk sensibly about "genes for " different traits because there will be genes the presence or absence of which, other things being equal, will make a difference in development (37). Dawkins admits that the effect of a gene depends on its background, which is comprised of genetic and environmental factors, but the idea of a "gene for" a trait still makes perfect sense in this context. To show that talk of individual selfish genes makes sense in light of the complexity of development Dawkins resorts to an extended metaphor of selecting a crew for a rowing competition:

> Rowing the boat is a cooperative venture, but some men are nevertheless better at it than others... Every day he [the coach] puts together three new trial crews, by random shuffling of the candidates for each position, and he makes the three crews race against each other. After some weeks of this it will start to emerge that the winning boat often tends to contain the same individual men. These are marked up as good oarsmen. Other individuals seem consistently to be found in slower crews, and these are eventually rejected. ...The oarsmen are genes. The rivals for each seat in the boat are

> alleles potentially capable of occupying the same slot along the length of a chromosome. Rowing fast corresponds to building a body which is successful at surviving. The wind is the external environment. The pool of alternative candidates is the gene pool. As far as the survival of any one body is concerned, all its genes are in the same boat. Many a good gene gets into bad company, and finds itself sharing a body with a lethal gene, which kills the body off in childhood. Then the good gene is destroyed along with the rest. But this is only one body, and replicas of the same good gene live on in other bodies which lack the lethal gene. (38)

To be successful, a gene needs the right background and a gene that cooperates with most of the other genes that it meets in successive bodies is one that will have an advantage over others in making copies of itself. Selection works only on separate selfish genes, but those genes that are selected are those that do well in the presence of the right set of other genes. Granted, genes are responsible for important parts of protein synthesis required for making bodies, but in what sense should we consider behaviours to be "programmed"?

5.8 Genes, Programs and Behaviour

Dawkins explains that genes program bodies in much the same way that computer programmers can be said to program computers to play chess (for Dawkins, brains are analogous in function to computers and *both* are a kind of machine) (49).

Chess programmers cannot provide a list of all anticipated games, an unrealistically time-consuming, if not impossible, endeavour. Instead, they set everything up beforehand with a list of rules for moving pieces and "advice" on moves to attempt (52). The reason genes must set things up beforehand and then sit passively aside is because of the problem of "time lag". The control genes exert over protein synthesis is powerful, but slow. Behaviour, the body's response to environmental stimuli is very fast and must be so if the body is to survive. "Like the chess programmer, the genes have to 'instruct' their survival machines not in specifics, but in the general strategies and tricks of the living trade." (55) Thus, genes play a double role, they are not only the makers of the instructions; they are the instructions themselves.

Making instructions for behaviours that are capable of dealing with future contingencies is analogous to gambling, but gambling with one's own future welfare. Winning is being "correctly" programmed to deal successfully with one's environmental contingencies and reproducing successfully (56). Like their human counterparts, gamblers in the biological world must consider the stake, the odds and the prize.

The ability to learn is one capacity that helps to solve the problem of programming for unpredictable environments. Dawkins gives an example of the type of instruction that might underlie learning:

> Here is a list of things defined as rewarding: sweet taste in mouth, orgasm, mild temperature, smiling child. And here is a list of nasty things: various sorts of pain, nausea, empty stomach, screaming child. If you should happen to do something that is followed by one of these

nasty things, don't do it again, but on the other hand repeat anything that is followed by one of the nice things. (57)

Instructions are not given in this form, but in their biological equivalent – a set of instructions to produce certain proteins that either produce certain structures or that produce tendencies to act in particular ways. Essential to Dawkins' explanation is the ability to talk about behaviour being purposeful, whether literal or metaphorical, and he is quick to warn us that it is only "as if" animals were following instructions. ("As if" stands for some unspecified mechanism (50).)

A powerful way of thinking about the behavior of individuals whose genes have a conflict of interest, Dawkins says, is John Maynard Smith's concept of the evolutionary stable strategy (ESS) (69). This way of thinking can be used to try to explain threat and bluff displays in animals that, in the absence of a "good of the species" explanation, seem puzzling. Solutions to problems about the evolution of such behaviours can be found by thinking about the costs and benefits of adopting particular strategies:

> Suppose that B and C are both my rivals, and I happen to meet B. It might seem sensible for me as a selfish individual to try to kill him. But wait. C is also my rival, and C is also B's rival. By killing B, I am potentially doing a good turn to C by removing one of his rivals. I might have done better to let B live, because he might then have competed or fought with C, thereby benefiting me indirectly. (1989, 68)

A good strategy is one that provides the highest pay-off, where pay-off is understood in terms of numbers of copies (or offspring). Strategies that are good will increase in the population, while bad ones will decline. Thus, a good strategy will encounter more of itself and will become stable if the population is uninvadable by any alternative strategy and the strategy does well against copies of itself. Dawkins asks us to consider an imaginary population of two behaviours, hawks and doves. Hawks are programmed to attack whomever they meet, while doves always run away when confronted. Assigning points to the different outcomes for the three possible pairings of hawks and doves relative to the costs and benefits of each strategy (for example, if a hawk and dove meet the hawk wins 100 points and the dove runs away unharmed for 0 points), there is a stable population based on the average costs/benefits -- 5/12 of the population are doves and 7/12 of the population are hawks (1989, 70-1). At this ratio of doves to hawks no invading hawk or dove can improve on the pay-off achieved by any other member of the population. This ratio of behaviours is evolutionarily stable. If group selection were true, ESS thinking suggests that a "dove conspiracy" should evolve because the average pay-off in an all-dove population is greater than one in which 7/12 of the population are hawks. The conspiracy, however, is susceptible to "treachery from within." If greater than 5/12 of the population are doves, then hawks gain an advantage and will begin to proliferate. Humans can form conspiracies because they have conscious foresight, and can subject themselves to rules that, in the short term, are sub-optimal. This is one place where viewing genes as little reasoners might mislead us. "In wild animals, controlled by the struggling genes, it is even more difficult to see ways in which group benefit of conspiracy strategies could ever evolve. We must expect to find evolutionary stable strategies everywhere."(73)

Talk in terms of soliloquies also helps to explain territorial defense. If there is an asymmetry in nature where residents have a practical advantage over intruders, we should expect the behavioural rule "resident wins, intruder retreats" to be evident in behaviour. The opposite strategy is "paradoxical" in the sense that if intruders won, selection would work to make everyone an intruder and, eventually, there would be no residents because everyone would be running aimlessly around! On the first strategy, however, natural selection would favour individuals who strove to be residents, leading to behaviours like leaving home as little as possible, "defending" one's home, etc (79).

There are equivalent stable strategies to the above "pure" strategies that have mixed behaviours. In a strategy that uses mixed behaviours, individuals would sometimes act like doves and sometimes like hawks. Using the population from the previous example, individual behaviours would be biased toward a 7:5 ratio of hawk:dove behaviour (73). Dawkins asks us to consider increasingly complex behaviours, behaviours that depend upon the behaviour of one's opponent. These conditional behaviours more accurately portray behaviours found in nature.

5.9 Altruism and Cooperation: The Prisoner's Dilemma

Dawkins is now in a position to explain how and why colonies of essentially selfish genes and their vehicles form groups of seemingly cooperative and even altruistic individuals. His explanation combines Hamilton's theory of kin-selection, Triver's model of reciprocal altruism and, in the 1989 edition of *The Selfish Gene*, game-theoretic models of rational decision-making based on the Prisoner's Dilemma.

Copies of genes are distributed between bodies. If genes are trying to become more numerous in the gene pool, then we should predict that if they could somehow *recognize* copies of themselves in other bodies, they would assist them. The fact that close kin have a greater than average chance of sharing genes provides just such a way for genes to act as if they recognize each other (89). Related individuals can be assigned an index of relatedness, a measure of the average expectancy that two individuals might share a gene. For example, because brothers have each inherited half of their father's genes, there is a fifty percent chance that they share any one of their father's genes. The same is true of their mother's genes. Therefore, brothers have an index of relatedness of ½. First cousins have an index of relatedness of 1/8. By the time you get to third cousins, the index of relatedness is approximately the same as the expectation that any two random individuals in a population share a gene (92). If altruistic behaviour evolves, there is a good chance that, by residing in the body of the receiver of the altruistic behaviour, the altruistic gene is the beneficiary. Individuals receiving assistance might have increased fitness, and the gene for that altruistic behaviour would preponderate (93). No appeal to group selection is needed to explain altruistic behaviour between kin.

The evolution of associations of mutual benefit seem easy to explain when the benefit is simultaneous as in the case of cleaner fish, aphids and ants and lichen. In these scenarios, however, the choice of how to respond is immediate. But what happens when there is a time lag between receiving the benefits and having to respond? It seems that there is plenty of opportunity for the individual who receives the benefit first to "cheat" and refuse to return the favour. For example, mutual grooming initially seems difficult to explain because individuals who get the benefits of being groomed, but who bear none of the time or energy costs of grooming others, do better than those altruistic individuals that spend their time and energy grooming others. At no population ratio will groomers do better than cheats. Eventually the population will come to consist entirely of cheats and drive itself to extinction. However, if we introduce a strategy of grudging behaviour (individuals that have a memory and groom everyone except those whom they have groomed and then been cheated by), though a small population of grudgers would not do well, there is a critical proportion at which grudgers meet enough of themselves to begin to have a better pay-off than cheaters. Grudgers will then increase in the population and cheaters will decline. Once this critical proportion is reached, grudging becomes an ESS. Thus, a program of ruthless selfishness can result in a stable population of reciprocal altruists, individuals whose behaviour benefits others, but who also receive benefit from others.

Dawkins suggests that the relationships between individuals in situations like grooming - the receipt of benefits and costs, and the symmetry between to the two actors - is like a Prisoner's Dilemma. The Prisoner's Dilemma is a situation in which two people confront one another and must decide whether to cooperate or not in some venture (lying/telling the truth, checking for ticks on one another, etc). The possible decisions one might make come with "pay-off" values and "costs". Here is a modified version of Dawkins example (Figure 5.1):

		What you do: Cooperate	What you do: Defect
What I do	Cooperate	Get my ticks removed and bear cost of removing yours. **300**	Keep my ticks and bear cost of removing yours. **-100**
What I do	Defect	Get my ticks removed without bearing cost of removing yours. **500**	I keep my ticks but don't pay cost of removing yours, either. **-10**

The Prisoner's Dilemma

Fig. 5.1

In a Prisoner's Dilemma, one does not know the decision one's opponent will make. So, if one is attempting to maximize the possible pay-off one might get (a rational decision) then one will consider the possible decision each player might make and pick the best. For example, if I cooperate and you cooperate then we both receive 300 "points ." (In nature the points are probabilities of gene propagation.) If I cooperate and you do not, then I get –100 and you get 500. The logical outcome of such a scenario is that, regardless of what my opponent does, my best strategy is to always defect. Since both players in a Prisoner's Dilemma are rational and reason the same way, they both end up defecting and receive a score that is not only sub-optimal, but worse than if they had both cooperated, thus the dilemma.

The possibility of adopting new strategies emerges if the game is repeated or iterated. Dawkins says, "Successive rounds of the game give us the opportunity to build up trust or mistrust, to reciprocate or placate, forgive or revenge."(206) The difficulty, Dawkins says, is that it is by no means obvious which strategy will work out best in the long run.

In a series of publications in the early 1980's, Robert Axlerod and William Hamilton used the Prisoner's Dilemma in an extensive investigation of the problem of the evolution of cooperative behaviour. Axelrod ran several tournaments in which people submitted computer programs each one of which constituted a different strategy. He then played them off against one another over several hundred games assigning points for the outcomes of various interactions between strategies, and summed the points earned by each strategy. He discovered that over a large number of rounds, the best strategy was Tit-for-Tat. "This strategy is simply one of cooperating on the first move and then doing whatever the other player did on the preceding move."(Axelrod and Hamilton 1981, 1393) Tit-for-Tat, though not a true ESS (evolutionary stable strategy) because it can be "invaded" by other strategies that begin by cooperating, is collectively stable. That is, its resistance to invasion depends on what other strategies are present in its environment. Generally speaking, however, Tit-for-Tat is happy to cooperate and thus does well in the long run, even in a global population that might be predominantly selfish (224). In a logical extension of his earlier writing, Dawkins added a chapter based on this work to the second edition of *The Selfish Gene* to show how behaviours in the wild might be predicted by models based on the iterated Prisoner's

Dilemma. This model is so successful that Dawkins proclaims, "[M]any wild animals are engaged in ceaseless games of Prisoner's Dilemmas, played out in evolutionary time."(1989, 203) The application of the Prisoner's Dilemma to symbiotic relationships (where two different species cooperate, e.g., shrimp that clean the teeth and gills of fish or ants "milking" aphids) allows us to extend the thinking of ESSs. The Prisoner's Dilemma gives us a powerful method to predict and explain altruistic behaviours between parasites and their hosts, like bacteria, wasps and fig trees, and the alternation of sexes in sea bass (229).

The gene's eye view also provides an explanation for other putatively altruistic behaviours such as unequal investment by parents in rearing offspring. In many cases it is the female who invests the most time and energy rearing offspring. Genetically, parents have an equal investment in offspring (each investing 50%). But, if one can exploit the other, it leaves the "cheating" partner free to propagate more of his genes (140). Generally speaking, eggs are much larger than sperm and constitute a much greater investment. By comparison males contribute much less. Female exploitation, Dawkins claims, begins here (142). Though the initial difference between gametes may have been small, natural selection would work so as to increase that difference. The initial asymmetry in investment gives us a reason to expect that females will invest more throughout development (146).

The idea that behaviours are programs, combined with the assumption that adaptive traits are to be explained in terms of how they benefit the replication of the genes for those traits, provides Dawkins with a device with which to explain not only altruistic behaviour, but also the existence of sterile castes in insect colonies, herding, mating preferences and various animal artifacts - long-standing challenges to the individual view of evolution

5.10 Explaining Individual Organisms

Dawkins recognizes that there exists a tension between his view of the selfish gene as an agent and individual organisms as agents. A large part of Dawkins' argument for the selfish gene is that it makes sense to think of individual organisms striving to maximize their reproductive success and acting "as if" they were making complicated cost/benefit calculations. The impulse for organisms to "care" about their reproductive success rather than their own longevity, however, does not make sense at the level of the individual organism; it only makes sense from the gene's-eye view. Generally, this tension does not make itself felt because both body and genes are successful – they share the same fate. This symmetry allowed orthodox Darwinians to focus on the success of bodies, that is, individual organisms. However, when the good of the gene and the good of the body come apart the tension is felt. For example, meiotic drive genes distort the segregation of genes at meiosis to favour them being included in a disproportionate number of gametes (well over 50%) and, because the meiotic drive gene works at the genic level, natural selection will favour it even though it might have deleterious effects on the individual organism. For example, the t gene in mice is lethal when it is homozygous, causing death or sterility to any individual unlucky enough to have it.

Despite its lethal effects, it spreads quickly throughout a population driving the population to extinction (1999, 236).

This tension produced by the logic of the gene's-eye view forces Dawkins to explain why individual organisms exist at all. As Dawkins explains near the end of *The Selfish Gene*,

> The individual organism is something whose existence most biologists take for granted, probably because its parts do pull together in such a united and integrated way. Questions about life are conventionally questions about organisms. ...[Biologists] don't ask – though they should– why living matter groups itself into organisms in the first place. Why isn't the sea still a primordial battleground of free and independent replicators? Why did the ancient replicators club together to make, and reside in, lumbering robots, and why are those robots – individual bodies, you and me – so large and so complicated? (237)

Dawkins' answer is that the evolution of complex adaptations requires cyclically repeating developmental processes. This developmental recycling offers a return to the drawing board, a necessary step for change to be cumulative and adaptive. "Replicators that gang-up in multi-cellular organisms achieve a regularly recycling life history, and complex adaptations to aid their preservation, as they progress through evolutionary time."(1999, 259) An individual is the physical unit associated with one single lifecycle. For sexually reproducing individuals this lifecycle begins as a single-celled zygote. Dawkins concludes, "The integrated multicellular organism is a phenomenon which has emerged as a result of natural selection of primitively independent selfish replicators." (264)

5.11 The Extension of the Phenotype

"Phenotype" refers to the bodily manifestation of the effects of genes. Dawkins argues that, regardless of the length of the causal pathway from gene to effect, we can always intelligibly talk about a gene for a trait. If we can reasonably talk about a gene that produces a protein that colours skin black, we can equally reasonably talk about a "gene for black skin" despite the fact that the protein coded for by the gene acts as an enzyme only one of whose effects is to produce the black pigment (1999, 196). Though the causal chain is longer, only by arbitrarily limiting the phenotype to the boundary of the skin can we prevent the logic of genes-for talk from being extended to behaviours and their products. Thus, a stream-dwelling insect that makes a home out of pebbles from the stream bed, should have the house included as a product of its behaviour. If we could experimentally determine genes for choosing light stones over dark stones then "the rules of existing terminology imply that the artefact itself should be treated as part of the phenotypic expression of genes in the animal." (Dawkins 1999, 198) Following this logic, we ought to view spider webs as part of a spider's phenotype, and we ought to be able to talk intelligibly about "genes for building webs".

On Dawkins' view, there is no difference in principle between genetic control of morphology and genetic control of behaviour. Animal artifacts should be regarded as phenotypic tools by which genes lever themselves into the next generation (199). Parity of reasoning suggests that the phenotype can be extended from artifacts in close proximity to the organism to the chemical control of host organisms by parasites and even further to the near-hypnotic power of a cuckoo baby to fool a much smaller adult. The logic of the selfish gene leads to the conclusion that it does not matter in whose body the gene resides "An animal's behaviour tends to maximize the survival of the genes 'for' that behaviour, whether or not those genes happen to be in the body of the particular animal performing it." (1989, 253) So, when an organism's behaviour is puzzling from the perspective of the individual, the most likely explanation is that either, like Hymenoptera, reproduction does not involve the chromosomes from both parents, or, like the cuckoo's adoptive parent, the behaviour is to the benefit of a gene that resides somewhere else, i.e., in another organism.

5.13 For Brevity Only – Dawkins' View of Metaphor

My summary of Dawkins' argument has necessarily been brief, but I think that I have captured the essential points of his view. The more abstract reformulation of "survival of the fittest" gives Dawkins a way to link together the qualities that must be possessed by any unit of natural selection (replication, longevity and fidelity) with views of the origins of life. The process of natural selection dictates that any successful unit will be selfish in Dawkins sense - it will do all it can to increase its chances for successful propagation. By definition, then, altruism is bad. What appear to be altruistic behaviours are really programs or strategies that have been shaped over the long course of evolutionary time to benefit the genes that are responsible for them. Thus, the only altruism that exists is a tit-for-tat kind of altruism, reciprocal altruism. The gene's-eye view stays true to the orthodox Darwinian insight that only individuals can be units of selection, but modifies this view by suggesting that individual organisms are too temporary to be units of selection, and replacing the individual organism with the individual gene. The replacement allows for a unified and consistent explanation of traits that seemed to challenge the logic of orthodox Darwinism.

In the following chapter, I will take up a more sustained examination of the positive and negative aspects of Dawkins' metaphors. Before doing so, I want to briefly outline what I take to be Dawkins' view of the nature and role of metaphor in his writing. What Dawkins thinks about his metaphors is not of great consequence for my argument, but given his adroit use of metaphor and thought experiment, I think that it is revealing to see how conservative Dawkins' view actually is regarding the use and function of metaphor in scientific writing. Such a view is by no means uncommon among scientists.

Dawkins' metaphor for metaphor is the Necker cube, a drawing of a cube that can be seen at one time as from above and at another as from below. The Necker cube belongs to the family of perceptual puzzles that includes duck/rabbit, a puzzle well known to philosophers. Dawkins describes the shift from the individual view of evolution to his gene's-eye view as equivalent to the shift in perspective one achieves when looking at a Necker cube. To see the individual or the gene as the unit of natural selection is to see two faces of the same cube. Thus, our new perspective is a way of reorganizing what we know about evolutionary biology through the lens of the replicator. Through this flip in perspective, Dawkins hopes to persuade the reader of his thesis that adaptations are best seen as for the benefit of the active, germ-line replicator. In a single paragraph, Dawkins states much of what he believes about metaphor:

> Expounding ideas that have hitherto appeared only in the technical literature is a difficult art. It requires insightful new twists of language and revealing metaphors. If you push novelty of language and metaphor far enough you can end up with a new way of seeing. And a new way of seeing…can in its own right make an original contribution to science. (1989, xi)

Superficially, it sounds as though Dawkins has given metaphor a leading role in understanding the world. A closer look, however, and we see that what he has given with one hand, he has taken away with the other. Yes, metaphor is capable of providing a new way of seeing, but metaphor that has this function is restricted to novel uses of language. Presumably, once the metaphor is explored and the novelty is gone, it is replaced by a literal description.

Dawkins claims that metaphor is capable of laying bare unimagined facts (xi). However, he offers no hint of how it is that metaphor might have this function and, further, he does not seem to think that metaphor is cognitively important. Metaphor has, at best, a heuristic function, allowing us to think of the same facts in new ways, and allowing us to make new connections. This is indeed an important function, and part of the need for metaphor arises from the attempt to recast technical ideas and language in non-technical language. But there is no mention by Dawkins of whether or not the technical descriptions don't also suggest a different way of organizing the facts at hand or whether their formulation somehow prevented or obscured seeing them in a new way. For Dawkins, metaphors, whether talking about conscious genes or calculating individuals, and other as-if talk, are just descriptions for the sake of brevity. They are always translatable into the non-metaphorical language of gene frequencies and the mechanisms of natural selection.

The explanation of altruistic and selfish behaviours that Dawkins offers is engaging and wide-ranging. He has sometimes declared his descriptions literal, sometimes declared them metaphorical. Dawkins' assertions aside, is it possible to determine which of his descriptions are in fact metaphorical and which literal, which ones are necessary and which ones are eliminable? In the next chapter, I will argue that facing up to more recent challenges in evolutionary biology raises questions about the literal status of some of Dawkins' explanations and turns us squarely toward the central question of this thesis: What is the nature and function of metaphor in biological explanation?

Chapter Six

Replicators, Rulers and Reasoners

6.1 Multiple Metaphors

There are three metaphors that play a prominent role in *The Selfish Gene* and *The Extended Phenotype* and that shape the way in which we conceive of genes. These metaphors are the Replicators, the Rulers and the Reasoners.

The "Replicator" metaphor describes the genes in terms of something that is copied and that plays a role in a particular copying process. As we saw in the last chapter, the genetic replicator has unique properties that Dawkins thinks ought to privilege it in heredity - its copying ability and its copying fidelity. Because of the gene's unique role in transmitting biological traits from parent to offspring by virtue of the part is plays in protein synthesis, Dawkins declares it to be the fundamental unit of heredity. But the criterion upon which traits are determined to be fundamentally genetic is developmental: Whether traits are passed on or not is something determined by watching the development of the offspring to see if they manifest the traits of interest. Making assumptions about how development proceeds, and what development requires, influences what is seen as relevant to the inheritance of particular traits.

The solution to the problem of altruistic behaviours that Dawkins is seeking in *The Selfish Gene* lies in finding the entity that ultimately benefits from altruistic behaviour by allowing it to reproduce more prolifically than those that do not. The ability of genes to be copied, their relative fidelity, and their fecundity make them ideal units of selection. Individuals, groups and larger entities do not have these characteristics to the required degree to allow them to be subject to natural selection, so Dawkins proposes that "... the basic unit of natural selection is ... some small unit of genetic material which it is convenient to label the gene."(Dawkins 1989, 39) These same abilities make the gene the beneficiary of the long-term process of the selection of adaptations. Asking what the units of selection are, however, tends to obscure the fact that the question, "What is the unit of selection?" has four different formulations, each one calling for a different kind of answer (Lloyd 1992, 334). Limiting the definition of the unit of selection to replicators and beneficiaries not only ignores these other questions, but also ignores the question of whether or not the gene is the only entity that can plausibly be said to play the roles Dawkins ascribes to it. The main theme of the first two sections, then, is to determine just how the metaphor of replication (a copying process) and replicators (things that are copied) can be reasonably applied and limited to the gene for the purposes of heredity and selection.

The "Ruler" metaphor describes genes in terms of giving instructions to other things and as possessing information that other things do not. Dawkins frequently speaks of genes being sets of instructions for how to make a body (22), and as setting up programs (62) for how bodies behave. Genes are also the makers of these instructions.

By use of the ruler metaphor, Dawkins assigns the gene a privileged role in determining developmental outcomes and gives it agency: The gene is the initiator of a causal biochemical pathway that ultimately leads to the effects that help its propagation (Lloyd 1992, 338). Do these metaphors automatically discount other entities, like cells, from being considered as units of development? Do genes play a role significantly different in development from hormones, habitat and nutrition? Describing genes in computational and informational terms not only positions the gene as the ruler of development, it serves to reinforce the gene's privileged position as the unit of selection and heredity.

Cognitive metaphors cast the genes and other entities as "Reasoners," the third metaphor referred to in my title. Dawkins frequently appeals to language that portrays genes as possessing human cognitive capacities: "Genes are *trying* to become more numerous in the gene pool by *programming* bodies to survive and reproduce"(Dawkins 1989, 88 my emphasis); they *manipulate* bodies for their own preservation (172); they are unconscious *strategists* (228); and, they *recognize* copies of themselves (88). These cognitive metaphors are not limited to the genes; Dawkins also attributes these states to non-human animals when, for example, he talks about them entertaining complicated cost/benefit calculations in the form of soliloquies: "I'll fight him for the harem in the end, but I may have a better chance of winning eventually if I wait, rather than rush in now." (68) Admittedly, Dawkins is constantly reminding us that talking about genes as if they had conscious aims is only a convenient way of talking (for example, 1989, 88-9). At the same time, however, these attributions allow us to ask questions that are unavailable to us if the same phenomenon is described in more "respectable terms."

What is of great interest is that this "as-if" talk is not only illuminating, but also gives us an effective method for predicting various behaviours and for generating testable hypotheses. The application of game theoretic concepts is permitted by the metaphorical redescription of genes and organisms as cognitive agents. Attempting to explain why these metaphors are so effective is a task I leave to the next chapter. I will restrict my discussion here to the ways in which these metaphors function in Dawkins' explanations.

6.2 Replicators: The Units of Heredity

All offspring share common sets of traits or observable characteristics with their parents. It is generally accepted that some traits, such as the language one speaks or the clothes one wears, are transmitted culturally. Other traits, such as hair colour, susceptibility to disease, etc., are the result of biological inheritance. Still others, such as one's height or weight, are the result of cultural and biological inheritance. Biological traits are determined by the elements of heredity that are transmitted to the offspring from the parent in reproduction (Hartl and Jones 1999, 2). The elements, or units of heredity are the genes, segments of DNA. Dawkins' gene has a privileged position in heredity because of the role it plays in the reliable replication of trait differences across generations. However, *replication* may not be the most fruitful metaphor through which to view either the hereditary process or its unit.

Dawkins originally defined a replicator as something that creates copies of itself and that acts as a mould or template for the production of a copy of itself (1989, 15). Presumably because of criticisms that this definition fails to recognize the complexity of the resources required for DNA to be replicated, he shifted his definition to "anything in the universe of which copies are made"(1999, 83). Refining his definition still further, Dawkins (1999) makes a distinction between two distinct categories of replicator. In the first category are active replicators that exert an influence over their probability of being copied and passive replicators that do not exert such an influence. In the second category are germ-line replicators that might conceivably have an indefinitely long line of descendants and dead-end replicators that will have only a finite number of descendants. Genetic replicators, genes, are active germ-line replicators – fragments of DNA that are long lived through copies and that exert an influence on the chance that they might be copied (Dawkins 1999, 84).

Dawkins likens the replicator's capacity to be copied to a xeroxed piece of paper (1999, 83). Like a xeroxed paper with an ink smudge, the genetic replicator will be copied "warts and all," as Dawkins puts it, a condition that, in the biological world, is uniquely satisfied by DNA. Changes in bodies, for example, the loss of an appendage or sun-bleached hair, or any other changes at a level higher than genes are not transmitted across generations. The genetic replicator is thus *the* unit of heredity.

The assumption that genes, on their own, can be considered replicators seems misleading because the ability for DNA to replicate is a property not of the gene, but of the cellular system. The initiation of the processes of replication, namely, mitosis and meiosis and the "proofreading" of the resulting copy, require enzymes and proteins that are in the cellular environment, not in the DNA itself. Dawkins view of the gene as the primary unit of heredity, based as it is on replication of DNA, is just too simplistic to capture the complexity of our developing understanding of inheritance and development.

Replication is used to describe the copying of DNA and it is copied in meiosis with a surprising degree of fidelity. But there are difficulties in construing the gene as a discrete unit. The material gene, Fogle argues, is, at best only an instrumental construct because there is "no feature in common among described protein coding genes of eukaryotes that unites them as material units of structure or function." (Fogle 1990, 366). In attempting to restrict the gene to a particular locus, the gene concept seems to mask a host of diverse structures and functions.

> [S]tructural regions can be shared, overlapping on opposite or like strands, nested, and even physically split. Introns [sequences of DNA that do not code for proteins] may or may not be present, the same nucleotide sequence can be both coding and non-coding, alternative splicing can produce multiple products of translation, proteins from one transcript can have multiple functions as either an intact structure or be cleaved into separate polypeptides, frame shifting can create a second translatable product from one nucleotide sequence, information crucial to gene expression can reside externally or internally with respect to the transcribed message… To place

all of these under the rubric of being a 'unit of inheritance,' in some material sense, strains credulity. (Fogel, 367)

The term "unit" was only appropriate for pre-Mendelian genetics and its use cannot be made consonant with current knowledge of DNA structure and function. One critic of Dawkins' discrete gene has argued that it might be more appropriate to describe genes as sets of domains - sequences of nucleotides that are identified by structural properties or activities and that distinguish them from other sequences (Fogle 1990, 367). Domains may be nested and overlap and many domains may comprise what we think of as a single gene; and domains may lack any empirical relationship to a single unit (369).

Dawkins explicitly limits his "genes for" talk to differences between phenotypes that can be traced to differences in genotype. But this talk has implicit in it the assumption that the measure of the difference and resemblance between parent and offspring - the degree of heritability - can accurately distinguish between the contribution of the DNA to the expression of a trait and the contribution of the environment to the expression of this same trait. It is true that geneticists have had great success predicting the extent to which some traits are inherited, especially those traits with a relatively simple genotype-phenotype correspondence. But Dawkins' hope to explain behaviours, in a way that is effective for more simple traits such as eye colour or sickle-cell anemia, seems quixotic. Behaviours are multi-factorial traits, the result of multiple genetic and environmental factors, and so no difference in phenotype will correspond in any direct way with a difference in a single gene. "With a multifactorial trait, a single genotype can have many possible phenotypes (depending on the environment), and a single phenotype can include many possible genotypes." (Hartl and Jones, 494) Further, our increasing knowledge of the variety of epigenetic effects suggests that "gene for" explanations will be limited. Cells acquire information that they can pass on to their progeny that is not contained in DNA. For example, cell 'memory systems' can transmit patterns of gene activity from parent to daughter cell with the consequence that cells with identical DNA can differ with respect to the genes that are active or inactive (Jablonka and Lamb forthcoming, Chapter 4). Some cells also possess 'architectural memories' and transmit changes in cell phenotype from one generation to the next, as in the case of paramecium that have had a section of their cortex altered and transmit this change across generations. There are also non-DNA features of chromosomes that are transmitted from generation to generation. These 'chromatin marks', such as DNA methylation, effect states of gene activity and inactivity and the mutation rates of those stretches of DNA in which they occur (Hartl and Jones 1998, 243).

On one hand, however, Dawkins' description of a "gene for" a trait, even in complex traits, is simply an extension of our everyday notion of a cause as "one condition selected from a complex set of conditions which…are together sufficient to produce the consequence." (Cranor 1994, 127) Of course our interests and context may help to determine what we select from many competing things as the contingent factor, but in this case we know the context and our interests. Drawing attention to a gene as *the* cause or *a* cause of a trait is a reasonable use of this ordinary notion of causation. We can recognize the complexity of inheritance and development and still draw attention to one contingent factor for purposes of understanding heredity.

> "The fact that a complex set of conditions is sufficient to produce an event does not detract from drawing attention to one of the contingencies as a cause or the cause for certain purposes. What matters is the context and the purpose and that we do not lose sight of the complexity of the processes involved." (Cranor, 131)

However, fixing on a gene as the causal factor for a trait may be the result of fixing attention on what is most obvious in making a causal contribution and not what is making the greatest causal contribution (133).

Despite the fact that this kind of talk can be understood as a relatively straightforward use of our notion of causation, there is the danger that focusing attention on one cause could unduly influence the formation of hypotheses, experiments and the determination of relevant data considered. So, even careful use of "gene for" talk drives research in particular directions and closes off alternate paths of inquiry (Cranor 1994, 132).

Focusing on the gene-as-replicator downplays the importance of the results of knockout experiments that have shown that genes known to be particularly important in developmental pathways could be eliminated with no effect on the phenotype. The genome was somehow able to compensate for the loss of these particular genes through the existence of multiple copies of those genes that were knocked out, or through multiple effects of other genes or through the dynamic regulatory structure of the developmental network. This suggests that gene replication may not be necessary for trait replication (Jablonka and Lamb, Chapter 2).

The replicator metaphor downplays the causal influence of other systems involved in the inheritance and expression of traits. Structural inheritance, chromatin-marking systems, and RNA silencing are not just transmitted across cell lineages, they are transmitted across generations of organisms and by doing so play a role in adaptive evolution (Jablonka and Lamb, Chapter 2). The importance of any gene in the causal pathway from genotype to phenotype cannot be determined on the basis of its presence or absence alone. The gene cannot be considered a single causal agent. "The idea that there is a gene *for* adventurousness, heart disease, obesity, religiosity, homosexuality, shyness, stupidity or any other aspect of mind or body has no place on the platform of genetic discourse."(Jablonka and Lamb, Chapter 1)

The complexity of inheritance and developmental pathways suggests that the identification of a gene, the replication of which serves to perpetuate traits or patterns of inheritance, misrepresents the processes involved. "Genes interact within and between loci in the development of most phenotypes, and the existence of these interactions prohibits a simple interpretation of the observed additive genetic variance in terms of additive gene action."(Wade 1992, 150) Though a quantitative measure of genetic variance between relatives may help to determine the importance of genetic factors, a recent textbook in genetics cautions the reader that genetic interpretation of familial resemblance is not always straightforward, especially in the study of complex behavioural traits in human beings, because of the possibility of non-genetic, but nevertheless familial, sources of resemblance (Hartl and Jones 1999, 510).

An alternative view to gene-selectionism, Developmental Systems Theory (DST), claims that the causal responsibility for the formation of traits cannot be partitioned into additive components of genetic and non-genetic origin. Instead, "DST views both development and evolution through the metaphor of processes of *construction* and *reconstruction* in which heterogeneous resources are contingently but more or less reliably reassembled for each life cycle." (Oyama, Griffiths and Gray 2001, 1) If it seems plausible that organisms are reconstructed in this way, then there is little reason to construe the gene as *the* unit, or even the most important unit of heredity.

Claiming the gene as the primary unit of heredity is a result of an unwarranted emphasis on the notion that structural identity must be transmitted across generations. This emphasis is a result of selecting the replicator metaphor and with it implications about the source of developmental activity that discount earlier tendencies to conceive of inherited components in a wider sense (Keller 2000b).

In a recent work, Eva Jablonka and Marion J.Lamb (forthcoming) argue that contemporary work in genetics strongly suggests that genetic inheritance is best conceived as only one of four kinds of inheritance. Information that will affect the development of the phenotype is passed along not only through the genetic system of inheritance, but also through epigenetic inheritance systems, through behaviour (e.g., learning and imitation) and, in the case of humans, through symbol systems.

The metaphor of replication captures something crucial to the process of evolution. But in his emphasis on DNA as a replicator, Dawkins has oversimplified the complexity of facts about the inheritance of traits. Even though geneticists can frequently make educated assessments of the degree to which complex or multifactorial traits are genetic or environmental, for example, by path analysis or variance analysis (Feldman 1992, 153), I think that the reasons above are persuasive enough to cause us to question the asymmetry that singles out the gene as a privileged unit of heredity.

Dawkins does have a reply to critics who want to enlarge the group of entities that ought to be included as units of inheritance:

> When we are talking about *development* it is appropriate to emphasize non-genetic as well as genetic factors. But when we are talking about units of selection a different emphasis is called for, an emphasis on the properties of replicators. The special status of genetic factors rather than non-genetic factors is deserved for one reason only: genetic factors replicate themselves, blemishes and all, but non-genetic factors do not (Dawkins 1999, 99)

Thus, Dawkins ties the replicator as a unit of heredity, as *the* transmitter of information about traits to another role – the unit of selection.

6.3 Replicators: Units of Selection

A textbook on evolutionary biology defines the necessary attributes of a unit of selection: "Selection can operate whenever different kinds of self-reproducing entities that beget descendants like themselves differ in their rate of survival or reproduction." (Futuyma, 1986, 150) Heritability, the degree of resemblance between offspring and parents, is fundamentally a product of the genes. Unless offspring resemble their parents to some degree, the effects of selection in one generation (the parents) will not be transmitted to the subsequent generation (the offspring) making heredity necessary for evolution by natural selection (Wade 1992, 149). The same factor that makes a replicator the unit of heredity, its ability to survive for a long time in the form of copies, is the same characteristic that Dawkins thinks allows it to satisfy the definition of the unit of selection given above, and that discounts other entities such as individuals, groups or species from being included as units of selection.

As we saw in Chapter Five, David Hull (1981) challenged Dawkins' description of the necessary and sufficient characteristics of the unit of selection. Dawkins' description of replicators, Hull argued, conflated two distinct processes, namely, replication and interaction. Replicators are differentially perpetuated because of the relative success of the interactors acting directly on the environment. Hull concluded that the process of selection requires the presence of both processes and their respective entities. Adopting this language expands the identification of relevant units of selection because the biological categories that satisfy the criteria for interactor and replicator will wander. "In certain groups nothing more inclusive than single genes may function as replicators, in others entire genomes, in others possibly organisms. In certain groups, one and the same entity may function as a replicator and as an interactor."(41) Hull's distinction is an important contribution to the debate about the role of metaphor in explanation because it emphasizes the fact that genuine biological explanation must refer to causal actors, and some metaphors might help us to identify them (interactors), while others may not (replicators). However, at whatever level they are identified, replicators and interactors must be individuals if they are to perform their function as units of selection because selection acts on individuals.

Recall that Dawkins made a conciliatory gesture to this criticism by making a distinction between the processes of replicator survival and vehicle selection. However, I suggested that Dawkins continued to hold that genes replicate *and* interact with the environment, albeit at a distance. Elisabeth Lloyd (1992) agrees with this assessment (335).

It is not only questionable whether genes can be causal actors in the process of selection, but also whether genetic replicators can be reasonably construed as the sort of individual that takes part in the process of selection. Darwinian individuals are discrete, functionally cohesive, have a birth and death and are stable across time, characteristics that make them readily identifiable as vernacular individuals (Gould 2002, 602-603). As well, Darwinian individuals must satisfy the necessary and sufficient characteristics for being involved in the process of evolution by natural selection: they bear offspring; offspring must be more like their parents than they are like anything else; and these individuals vary in their interaction with their environments in ways that effect their reproductive success relative to others as a causal result of the heritable properties not manifested by less successful individuals (611). Genes may satisfy the first three of these conditions of Darwinian individuality, but they fail to satisfy the directness of interaction with the environment and, for this reason, should not be considered as units of selection.

Focussing on faithful replication makes it easy to overlook the fact that the real goal of a unit of selection is to increase the representation of its heritable attributes (Gould, 621). Passing on these attributes can be done whether that unit is passed on as a whole, as in the case of asexually reproducing organisms or in disaggregated form as sexually reproducing organisms do (Gould, 621).

A better metaphor for revealing the required characteristics of the unit of selection, Stephen Jay Gould suggests, is the metaphor of the sieve.

> Sieving represents the causal act of selection – the interaction of the environment (shaking the sieve) with varying individuals of a population (particles on the sieve). As a result of this interaction, some individuals live (remain on the sieve, while others die (pass through the sieve) – and survival depends causally upon variation in emergent properties of the particles….(621)

A unit of selection is any entity that lies on the webbing of the sieve, and, by helping to make more entities with its characteristics, increases the number of entities like it relative to others. Any entity that is causally responsible for enabling a relative increase of its hereditary contribution to the next generation qualifies as a unit of selection (622). Selection occurs, then, when plurifaction - Gould's term for making more of things with your attributes - is the result of a direct causal interaction between traits of an evolutionary individual and the environment in a way that enhances the differential success of the individual. Genetic replicators do not have such traits, but individuals do, and so do higher-level entities such as species.

Like Hull, Gould concludes that direct interaction with the environment is the proper criterion for determining the unit of selection. Even if genes are the primary or even sole unit of replication, interaction may take place at many levels from the gene to organism, and from demes to species. The causal nature of selection is only recognized when selection is talked about in terms of those entities that interact directly with their environment and make more of their contribution to the next generation than others (622). Genes could only interact directly if they built organisms additively. If they did that, then the interactions of organisms with their environment could be causally reduced to the properties of individual genes (620). Traits are not the additive properties of genes, but the result of complex interactions. They are emergent properties. Dawkins, of course, recognizes that genes do not act directly with the environment in the way that bodies do. But because genes are at the source of the biochemical pathway that initiates the development of traits that do act directly, he thinks that the indirect action of genes is action enough to secure genes the status of the unit of selection: "Replicators don't behave, don't perceive the world, don't catch prey or run away; they make vehicles that do all those things."(Dawkins 1989, 254)

Though replicators are important in evolution, we could view them instead as items for tracking the selection history of particular traits rather than as causal actors in the process of natural selection (Gould 2002, 633). Genes do not form the "concrete reality" that confronts biologists. This role is confined to higher-level phenomena like individual organisms. The survival of replicators it is argued is only an outcome of the causal process of selection on vehicles. The replicators themselves do not take part in this process because they are invisible to selection, effectively hidden away in their vehicles. While changes at the genotypic level do not always manifest themselves at the phenotypic level, changes at the phenotypic level are always represented at the level of the gene. This aspect of the relation between genotype and phenotype, what Gould refers to as "downward causation," recommends genes as the ideal unit for evolutionary accounting. "No single unit of bookkeeping can monitor every conceivable change, but the gene becomes our unit of choice because the nature of hierarchies dictates that genes inevitably provide the most comprehensive record of changes at all levels." (Gould 2002, 636)

Dawkins attempted to dissolve the debate over the levels of selection by making a distinction between two ways of looking at selection. One way is to view it in terms of replicator survival. On Dawkins' view, there is only one entity that satisfies the criteria for replication in this context – the gene. But because genes are not "naked," and are instead hidden away inside larger units, we often find it convenient to talk about selection in terms of bodies, bundles of the phenotypic effects of genes (1982, 168). Bodies are vehicles and there is a legitimate debate about whether bodies are the only vehicles or if family groups, demes or even species might also be considered vehicles. So, on Dawkins' view, there is a real argument over whether individuals or groups are correctly seen as vehicles. But there is no debate between the gene and the individual as a unit of selection because individuals cannot be replicators. There may be a hierarchy of biological complexity, but it is incorrect to think of the units of selection as though they were arranged on rungs on a ladder (Dawkins 1999, 82). The debate over the levels of selection is a debate over what to count as a replicator and, for Dawkins, there is simply no debate. "Replicator survival and vehicle selection are two views of the same process. They are not rival theories." (178)

The debate over units of selection, what is selected, brings with it a corresponding debate over the appropriate level at which selection takes place, that is, what model of selection - gene-selectionism, individual selection, species selection or some other - properly describes the process of natural selection. It is often thought that Dawkins' appeal to the Necker cube is an admission of a pluralist view of the models of selection, equally valid ways of describing the same process. For Dawkins, however, the gene's eye view gets the biological facts right. While the individual view allows for descriptions that are parallel to gene-centred explanations, as in the case of reciprocal altruism, the real explanation is to be had at the level of the gene.

Wilson (2003) makes the case that levels-talk might best be seen as metaphorical. Levels of selection are often invoked by model pluralists, those who hold that distinct models of selection "are actually non-competing accounts of one and the same process."(Wilson 2003, 533) The main claim that is advanced in support of pluralism is that formal expressions of different models are intertranslatable. The implication is that "if alternative models of natural selection can be represented either in a common mathematical framework or by the very same equations, then these models differ at most heuristically, not on some deeper level..."(Wilson, 537). Dawkins seems to espouse such a view with his assertion "The selfish organism, and the selfish gene with its extended phenotype, are two views of the same Necker Cube."(1999, 6) And, like the two views of the Necker Cube, neither the gene's-eye view nor the view of the selfish individual is the correct or "true" one (1).

But as Wilson points out, the fact that any two models are expressible in the same formal expression does not entail that two processes do not or cannot differ from one another. Mathematical models capture only some aspects of the processes they model, and can serve as models only given further assumptions. Decision theory for example, has a wide application and can be used to model humans, insect societies, organizations, etc. This "pluralism," however, says nothing about the how these domains might differ (538). Even granting structural isomorphism to two models, their expressions fail to capture the fact that the quantificational range of the variables in the expression may be radically different, for example when individuals are agents in one model, while groups are agents in another. At least sometimes, then, different models are not simply different perspectives because each may make very different and sometimes incompatible ontological commitments (540). Wilson suggests we consider replacing the levels metaphor, one that implies that levels are neatly segmented and separable, with a metaphor of "entwinement". Levels are co-instantiated and may not in fact be readily isolatable, or make distinct contributions to fitness. Our inability to clearly separate them is a result of the complexity of the biological world (532).

Though Dawkins' notion of the replicator and vehicle are illuminating metaphors, highlighting some aspects of genes (their power to copy) and bodies (their "directness" in interaction with the environment), there seems to be enough disagreement over just what satisfies the criteria of each to consider the case still open. Dawkins' replicator concept and his gene's-eye view of evolution do not settle the debate over the units of heredity or the units and levels of selection.

6.4 Rulers: Genetic Information and Genetic Programs

The metaphors of instructions, information, and programs employed by Dawkins makes talk of complex social behaviours being under the control of the genes logical. Genes set things up in advance, like chess programmers (1999, 16). The program metaphor unites the two roles of the replicator as the unit of heredity and as the unit of selection by the unique role this program plays in development. The gene's special status as the *transmitter of developmental information* is the reason that the gene has a privileged role in the processes of heredity and selection. Recognizing the intimate link between the unit of selection and development, Dawkins attempts to counter any objections by dissociating his unit of selection from the unit of development: "...frankly facing up to the fundamental genetic nature of Darwinian *selection* is all too easily mistaken for an unhealthy preoccupation with hereditarian interpretations of ontogenetic *development*."(1999, 28) However, the metaphor of genetic information and programming, though not explicitly upholding genetic determinism, privileges the developmental role of genes over the role of other things that might play a role in development such as the cellular environment, the environment of the developing individual, and even the interactions of the individual with others. It is worth asking if this metaphorical description is to be taken literally, as Dawkins suggests, or if it too, like replication, risks mischaracterizing the complexity of the process it attempts to describe.

Talk by geneticists of information and code is common enough; it is a standard method of description: "only four bases in DNA can code for the huge amount of information needed to make an organism. It is the *sequence* of bases along the DNA that encodes the genetic information..."(Hartl and Jones 1999, 7 emphasis in original). Isn't this a legitimate description of the nature of DNA and its relation to inheritance and development? In his history of heredity, *The Logic of Life*, Francois Jacob (1973) explicitly recognizes the way in which the program metaphor unites heredity and development. The influence the program metaphor has had on descriptions of hereditary and developmental processes, and the extent to which this metaphor has been developed, makes it worth quoting Jacob at length:

> Heredity is described today in terms of information, messages and code. The reproduction of an organism has become that of its constituent molecules. This is not because each chemical species has the ability to produce copies of itself, but because the structure of macro-molecules is determined down to the last detail by sequences of four chemical radicals contained in the genetic heritage. What are transmitted from generation to generation are the 'instructions' specifying the molecular structures: the architectural plans of the future organism. They are also the means of

executing these plans and of coordinating the activities of the system. In the chromosomes received from its parents, each egg therefore contains its entire future: the stages of its development, the shape and the properties of the living being which will emerge. The organism thus becomes the realization of a program prescribed by its heredity. (1973, 1-2)

Reproduction is the fundamental characteristic of living systems, and because of this characteristic and the propensity for mistakes to be made in this process, evolution by natural selection occurs. Jacob, like Dawkins, picks out fecundity as the one criterion that makes the difference in the "contest for progeny," and any improvement in the program that promotes the reproduction of descendants will be passed on to those descendants (Jacob, 5). Though Jacob was partly responsible for initiating talk of the developmental program and popularizing it (Keller 2000b, 160), I think that he is modestly aware of the metaphorical nature of the description, in a way that Dawkins is not, even as he recognizes its power as an explanatory concept.

Dawkins' argument for the primacy of the gene in heredity is a version of what we might call the covariance argument. All other things being equal, a change in a specific gene at a specific locus will cause a change in the phenotype of the organism relative to those individuals who have not had such a change (1999, 21). The modesty and simple logic of this argument lies behind Dawkins' examples of a "gene-for" complicated behaviours such as reading (1999, 22).

However, covariance arguments have been weakened as general arguments for gene primacy by the recognition of parental effects on offspring development. For example, the particular species of plant upon which a butterfly lays her eggs will affect the degree to which her offspring resemble her (Sterelny 2000, 195), and the sex of a crocodile's offspring is at least partly determined by the temperature of the egg, becoming male at intermediate temperatures and female at extremes (Griffiths 2001, 396). Pervasive in Dawkins' descriptions is the assumption that "of course" genes require the proper background conditions to operate, but when those conditions are present, the genotype carries the fundamental information for the organism (Oyama 2000b). As I noted in 6.2, genetic causation is, in cases of multifactor traits, *a cause*, but not necessarily *the cause* for the appearance of traits.

Dawkins' other argument for the primacy of the genetic replicator is his bottleneck argument. The only entities reliably replicated across generations are genes because of the fact that, in sexual reproduction, any organism is "reduced" to a single cell carrying the genes. As I noted in section 6.2, however, the idea that the genes are the only entity that is reliably inherited is open to question. There is, however, a third argument for the idea that the genes are *the* instructions or programs for development, the teleosemantic argument.

6.5 Rulers: Teleosemantics

Teleosemantics is a concept in the philosophy of mind that was originally formulated to show how non-semantic elements could give rise to the semantic properties of mental representation (Rowlands 2004, 3). The teleosemantic account rests on Ruth Millikan's notion of the 'proper function' of a mechanism or trait; that is, what the trait or mechanism is supposed to do. Millikan's definition of a proper function is an attempt to provide a theoretical definition of purpose or function that captures the fact that something can be a member of a function category even if it is defective and fails to fulfil its function (Millikan 1998, 304). "It is an attempt to describe the phenomenon that lies behind all the various sorts of cases in which we ascribe purposes or functions to things…"(301) An example of a proper function in Millikan's sense is the function of the heart to pump blood.

The proper function of a trait or mechanism is determined by its history rather than by its present character (Rowlands, 2). "In particular, the possession of a proper function F by an item depends on that item existing because it possesses certain characteristics that have been selected for because of the role they play in performing F."(Rowlands, 2) This definition is enlisted to formulate a naturalistic account of mental representation: "If a cognitive mechanism M has evolved in order to detect an environmental feature E, then this is what makes it an appropriate state S of M about E; this is what makes state S mean E…"(3) The fact that M is about something, E, makes it intentional and this is what we will mean when we say that DNA has "intentional" information, for example, that the DNA sequence CAA is information "about" the amino acid Valine. And, because amino acids have been selected to bring about particular states in a developing organism, DNA sequences are about those states as well.

John Maynard Smith (2000a) enlists this notion of teleosemantics in order to develop an account of genetic information, one that allows him to say that DNA has "meaning" and so can be "misread" or "misrepresented" while other things cannot be. A human information system that uses code has the following components: a message, an encoder, a transmitter, a channel or pathway across which the message is sent, a receiver, and a decoder. It is common in genetics to describe the processes of genetic transcription (DNA⇨ mRNA) and translation (mRNA ⇨ to protein) in terms of coding and decoding. Conspicuously absent in this process, however, is a "coder." Maynard Smith argues that the information is in fact coded by natural selection (2000a, 179). Following the teleosemantic account, the working of natural selection on an arbitrary string of nucleotide bases gives DNA its "meaning." The proper function of DNA sequence TTT is initiating the production of the amino acid Lysine, a function it has been selected for because it favours the survival of the organism of which it is a part. This meaning is transmitted across the channel of DNA to RNA to protein (179). The shape of DNA has been generated randomly (by mutation) and selected because it makes proteins that build bodies that have higher fitness relative to other organisms (190). The local environment in which the proteins are formed represent the channel conditions. When organisms change because of changes in their environment, it is because they have inherited the genetic mechanisms that allow them to do so.

Environmental conditions cannot have meaning because they are not the result of some mechanism that has been selected and so they cannot have proper functions. The environment is not a source of information, but of 'noise' (192). Because the teleosemantic account enables one to make a distinction between the proper function of DNA and the environment, it justifies the biologist's distinction between genetic and environmental causes in development and in inheritance. The other benefit of the teleosemantic account is that, because genetic instructions carry intentional information, they can be misread, they can be wrong. No one suggests that the genes contain a program that says "when exposed to thalidomide grow rudimentary limbs."(Griffiths 2001, 397) In such a case the information has been misread or there was too much noise for the information to be correctly interpreted. This is not a distinction we make with other so-called developmental resources.

One challenge to the teleosemantic argument has been the appeal to a parity thesis. A fundamental feature of a human information system is that the source and channel conditions can be reversed. An analogous reversal between gene and protein development, however, is tantamount to a violation of the 'central dogma' of genetics. "'The transfer of information from nucleic acid to nucleic acid, or from nucleic acid to protein may be possible, but transfer from protein to protein, or from protein to nucleic acid is impossible…' "(Crick quoted in Maynard Smith 2000b, 214) The parity argument implies, however, that teleosemantic information will be present in *any* inheritance system that is the product of evolution by natural selection and so this concept applies equally well to epigenetic inheritance systems as well as to genetic ones (Griffiths 2001, 400). Changes in other inherited resources such as DNA methylation patterns, which are important in tissue differentiation during the life of a single organism, can cause heritable variation in all cells descended from an egg in which these structures have been altered (400). Since methylation patterns are also a product of natural selection, they too are about the states that they help to produce.

Developmental Systems Theorists, too, have tried to argue for a widening of developmental resources to include micro-organisms inherited only by some aphid castes that manufacture a chemical necessary to a growth spurt in aphids, and dispositions such as preferences for particular nests that are "imprinted" in the offspring of some birds. Since these mechanisms most likely have biological functions and were shaped by natural selection, on the parity thesis they, too, must contain information (401). Instead of arguing that all of these things constitute information, however, Developmental Systems Theorists like Susan Oyama argue that a change in metaphor is required. "[W]hat is transmitted between generations is not traits, or blueprints or symbolic representations of traits, but developmental *means* (or resources or interactants)." (2000a, 29)

As a result of the emphasis on a wide variety of developmental resources and parity between these different resources and developmental contexts, some Developmental Systems Theorists deny that there are any pre-existing instructions or programs that shape the organism from within. Information is constructed anew in each individual developmental cycle (Robert, Hall and Olsen 2001, 955).

> "...Information "in the genes" or "in the environment" is not biologically relevant until it participates in phenotypic processes. *It becomes meaningful in the organism only as it is constituted as "information" by its developmental system.* The result is not *more* information but *significant* information."(Oyama 2000b, 16 original emphasis)

Other Developmental Systems Theorists accept that there is some information prior to individual development (ontogenetic information), but that it is not exclusively genetic (Robert, Hall and Olsen, 955).

Like Dawkins' defense of the genetic replicator as the unit of selection, defenders of teleosemantic accounts of information appeal to the instability of the environment across generations relative to the genetic information to give genes special status. "Differences due to nature are likely to be inherited, whereas those due to nurture are not; evolutionary changes are changes in nature, not nurture; traits that adapt an organism to its environment are likely to be due to nature."(Maynard Smith 2000a, 189) However, defining inheritance in terms of nature and nurture already biases an explanation about what is important. The meaning of "inheritance" should be widened to include all those things that pass from one generation to another in such a way that selection can act on all variations (Griffiths 2000, 402).

Emphasizing the symbolic nature of DNA is also meant to counter the parity argument by clearly separating genetic from non-genetic factors. The information metaphor allows us to talk about genes sending signals to other genes in order to control genetic regulation in development. There is a mouse gene referred to by Maynard Smith as "*eyeless*." Mutations in this gene cause mice with the mutation to develop without eyes (Maynard Smith 2000a, 188). This gene has been transferred to various locations in fruit flies and, just about wherever it is located - in a leg, for example - a compound fly eye develops. It seems strange that a mouse gene should regulate the development of a compound fly eye, but Maynard Smith suggests that this is not so strange if we think of the gene as sending a signal to other genes to "make an eye here." By sending this signal, the gene is switching on the genes for making a fly eye rather than a mouse eye (188). Eyeless is symbolic in the sense that there is no necessary chemical connection between the nature of this inducer gene and its effects; the genes for making eyes could have evolved to have been switched on by a different gene (188). Maynard Smith suggests that information talk in this case is indispensable if one is to construct an evolutionary story for the conservation of genes controlling eye development. "[I]t is hard even to think about the problem if one does not think of genes sending signals, and if one does not recognize that the signals are symbolic." (189)

Against Maynard Smith's characterization of inducer genes as symbolic and information bearing, it has been argued that extending the information system metaphor implies that, if there is a code and information, there must be something that reads and executes that code. The *eyeless* gene mentioned above is such a "reader." Because it can make an eye that is appropriate for any organism within which it is placed, it should not be thought of as sending the signal "make an eye here," but as reading and executing many eye-making instructions (Sterelny 2000, 199). Rather than coding arbitrarily, we ought to regard *eyeless* as a reader that maps genotypic sequences onto phenotypic differences, but one that has predictable effects. *Eyeless* is not a gene that carries information; and so, not all genes carry developmental information. This conclusion weakens the general argument that *all* genes carry information, a characteristic that other candidates for units of heredity and development lack.

Further, the above conclusion should inspire us to work at other ways of looking at and extending the information metaphor and to ask if DNA is in fact the only such system. As our understanding of the role of information bearing structures and readers/executers changes, other metaphors might suggest themselves and be explored (Sterelny 2000, 200).

Strangely, it is John Maynard Smith, one of the strongest defenders of the non-metaphorical relation between human and genetic information systems, who acknowledges that scientists need to get their ideas from somewhere and it should not surprise us if they turn to current technology to get them (2000a, 179). What difference might the choice of an alternate technology make in increasing our understanding and yielding new insights into the structure and causal mechanisms of nature? The assumption that DNA is a program and that genes contain information, confuses the explanation with what is being explained (Keller 2000, 160). Realists assume that ontogenetic processes fit our notion of programs, rather than asking if our metaphors of programs continue to describe adequately ontogenetic processes and the dependency of genes on their cellular context (Oyama 2000b, 73; Keller 2000, 175). Despite the undoubted instrumental success of the genetic program metaphor, Maynard Smith might have been prompted to ask the same question with which Francois Jacob concludes his history of the study of heredity: "Today the world is messages, codes and information. Tomorrow what analysis will break down our objects to reconstitute them in a new space?" (1973, 324)

It may be that a better understanding of the history of explanation of biological disciplines would temper the tendency to confuse interaction metaphors with a description of nature's structure finally laid bare. Evelyn Fox Keller has produced a brief history of the way in which the state of technologies and the influence of interdisciplinary work have helped to locate the developmental program squarely in the genes, despite the recognition that development is a complex process with many requirements. The genetic program represents an entrenched "conceptual shortcut" which suggests that research metaphors, and the research programs they help to guide, are contingent on the state of technologies in other areas, in this case, automata theory, cybernetics, artificial life, and molecular biology (Keller 2000a; 2000b).

6.6 Reasoners: Strategies and Prisoner's Dilemmas

Evolutionary Stable Strategies are ways of thinking about the evolution of certain behaviours. For Dawkins they are strictly a matter of convenience. "[Y]ou don't actually have to use ESS language, provided you think clearly enough." ESSs are "a great aid to thinking clearly, especially in those cases...where detailed genetic knowledge is not available."(1989, 287)

The soliloquies and rational calculations Dawkins makes use of are, he says, just a way of pointing out that decisions about what to do should "ideally be preceded by a complex, if unconscious, 'cost-benefit' calculation."(1989, 69) Being ideal, however, means that they are not found in nature. The strategies are behavioural rules pre-programmed by the genes, behaviours that are manifested in dispositions to act in particular ways in particular circumstances. So, we can think about these rules as if they were embodied in conditional commands: "Attack opponent; if he flees pursue him, if he retaliates run away."(69) What is most useful about this way of thinking is that we can successfully use this method in the absence of any specified method by which the rules themselves are made. If ESSs are simply a language of convenience, a substitute for *real* thinking in terms of changes in gene frequencies, why would they work well enough, even in the *absence* of detailed information about their genetic basis, to be able to predict behaviours that one can observe in nature?

Speaking about plants and animals as if they were making conscious decisions about how to increase their success is a common and relatively harmless explanatory device in biology (Dawkins 1989, 278). The use of ESSs and Prisoner's Dilemmas allows us to consider how fish and bats and genes might behave if they were rational optimizers. As Dawkins points out, however, being fit does not entail acting rationally, behaviour is functionally equivalent to acting "as if" it was done with conscious intention (1989, 64). Unthinking behaviour that appears as if it is rational does not require a brain. It requires only that the organism that displays it be responsive to its environment, to the behaviour of other organisms, and that its behaviour is inherited and that these behaviours differentially affect the fitness of the organism and those around it (Axelrod and Hamilton, 1392). Game theory does not need to deal with preferences or intentions at all. "Natural selection 'chooses' or 'prefers' the behaviours that will maximize survival value. This is all we need to apply the mathematics of game theory, even though no conscious choices or preferences may be involved." (Poundstone 1992, 235)

The use of ESSs and conscious intentions is not just convenient. Though their use sacrifices the mathematical rigour of other models; they are *explanatory*. Dawkins describes the hawk and dove story as a simple model, not something that happens in nature. Nevertheless such models help us to understand things that happen in nature and, as they are gradually elaborated and made more complex, they come to resemble the real world more (74). This strikes me as a peculiar thing to say. If the simple model does not already resemble something happening in nature, how does it help us to understand nature? It appears that Dawkins has unwittingly thrown himself upon the horns of a dilemma: If the simple model has no connection (that is, resemblance) to the actual working of nature, then there seems little reason to think that a more complex one has greater resemblance. (Why would complexity, per se, be a criterion that shifts one from no resemblance to resemblance?) Or, the complex model resembles the real world and (absent some mechanism that specifies how one moves from no resemblance to resemblance), so does the simple one. I think that Dawkins is pointing to superficial details here and neglecting the deeper structural reasons for why any model works at all. The same criticism can be made of his metaphors, recognized and, especially, unrecognized. The structure of the simple model already specifies the structure that is then further revealed by elaboration. Perhaps this is what Dawkins means.

The Prisoner's Dilemma gives us a new way to think about the evolution of non-human behaviours by employing intentional language. It allows us to predict the sorts of strategies or "pre-programmed behaviours" that we might expect to find in nature, and it seems to do quite well. But if anything is metaphorical, surely the attribution of rationally calculating genes is. Dawkins notes,

> A computer program can behave in a strategic manner, without being aware of its strategy or, indeed, of anything at all. We are of course, entirely familiar with the idea of unconscious strategists, or at least of strategists whose consciousness, if any, is irrelevant. Unconscious strategists abound in the pages of this book. Axelrod's programs are an excellent model for the way we, throughout the book, have been thinking of animals and plants, and indeed of genes. (1989, 228)

I think that there are two points here worthy of notice. First, there is a metaphor at work. The metaphor that guides the construction and application of the models of the Prisoner's Dilemma and Evolutionary Stable Strategies is the metaphor of non-human animals and genes as rational decision makers. The second point is that it is an excellent metaphor/model. Excellent because it predicts the sorts of situations and behaviours that we find in nature and, by virtue of this fact, gives us a way to construct a causal story that explains them. "Selection on generations of bacteria has presumably built into them an unconscious rule of thumb which works by purely biochemical means."(Dawkins 1989, 229) ESSs and Prisoner's Dilemmas are parts of the evolutionary explanations that we give about why the living world is as it is.

Being metaphorical, ESSs and Prisoner's Dilemmas cannot avoid sometimes misleading us. The use of soliloquies in the determination of ESSs can lead to the assumption that populations will oscillate wildly between two kinds of strategies rather than reaching an equilibrium of mixed strategies or a stable mixed strategy, a conclusion one reaches using more rigorous mathematical analysis (Dawkins 1989, 76). The predictive power of these models is so strong that Dawkins asserts, "We must expect to find evolutionary stable strategies everywhere."(73) But why, if such soliloquies are a language of convenience should we expect to be able to find them everywhere?

The reason that this metaphor works so well and why should we expect them to be ubiquitous is that the process of evolution by natural selection mimics the process of trying both smart and stupid moves, leaving only the "smart" moves for us to see. This result gives us the illusion that natural selection has deliberately chosen some "designs" rather than others because they would be good at what they do; they are well "designed." Creating explanations for why those things are well designed involves uncovering the reasons embodied by the success of organisms with those traits (Dennett 1987, 190).

6.7 Reasoners: Cognitive Metaphors

Dawkins' description of genes as *programmers*, as *cooperating* in their effects on behaviour, and as *giving instructions*, are attributions of human cognitive qualities to entities that we are sure do not possess them. Why then does he use them? His predominant reason is that we are used to talking in this way about behaviour. It is a language of convenience, but one that can be shed at any time for the language of the population geneticist. "At times gene language gets a bit tedious and for brevity and vividness, we shall lapse into metaphor. But we shall always keep a sceptical eye on metaphor, to make sure they can be translated back into gene language if necessary."(1989, 45) Describing genes and non-human organisms as if they had conscious aims allows us to ask questions that might illuminate the evolutionary processes that are at work (88). However, if it is a language of convenience only, it seems miraculous that it might produce questions that lead to explanations at a deeper level, suggesting hypotheses for detailed experiment, experiments that yield new and important understanding. The metaphor of the intelligent gene scheming to secure its own survival is powerful and illuminating (Dawkins 1999, 15). So, perhaps it is not convenience *only*, as Dawkins so often claims. In her discussion of the use of cognitive metaphor in biology, Margaret Boden (1981) describes what she sees as the positive and negative aspects of the use of cognitive metaphors in biological descriptions.

> Despite the extension of use that is involved...a cognitive approach to biology can be defended on both positive and negative grounds. Negatively, it is unlikely to mislead people into mistakes at the psychological level, because the two domains are prima facie so different. Positively, it can encourage biologists to ask empirically fruitful questions, questions that a purely physico-chemical approach might tend to leave unasked. (108)

I want to make a slightly stronger claim that it is not that the physico-chemical approach would tend to leave such questions unasked, these questions could not be recognized as ones worth asking if they were not posed within the terms of these metaphors. In a similar vein, Grant (1991) argues that the etiology of the metaphors of cutting losses, investment strategies, risk calculations, etc., lies in a direct application of modern business executive strategy to the biological sphere (441). Acknowledging this fact, however, raises the spectre of the inventive aspect of metaphor that is so troubling for the realist. "The only question is whether these are simply vivid ways of illustrating what is going on in nature or whether they represent an imposition on nature from the human realm." (Grant, 441) Describing genes and non-human animals in these ways leads us to ask further questions. For example, if genes are programs, who is the programmer? Assigning the role of programmer to the gene as well makes the appearance of a meaningful program theological – a trap into which Dawkins falls with his declaration that "DNA just is." (Grant, 433) If Dawkins wanted to avail himself of this metaphor without sounding theological, he needed to attribute the programming role to natural selection. As I tried to argue in sections 6.4 and 6.5, however, elaborating the program metaphor in this way does not help matters.

Dawkins has employed a host of metaphors that privilege the gene in many ways. I have tried to show that all of these metaphorical descriptions - of genes as replicators, reasoners and rulers - illuminate as well as obscure, expand as well as restrict our view of biological phenomena. Traditional assumptions about how metaphors work lead us to neglect this fact. Accepting descriptions that privilege the causal role of genes in heredity, selection and development as non-metaphorical is incompatible with our growing understanding of genetics and evolutionary biology. Equally, however, antirealist views about metaphor in biology must confront the predictive and explanatory power these metaphors seem to possess. Evelyn Fox Keller has examined in detail the function of the metaphorical description of the gene and it seems fitting to end this section with her observations:

> The assumption [of] materiality (and especially particulate materiality) lent the gene its fixity, that is, the permanence it required as a stable unit of transmission...But endowing the gene with the power to act added to the property of materiality the further implication of agency. The capacity to reproduce itself – traditionally the defining property of life – lent the gene vitality. And finally, attributing to the gene the capacity to direct or control development effectively credited it with a kind of mentality – the ability to plan and delegate. (2000a, 47)

6.8 Replication: An Interaction Metaphor

Let's begin by reviewing the main features of the interaction metaphor. In the metaphor "man is a wolf," an implication-complex about wolves is imposed on a corresponding implication-complex about man. The wolf metaphor brings into sharp focus the aspects of man that can be easily talked about in "wolf language" and pushes into the background those things that cannot. Through this interaction, humans become more wolf-like and wolves become more human. It is not necessary that the set of implications already be in mind when the metaphor is constructed and it is an essential aspect of the interaction view that they need not be. The metaphor might allow for a series of implications to be explicated ad hoc, but any attempt to translate aspects of the metaphor into some non-metaphorical man language will change the understanding of man that we achieved through the metaphor.

If Dawkins' descriptions are interaction metaphors, we should have some success constructing implication complexes for them. Attempting to make these complexes explicit will also help to suggest why the metaphors can be both illuminating and misleading. Here, then, is a possible implication complex for "the gene is a replicator" (Table 6.1):

Replicator	Interaction	**Gene**
(1) Thing of which copies are made.		DNA fragment is copied in meiosis
(2) Discrete thing	⟷	Fixed boundaries, discrete bit of DNA
(3) Copy is near perfect.	⟷	Copied relatively faithfully.
(4) Copied indefinitely	⟷	Copied for many generations
(5) Lasts long enough to be used to make more replicators	⟷	Lasts many generations and acts as template for other genes.
(6) Retention of structure in subsequent copies.	⟷	High degree of structural stability (low mutation rate).
(7) Causally related to appearance of next copy (e.g., paper to copy)	⟷	Acts as a template for the replication of descendants.
(8) Changes in original affect chances that replicator is copied.	⟷	Traits coded for by genes affect fitness of organism and chances that DNA will be copied.
(9) Requires other machinery/ apparatus for copying process (e.g., photocopier)	⟷	Requires enzymes (RNA polymerase), other proteins, and cellular environment for copying process.

The Gene is a Replicator

Table 6.1

The first and perhaps most obvious point to make about why we ought to view the "replicator gene" as an interaction metaphor, rather than as a comparison or substitution metaphor, is that Dawkins does not want to *compare* the gene to a replicator. He wants us to see that the gene *is* a replicator. It is paramount for Dawkins' view that we come to see the gene *as* a replicator; only an interaction metaphor has this power.

A second reason why this metaphor should not be seen as a comparison metaphor is that, if it were a comparison metaphor, we would have to be able to state the ways in which replicators and genes are similar prior to the construction of the metaphor. When Dawkins explicates the replicator metaphor, he does not ask about the ways in which a gene and a replicator are similar. He tells us that genes *are* replicators and proceeds to explicate just what it means for a gene to be a replicator.

The choice to describe the gene in terms of the replicator determines the system of implications that control the description of the gene. Those aspects of the gene that can be talked about in replicator language - in terms of copying - are highlighted. Implications (1), (3), (5) and (8) seem particularly amenable to replicator talk. The filtering of the gene through the focus of the replicator means that some aspects of the gene are pushed into the background.

Implication (2) poses a particular problem because replicators are concrete things, but as Dawkins stresses, his gene is a conceptual unit, not a physical unit. The physicality of the replicator lends to the gene a discreteness that it does not have and this explains in part why Dawkins is repeatedly arguing against those who object to his characterization of the gene as a discrete unit. The replicator makes this a particularly salient feature of genes even though it is misleading to see genes in this way.

Explicating the replicator metaphor more fully leads to an awareness of (9), an awareness that, I thought, would have made Dawkins more likely to acknowledge the resources required for replication. If a gene is to be copied, then all the relevant copying mechanisms must also appear in each generation. Recognition of this fact might make Dawkins less certain that only the gene has the required longevity, fidelity and fecundity to be the unit of heredity and selection. Making the implication-complexes explicit also makes it easier to see that we might see other systems of inheritance, such as epigenetic systems, through the replicator frame. It is also important to note that the "gene is a replicator" metaphor is particularly resonant. It may not be fully explicated, but that will not deter us from spelling it out further in an ad hoc way. It is likely that there are as yet unknown characteristics of genes that can be talked about in replicator language.

Because the focus and frame interact, the replicator is in some sense seen by us as more gene-like. Some evidence for this is the weakening of the requirement of fidelity in copying. The copying process need not be exact; there is a sliding scale of fidelity along which candidate replicators can be placed.

Finally, the implication-complexes are not necessarily joined by similarity, yet another reason why the "gene is a replicator" should not be seen as a comparison metaphor. The implications interact in various ways. The ideas in (1) might best be seen as a similarity relation. We might best see (8) as an analogy. For all their differences, however, the associated ideas interact as sets, not as particular correspondences.

I have found no indication in the literature that the other metaphors I have discussed have been examined in this way, but, if we are to continue with this form of analysis, we must proceed with caution. Comparing the implication complex point by point is to have already abandoned the interaction metaphor and to be verging on a formal comparison. To paraphrase Black, what we have above is DNA *compared* to something of which copies are made rather than DNA *as* something of which copies are made (Black 1955, 285 note). Viewing the "gene is information" metaphor as interactive makes prominent the fact that the focus can be used to further illuminate aspects of the frame. For example, a relevant question for the genetic information system was how the code has meaning. As we saw, this was answered in part by appeal to the meaning something can possess by virtue of having the correct history.

Changing metaphors imply that though interaction metaphors connect with some aspects of the world, they do not do so once and for all. Having come this far, the metaphysical and epistemological questions loom large. In the next chapter, I will outline a view of metaphor that I think explains why Dawkins' metaphors allow us to make predictions and to construct causal stories about the evolution of traits. I will end with a discussion of some of the metaphysical and epistemological implications of this view of metaphor.

Chapter Seven

Metaphor, Induction, and Patterns

7.1 Connecting Metaphor and Inductive Success

In Chapter Two, we saw that one form of explanation in evolutionary biology proposed causes for the traits that organisms display. These explanations take the form of historical narratives that explain how a trait helped an organism's reproductive success. These adaptive stories allow us to make predictions about other traits we might find in nature. We also saw that models, metaphors and analogies are closely related, and that models and analogies may even be derivatives of metaphors.

In Chapter Three, I explained why I think that questions about metaphor tend to motivate realist and anti-realist debates concerning the status of metaphorical entities and I described how metaphors could provide insight as well as foster obscurity. In light of these debates, instrumentalism was suggested as a suitably modest position from which to begin our examination of the role of metaphor. On an instrumentalist view, scientific theories are judged on the basis of their predictive success alone, and any claims to a theory being true or false are rejected. One detrimental aspect of the instrumentalist view was that it replaced the scientific goal of explanation with prediction, so a tension was acknowledged between adopting this provisional account and the view that evolutionary explanations provide causal accounts as well as making predictions. It was also noted that adopting instrumentalism as a provisional philosophy of science raised two problems - accounting for predictive success and clearly distinguishing between observational and non-observational entities.

Chapter Four began with an account of traditional views of metaphor - the substitution view and the comparison view. Both views were unduly restrictive of the role metaphor could play in telling us how the world is. The interaction view, it was argued, addressed the shortcomings of these traditional views by describing metaphor in terms of the mutual interaction of two systems of implication. Metaphors function as cognitive instruments that can inform us about how the world is. The strong creativity thesis was rejected on the grounds that metaphors are not immune to testing against our experience, but a tension was recognized here between the commitments of the instrumentalist view and the claim that similarities between metaphorical subjects might in some sense be present prior to their recognition. Setting this difficulty aside, the interaction view of metaphor was seen to be compatible with instrumentalism.

Having a provisional philosophy of science in hand and having sketched out what I took to be the most adequate view of metaphor, I turned in Chapter Five to a summary of Richard Dawkins' gene's-eye view of evolution. This summary was intended to illustrate the importance of metaphor in Dawkins' descriptions and to show that he holds many of these descriptions to be merely terms of convenience, terms always paraphraseable in non-metaphorical language.

In Chapter Six, I examined three of Dawkins' most prominent metaphors, what I called replicators, rulers and reasoners. I argued that the replicator in both heredity and selection is not necessarily limited to the genes and that emphasizing the process of replication tends unfairly to privilege the gene over other entities. The metaphor of the genetic program and genetic information was difficult to reconcile with our knowledge of heredity and development. The use of game theoretic concepts and cognitive metaphors was interesting because, though prima facie metaphorical, they proved very successful at predicting behaviours. I concluded the chapter with an articulation of a possible implicative complex for the metaphor "the gene is a replicator." Regardless of the other functions that they might have, e.g., pedagogical or heuristic, Dawkins' metaphors are essentially predictive instruments, but they also seem to provide causal explanations.

In this final chapter I want to try to bring together the various themes that have been discussed to offer a tentative answer to the question how metaphors, and the theories they help to generate, can be predictive and provide causal explanations of biological phenomena.

I will argue that we ought to view the predictive power of metaphor as a product of its ability to make manifest otherwise indiscernible patterns, patterns that connect with the causal structure of the world. It is this connection between metaphor and prediction that ultimately explains the presence of metaphor in explanations. Accepting this view, however, also entails modifying our provisional instrumentalist view in a way that gives it a mildly realist flavour.

7.2 Thomas Kuhn, Richard Boyd and The Function of Scientific Metaphors

To begin, let us look at two views, those of Thomas Kuhn and Richard Boyd, both of which accept Black's interaction view as an account of metaphor, but that reach different conclusions regarding the function of metaphor in scientific theorizing, one vaguely instrumentalist and the other realist. Boyd ties the interaction account of metaphor to a causal theory of reference and a notion of natural kinds that allows for a robust scientific realism. In contrast, Kuhn ties the interaction view to the incommensurability of scientific terminology and the production of resemblances based on exemplars, and he denies that we can arrive at a final description of the way the world is.

What is the connection between interaction metaphor and the world such that it can allow us to make successful predictions? To begin, consider the following remarks by Richard Boyd and Thomas Kuhn made in giving accounts of the use of metaphor in science:

> The causal structures to which our language is accommodated exist quite independently of our conceptual schemes or theory construction. We do not decide by convention where the boundaries of natural kinds lie. Neither do we, in any important sense, "construct" the world when we adopt linguistic or theoretical frameworks. Instead we accommodate our language to the structure of a theory independent world. (Boyd 1993, 532)

> What is the world, I ask, if it does not include most of the sorts of things to which the *actual* language spoken at a given time refers? Was the earth really a planet in the world of pre-Copernican astronomers who spoke a language in which the features salient to the referent of the term "planet" excluded its attachment to the earth? Does it obviously make better sense to speak of accommodating language to the world than of accommodating the world to language? (Kuhn 1993, 541)

7.3 Thomas Kuhn

For Kuhn (1993), Black's principle insight was his insistence that metaphors need not presuppose or supply a list of similarities between the two metaphorical subjects. The resulting open-endedness and imprecision of metaphor, Kuhn thinks, has a parallel with the way in which theoretical terms are introduced and elaborated in scientific theorizing and investigation. The determination of kind membership and the application of kind terms in theories is done on the basis of resemblance, in much the same way that Wittgenstein described our determination of members of the class of games. Famously, Wittgenstein argued that for some terms, like "game", there is no description or set of descriptions that would satisfactorily describe all things that we refer to as a game.

> What is common to them all?...if you look at them you will not see something that is common to *all*, but similarities, relationships, and a whole series of them at that...we see a complicated network of similarities

overlapping and crisscrossing: sometimes overall similarities , sometimes similarities of detail. (Wittgenstein 1968, sec.66)

We recognize games not by virtue of a set of necessary and sufficient conditions, but by their resemblance to one another, much as one might recognize members of the same family. Our application of kind terms is grounded by reference to a "dubbing ceremony." Whether a name refers to an individual or not is determined by looking back to see if there is the appropriate historical connection to an original naming event.

Paradigm kind terms, like that of any chemical element or biological species, such as nitrogen or beaver, are not always defined in terms of necessary and sufficient conditions so, they have a degree of indeterminateness that permits borderline cases. In introducing and applying theoretical terms, there is a metaphor-like process at work. "Family members" provide the basis for the creation of the application of kind terms; they invite the language learner to discover characteristics that make them similar (Kuhn 1993, 537). Kind terms, Kuhn says, require a number of ostensive acts or naming events to fix the feature space (the range of properties) and the salience (the relevant properties) for inclusion in a kind (537).

The discovery of the similarities that facilitate our delineations of kinds and our learning the range of reference of kind terms is much like what happens in the production of an interaction metaphor. "...The juxtaposition of examples calls forth the similarities upon which the function of metaphor or the determination of reference depend."(537) Just as a metaphor can be meaningful when the two subjects of an interactive metaphor do not share an explicit list of shared characteristics, kind terms may be meaningful without providing explicit definitions for their meaning.

Despite this similarity between interaction metaphor and theoretical terms, Kuhn sees the application of kind terms as prior to the construction of a metaphor because the focus and frames of a metaphor (frequently kind terms) must be understood prior to the formation of the metaphor. If you do not understand the various referents of "war" and "game" then you will not understand the metaphorical expression "war is a game" (537). I find this characterization of kinds puzzling. If both kinds involved in the metaphor must be known prior to the interaction of the metaphor, then one of the main functions of a scientific metaphor, that of organizing one's conceptualization of a phenomena in an attempt to understand it, does not seem necessary. This implies that metaphorical expressions are capable of reorganizing the properties of a thing, but would not be necessary for discovering what those properties are. The realism implicit in this interpretation of Kuhn's view of metaphor seems at odds with his historicism, a consequence of his rejection of our ability to describe "the joints of nature." "If Boyd is right that nature has "joints" which natural-kind terms aim to locate, then metaphor reminds us that another language might have located different joints, cut up the world in another way" (1996, 537). For Kuhn, the history of science shows us that theoretical terms do not zero-in on the joints of nature (541). On the view above, however, the reference of at least some things is fixed some way other than non-definitionally, perhaps definitionally.

The view of metaphor sketched by Kuhn suggests a way in which our provisional position of instrumentalism might be modified. The metaphors that we use and the corresponding "joints" that these metaphors locate should not be seen as true or false, they are useful. Metaphors are articulated and elaborated in science with an eye to solving technical problems. The articulation of these metaphors and the fixing of the referents of their terms allows for an increased facility in solving these problems, but not in pronouncing on the structure of the unobservable world. Though Kuhn's view might be suggestive of the modifications we might make, there is reason to think that the two positions are incompatible. Kuhn is willing to make truth claims about both observable and unobservable entities and processes and he has a tendency to describe things in such a way that suggests that with a change of instruments it is not only our view of the unobservable world that changes but also our view of the observable world.

7.4 Richard Boyd

Richard Boyd (1993) agrees with much of what Kuhn says. He focuses on Black's interaction account of metaphor, emphasizing the imprecision and open-endedness of metaphor. He also accepts family resemblances and the use of non-definitional reference fixing as an account of the introduction and application of theoretical terms. Where Boyd disagrees with Kuhn is over what the basis of reference implies for the epistemology and metaphysics of science. At the heart of Boyd's account of metaphor in science is the "theory-constitutive metaphor," metaphors that, at least for a time, scientists use to express theoretical claims and that cannot be phrased in some subject-specific, non-metaphorical language. As prime examples of such theory-constitutive metaphors Boyd singles out computational metaphors in cognitive psychology, metaphors that describe the brain in terms of a computer, and thought in terms of a computational language and information processing. At the time Boyd was writing, he believed that there were no non-metaphorical paraphrases available for making the same theoretical claims. So, "it is clear that these computer metaphors are theory constitutive."(487) It seems as though Boyd's theory-constitutive metaphors might make him a possible ally in describing theories as instrumental constructs. As we will see, however, his account of the elaboration of metaphor makes commitments that are incompatible with instrumentalism.

On Boyd's account, the underpinning of reference, the reason why we can have theoretical advances and make discoveries, is the epistemic access we gain to the world's causal structure by virtue of picking out relevant structures and properties in our delineation of kinds. Boyd thinks that interaction metaphors, in their role as theory-constitutive metaphors, may eventually be subjected to "denotational refinement," an ever more explicit formulation of their similarities, in the attempt to find a precise fit between language and a complex world (Boyd 1993, 482). Although scientists may not understand what features referred to by the metaphor are essential to the delineation of a given kind, they can make informed guesses about the important similarities and analogies between the two subjects. "The aim of the introduction of such terminology is to *initiate investigation* of the primary subjects in the light of an informed estimate of their properties."(494)

The explication of theory constitutive metaphors is, for Boyd, one of the ways that scientific language accommodates itself to the world through a socially coordinated refinement of partially understood phenomena. The precision of fit between theoretical terms and the causal structure of the world explains inductive success.

Because of the connection Boyd makes between metaphor and the structure of the world, his position raises the question of whether or not metaphors can support predictions and inferences, yet remain non-substitutional – i.e., incapable of being paraphrased. Is denotational refinement equivalent to the idea that theory-constitutive metaphors are amenable to paraphrase without loss of cognitive content? Recall that the implausibility of having such a paraphrase was one of Black's chief reasons for holding that interactive metaphors ought to be viewed as cognitively meaningful.

Accepting Boyd's view of metaphor would mean interpreting Dawkins' metaphor of replication as giving a detailed description of an actual, albeit very complicated, process. Our inability to specify the necessary and sufficient conditions of such a processes is part of the reason we need the metaphor to help accommodate our scientific language to the structure of the world. The reason why this metaphor works is that there is a cluster of properties in the process of inheritance and selection that match that of the copying process from which the metaphor is drawn (Boyd, 528). This feature of Boyd's "homeostatic property cluster" view raises some difficulties in separating it clearly from Kuhn's. The need for theory constitutive metaphors and the presence of homeostatic property clusters suggests to Boyd that "The properties that constitute the homeostatic property cluster may not even be finite in number and they may vary significantly from time to time or place to place, so that a finite characterization of them (much less a cognitively tractable characterization) need not be possible." (529) There is in principle no reason for us to expect that all successful theory-constitutive metaphors will become fully explicated (529). In those cases, I think, there is nothing to decide between Boyd's realist view and Kuhn's constructivism because there is no basis upon which one might say that a metaphorical description has been fully explicated and has become non-metaphorical or whether it is not fully unexplicated and some aspects of it remain metaphorical. More problematically, however, we need a clear specification of what it means for a metaphor to be fully explicated and how we might recognize such cases.

7.5 Disagreements

Kuhn does not share Boyd's optimism regarding metaphorical descriptions of the world. Though he does think that metaphors establish links between our scientific vocabulary and the world (for this reason, Kuhn takes both himself and Boyd to be realists (539)) those links, according to Kuhn, are never given once and for all. "Theory change, in particular, is accompanied by a change in some of the relevant metaphors and in corresponding parts of the network of similarities through which terms attach to nature."(539)

Interestingly, Kuhn and Boyd agree on the fact that the changes in similarities result in more effective ways of dealing with some natural phenomenon, but where Kuhn stops at this instrumental level, Boyd goes on to cast this effectiveness in terms of epistemic access to the causal structure of the world, an access which is marked by our ability to collect information about the kinds in question and to make new discoveries (Boyd, 528). The introduction of epistemic access and the accommodation of language to the world as a ground for reference is, on Kuhn's account, Boyd's response to the empiricist problem of induction. The problem of explaining how successful inductive generalizations are possible is thought to be solved by explaining that kind terms correspond to the things in the world that exist independently of our conception of them.

If our references to kinds were simply a result of our interests or fancy then we would hardly expect them to generate successful inductions. This sort of response is usually motivated by what I like to describe as "the incredulity thesis" and Boyd expresses it well: "It is impossible to understand scientists as being in the business of achieving non-accidental success at inductive generalizations without understanding them to be in the business of learning about (typically "unobservable") causal powers and underlying mechanisms of structures." (511) Remember that it is this knowledge of unobservables that the instrumentalist wants to deny. But if language is not anchored to the world via its ability to help us get into better relations with it by virtue of an ever-increasing accuracy, then successful prediction seems simply miraculous. The challenge that Boyd has raised for Kuhn, and for us, is how to explain successful induction while denying us ever better epistemic access to the structure of the world. Must theoretical language "fit the world" in order for us to have inductive success? If no plausible account can be given in instrumentalist terms, then we might be forced to adopt a realist philosophy of science – a philosophy that can provide an account of inductive success.

7.6 Explaining Inductive Success

If the argument from instrumental success to contact with the causal structure of the world is persuasive, this still may not lead us directly to an acceptance of scientific realism. Does Kuhn's answer have something that might recommend it to the instrumentalist?

The activities of scientists in between times of dramatic theory change are what Kuhn terms "normal science." "Normal science" is comprised of the development and investigation of hypotheses that take place within a particular theoretical framework. The results of this sort of work are an increase in the scope and precision with which a paradigm can be applied. "Paradigm," according to Kuhn, refers to several things. For our purpose, it is sufficient that one of the things to which it refers is an "exemplar", that is, the problems and problem solutions that act as models for educating scientists and evaluating work within a scientific discipline (Kuhn 1996, 187). In order for a scientist to be able to see problems and to solve them, he must first learn scientific theory and the rules for applying it. Exemplars help scientists to do this. A scientist learns a particular theory and the rules for its application through an educative process in which she learns to pick out the relevant features of an experimental situation (188). The exemplars, the accepted problems and problem solutions of a discipline, give laws and theories their empirical content. Becoming competent in using the exemplars establishes a "group licensed way of seeing" a situation (189). Repeated exposure to exemplars creates habituated dispositions to see particular situations as similar to or dissimilar to exemplars. One of the things a scientific community assumes with its paradigm or exemplar is a criterion for choosing problems that are assumed to have solutions (this is one of the analogues that scientific investigation has with other puzzle-solving activities). Training in the use of exemplars gives us a "feel" for puzzles much like training in artistic discrimination might give the art expert a "feel" for forgeries.

Educating scientists in the use of exemplars establishes a history of regularity required for a scientist to project theoretical terms. When a situation is recognized as exemplary, this recognition brings with it a history on the basis of which a candidate predicate is to be projected. Thus, for Kuhn, the expectation generated by the recognition of an exemplar is not a function of having found the "joints of nature"; it is a result of the linguistic practice and institutional nature of science. Such a view of inductive success does not deny that we make contact with the world, but it does deny that this contact results in a description of the one way in which the world is. This instrumentalist-sounding talk about scientific theories and the role of metaphors is reminiscent of the description of maps given in Chapter Four. A main feature of maps was that, despite their instrumental origins, they had to support successful navigation, and this was a result of making contact with the world, if even in a selective way.

Kuhn argues that the appearance of the increasing success of scientific work is in part a function of the fact that disciplines work around exemplars and in part a function of the creative aspect of puzzle solving. He contends that, from the view of any single community, the result of such creative work will be progress (Kuhn 1996, 162). Even art, what some might see as the least progressive of activities, has had periods of undoubted progress, such as that displayed in the increasing ability for artists to achieve realistic representation during the Renaissance (161). What Kuhn suggests, however, is that this progress is in some ways only apparent. It is the result of a lack of competing exemplars or paradigms within a discipline. What is true of art, Kuhn thinks, is true of science: "Scientific progress is not different in kind from progress in other fields, but the absence at most times of competing schools that question each other's aims and standards makes the progress of a normal scientific community far easier to see" (163).

Progress is easier to see in part because agreement on exemplars means that first principles are not being questioned, leaving scientists to work out the details of the exemplars. This focus on details leads inevitably to an increase in the efficiency and effectiveness of solving problems, appearing to us as increasing success (164). Like the direction of the process of biological evolution, however, the evolution of scientific thought is not *toward* anything, despite an increasingly detailed understanding of nature (170). "...scientific development is, like biological, a unidirectional and irreversible process. Later scientific theories are better than earlier ones for solving puzzles in the often quite different environments to which they are applied." (Kuhn 1996, 206) If Kuhn is justified in describing science as a "set of instruments for solving technical puzzles in selected areas" (1993, 541), then it seems that he can offer an explanation of progress and success, albeit one quite different from our intuitive picture.

The main lesson we might take away from the discussion above is that we can embrace the instrumentalist aspects of scientific theorizing and have an account of inductive success. Progress, our notion of which relies on a concept of induction that gets us ever closer to the one true description of nature, need not in fact make this commitment. Kuhn's response to Boyd's challenge has suggested that a reasonable explanation for this success, or lack of it, is that descriptions connect with the causal structure of the world. Making this concession does not necessarily bring with it claims of truth or falsity regarding theories and metaphors, only claims of instrumental success, a fact that we can point to.

I find Boyd's response to the problem of prediction unsatisfactory for the obvious reason that accepting it entails giving up on instrumentalism and adopting a realist position. But I find Kuhn's response equally unsatisfying because I think that he has relied too heavily on the structure of scientific communities in describing why progress is only apparent. There is an intermediate position, one that I think helps to elucidate the reason for the success of metaphor as a predictive instrument by virtue of its connection with the causal structure of the world, but that does not require that we give up completely on instrumentalism.

7.7 Predictions and Patterns

Both Boyd and Kuhn appeal to recognized similarities as the reason for the successful application and successful employment of metaphor. I think it is here where a more subtle formulation of the connection between metaphorical descriptions and inductive success is possible by asking ourselves what virtue similarities have that enable us to make predictions based on them. The answer I think, and this is one that was noticed by people long ago, but perhaps especially by mathematicians, is that similarities are important for the purposes of prediction when they form a pattern. Considering predictive power in terms of patterns might allow us to finesse the fine lines drawn between instrumentalism, Boyd's realism and Kuhn's constructivism.

I take it as uncontroversial that our predictive ability depends on our ability to find a pattern - faced with patternlessness one could not make successful predictions (Dennett 1991, 29). The most useful way to think of patterns in this context is, following Dennett (1991), to see a pattern as a description of data that is more efficiently communicated than a "bit map" description of the same data. A bit map is a description that contains as much information as what is described - "an effective verbatim quotation." (Dennett 1991, 32) For example, communicating a random series of numbers to someone, e.g., 7, 23, 52, 55, 1000, ..., will require that you tell them every number in the series. However, if I had a series in which I could discern a pattern, e.g., 2, 4, 6, 8, 10, ..., I would only need to specify the first number and a rule, "add two," and the number where to stop. As another example, consider a bit map of a human genome. This might be a description of the relevant DNA in terms of the elements and bonds involved in all the nucleotide base pairs. (A description in terms of the nucleotide base pairs is to already ascend to a pattern-level description.) Any description of data that is at a level higher than a bit map description is a pattern (Dennett 1991, 32). One might argue that to see things in terms of molecules is already to have discerned a pattern, but the important point is that a bit map can capture randomness, a patterned description cannot.

Basing the success of predictions on patterns, then, requires that there is a pattern in the world to exploit (30). This in turn implies that the patterns we discern are real; the relations between properties and processes are there, in the world, independently of whether they are perceived by us or not. Thus, discovering a pattern is finding out something about the world. The pattern view seems to push us toward realism. At the same time, however, this view implies that there will be many equally compatible patterns of the same phenomenon. Does the pattern view commit us to pluralism? It would seem so. It is possible that two individuals might find different patterns in the same data, and this is not in itself reason to assume at the outset that one is correct and the other wrong – recall the debate over the levels of selection. If both patterns are equally good at predicting for the same purpose (e.g., for betting, for preventing accidents or for predicting traits) then both are equally real (48). Patterns are a way of organizing data (44) and differences in knowledge will affect the properties and relations that are salient for us producing differences in our capacities to pick out patterns (34). Dennett even goes so far as to suggest that "Other creatures with different sense organs, or different interests, might readily perceive patterns that were imperceptible to us. The patterns would be *there* all along, but just invisible to *us*." (Dennett 1991, 34 original emphasis)

7.8 Interpretive Stances

The patterns that Daniel Dennett thinks we move between are really three different "interpretive stances" or predictive strategies that we take toward the world in order to generate sets of expectations about how something will behave. These three stances are the physical stance, the design stance and the intentional stance (Dennett, 1987).

When we want to predict the observable behaviour of a physical system, say the flow of water from the Athabasca glacier in the Rocky Mountains to the Mackenzie Delta in the Northwest Territories, its constitution (perhaps even to the microphysical level) and the physical impingements on it, we would use our knowledge of physics to predict it (Dennett 1987, 16). This strategy is the most detailed, and is always in principle available, though not always practical (16). If we ascend to the next level of pattern recognition, we adopt the *design stance*.

When attempting to interpret behaviour from the design stance, one ignores details of the physical description (e.g., what something is made of) and concentrates instead on what it is designed to do, making predictions based on this assumption. For example, I can predict how to work my VCR by assuming that it was designed to accept and play a certain size of video tape, that buttons labelled with arrows pointing in different directions are meant to tell me in which direction the tape will run, etc. (17) If I have trouble, I can descend a level in the design stance and refer to the manual, but I do not have to descend all the way to the physical stance to figure out how my VCR should behave unless there is an electronic or physical malfunction. The only behaviour that I will be able to predict from this level, however, is the behaviour for which my VCR has been designed. This strategy is not limited to human artifacts; it works well on anything that has the appearance of good design. "Not just artifacts but also many biological objects (plants and animals, kidneys and hearts, stamens and pistils) behave in ways that can be predicted from the design stance."(17)

If this stance is inaccessible, for some reason, then one can ascend to the third stance, the *intentional stance*. When predicting what an agent will do from the intentional stance, one treats an agent as if it had beliefs and goals that correspond to its place in the world and one presumes that the agent will act rationally to further those goals based on the beliefs attributed to it (17). This description should by now sound familiar because it is exactly the strategy Dawkins uses to interpret what genes "want" and how we should expect animals to behave. What is of particular interest about the intentional stance is that it will work well predicting the behaviour of things that we would not normally say are rational, believing agents.

> The strategy works on birds, and on fish, and on reptiles, and on insects and spiders, and even on such lowly and unenterprising creatures as clams (once a clam *believes* that there is a danger about, it will not relax its grip on its closed shell until it is *convinced* that the danger has passed). It also works

on some artifacts: the chess-playing computer will not take your knight because it *knows* that there is a line of ensuing play that would lead to losing its rook, and it does not *want* that to happen. More modestly, the thermostat will turn off the boiler as soon as it comes to *believe* the room has reached the desired temperature.(Dennett 1987, 22 my emphasis)

What is it about this stance that makes it so useful for predicting behaviour? Asking this question invites answers in terms of true and false descriptions of phenomena and in terms of descriptions that correspond in some direct way to things in the world. If, instead, we ask about the kinds of things whose behaviour we are talking about when this method of prediction works, then we discern a pattern: they are intentional systems whose behaviours are predictable by attributing beliefs and desires to them (1987, 242).

7.9 Interpretation and Explanation

The three interpretive stances above are three patterns on the basis of which we might make predictions. Explanation in terms of these stances has much in common with textual interpretation. Adopting the design stance to explain what the function or purpose of something is what Dennett calls "artifact hermeneutics."(Dennett 1990) As the name suggests, the methods and approaches that we use for deciding what something "means" or what it is "for" are identical with those employed in making sense of texts. In the case of biological artifacts, that a particular trait is an adaptation is a plausible assumption of interpretation generally, but not infallible. It is a good first guess.

> "The adaptationist strategy in biology seeks to answer "why?" questions in exactly the same way the intentional strategy in psychology does. Why, asks the folk psychologist, did John decline the invitation to the party? The presumption is that there is a (good) reason..."(1990, 187)

Likewise, when biologists come to explain why the giraffe has a long neck, or why ants are divided into castes, they start with the assumption that there is a good reason, and construct hypothetical scenarios to figure out what might have been the case in the organism's history that would explain the selection for this particular trait. However, these adaptationist stories are not immune to testing and criticism, their predictions are subject to falsification (188). On my instrumentalist view, this just amounts to saying that these stories prove to be empirically inadequate. Adaptationism gives us a ready rationale for design that is not dependent on the origin of the trait, an aspect of the trait that is in most cases unavailable (188). So, Dawkins can make use of the Prisoner's Dilemma to explain why we find certain behaviours in nature by virtue of the fact that they can be predicted. Using this method we cannot find out the proportions of different kinds of behaviours that might have been originally present in the population. And we do not have to know the mechanisms by which behaviours are transmitted from generation to generation to make successful predictions.

The attribution of human cognitive states to the products of evolutionary processes means that we attribute to the process of natural selection the ability to provide reasons for designing its products in certain ways (187). The assumption that something is well designed, that it is optimal for what it does, allows us to begin our interpretation; it does not mean that that is where our interpretation will end. Why is it that these higher levels of description, so successful in predicting human behaviour, work where we are sure there are no such reasons guiding design and behaviour? Dennett's answer is that the long process of selection over evolutionary time has produced the illusion of intelligent design because all the "stupid" moves have been tried, leaving only the "smart" moves for us to see. "All we see is the unbroken string of triumphs."(189) Creating explanations for why those things are triumphs involves uncovering the reasons embodied by the success of organisms with those traits, that is, interpreting them (190). To imagine these reasons, we jettison the low-level strategies in favour of a higher-level one in an attempt to generate testable hypotheses and to gain greater predictive power or generality at the cost of giving up detail (189).

With respect to different patterns or interpretations, Dennett suggests that it may be in principle impossible to say that one is better than another. Even if one has greater local explanatory power, there may be no pattern to successful predictions and falsifications that provide sufficient ground for saying one pattern is better than another (1991, 48). "There could be two interpretation schemes that were reliable and compact predictors over the long run, but that nevertheless disagreed on crucial cases."(48) What makes Dennett's view instrumentalist is his appeal to scientific utility, making more successful predictions makes a pattern better than its competitor. What makes his view a realist one is his contention that patterns at all three levels of description are "there" just because they can help us to make predictions. "There is no such state as quasi existence."(1991, 27)

7.10 A Modified View of Interaction Metaphor

Dennett's description of patterns allows us to wed our recognition and use of different metaphors (pattern recognition devices) with their utility (providing predictive power). The focus of the metaphor is a pattern that we have used to make successful predictions in the past and may be mundane (vehicle) or specialized (Prisoner's Dilemma, information system) and that has its source at the level of observable phenomena. Dennett's account also provides a possible response to Boyd's challenge to account for predictive success. Patterns are, as Dennett noted, by definition candidates for pattern recognition. So, the patterns revealed by metaphors are "there" whether anyone notices them or not, and in that sense they are independent of us and real. Accepting, as Kuhn did, that predictive success requires making contact with the world gives our instrumentalist view a realist tinge. But our metaphors say nothing in terms of truth or falsity about how the world is and say only that, in a broad sense, the world is how all the patterns attributed to it "fit". The fitting however, remains a feature of the observable world, a frequent feature of the focus of metaphor.

So, what does the interaction view look like if we adopt the view that metaphors of this kind are patterns that enable prediction? First, the implication complex of the interaction metaphor selects and highlights the features of one system or pattern to make it fit the other. Selecting and highlighting different features may make using this pattern for prediction easier, or more "computationally tractable," to borrow a phrase of Dennett's. Some patterns are better than others at prediction, but there are several dimensions across which they might differ. So, talking about one pattern as being closer to the way the world is doesn't make sense, but talking about one pattern being more adequate or suitable to the task at hand does. Viewing metaphors as predictive instruments is consonant with the view that theories are instruments for prediction and both of these views fit well with instrumentalism, as does the rejection of truth talk about metaphors discussed in Chapter Four. The "as-if" component implicit in instrumentalist theories is not a weakness of the view, as some have argued, but a recognition of the fact that by framing things in as-if terms scientific theories help us to make predictions and to give causal explanations.

One difficulty of holding onto an instrumentalist stance, however, is that it does not allow for the causal element of explanation. Attributing causal relations between things brings with it a commitment to the existence not only of the causes, but of the things between which there are causes. The pattern account pushes us in the direction of accepting the existence of such causal accounts because these accounts, too, are patterns revealed by metaphor. As we have seen, the ability to create causal stories is a conspicuous and essential feature of biological theorizing.

I depart from Dennett's description of patterns in my conclusion that this view espouses a form of pluralism. Dennett claims that physical level interpretations are the most basic, always available, and presumably the "real," though sometimes unwieldy, descriptions of the way things are. He also rejects the idea that patterns will agree point by point and so result in competing descriptions of the same phenomenon (1991, 46). I think that if we take Dennett at his word with respect to patterns there are going to be many competing physical level descriptions. If Dennett wants to say that there is only one detailed description, a "bit map" description, I do not see why a thoroughly realist view would not follow. Even so, such a description, by definition, would not display a pattern (it would be indistinguishable from patternlessness) and so could not be used for prediction. (A similar argument for the necessary presence of metaphor in scientific theorizing is offered by Mary Hesse in her *Models and Analogies in Science* (1966).) Further, in describing his lowest level of description, Dennett has resorted to a metaphor, the bit map, and in doing so has described a pattern of description even at this lowest level. The pattern he has discerned is a way of describing the lowest level of description that is drawn from one domain, computers, and that applies to a wide range of domains.

On my view, disagreements over interpretations of the same data at the same level - for example, the units of selection - will only be decided by the ability of competing metaphors to predict observable phenomena. If the debate is not joined to empirical work, then the status of one pattern being better than the other will not be decided. Of course, we could debate which one we think is going to be better on the basis of its consistency with all the facts we have, the possibility that it might have predicted past cases, etc.

The metaphors we use, the patterns we invoke, are tested against our experience, and it may be that one pattern is less adequate than another to capture all of the detail that we find and to make predictions. This is one reason why the pattern view may entice us to drop instrumentalism in favour of realism: In our attempt to use the metaphor-patterns we might find aspects of the focus that do not fit the frame and vice versa. This recalcitrance forces us to cast about, either reflecting on the frame and searching for a new metaphor, or inverting our investigation and exploring the metaphorical focus in light of the frame. But the decision to use one metaphor rather than another, to choose one pattern over another, is an instrumental one; it is a function of its predictive power and hypothesis generating capacity. Viewing interaction metaphors as pattern finders accords with Black's belief that interaction metaphors "embody insight expressible in no other fashion."

The pattern description also allows us to address the tension between Black's assertion that objects or relations exist prior to being revealed by metaphor and the anti-realist assertion that theoretical entities are useful fictions. The predictive success of these fictions still calls out for an answer and the one that I have given is that, in so far as these objects and relations constitute patterns for recognition, they must, by virtue of that fact, already be there in order to be candidates for recognition. Causal relationships are 'there' too, as patterns in the world. Accepting this fact does not commit us to saying that our metaphors help us to get closer to the real joints of nature, or that metaphors are true or false. In some sense, all these patterns would be true because all recognized patterns "fit" the world.

In his attempt to make the basic properties of the process of evolution and the development of complex adaptations salient, Dawkins appeals to metaphors of replication, information transmission, rational decision-making and other human cognitive states. But they are not the only ones; they are not privileged metaphors that have finally described how the biological world is. Adjudicating between metaphors used in biological explanation will be done on the same basis as any concepts in science, namely, their success in prediction and hypothesis generation. Their success at doing these things will help to reinforce our view that these metaphors are good ones for understanding the world.

The patterns that we find exist and the patterns suggest causal relationships. The existence claim forced upon us by the pattern view implies that, as the interaction view requires, the similarities revealed by metaphor are already there and not only our creation. That causal relationships are seen in patterns turns predictive strategies into causal accounts, stories about why some biological phenomenon appears the way it does or behaves the way it does. This existence claim is in tension with the instrumental account of theories and by implication metaphors. Now it appears that we ought to accept that theories are not merely predictive instruments. By the use of metaphor, for example, through the use of teleological language in biological accounts, we can provide causal stories that help us to make predictions, and to the extent that they give a causal account, they are also explanatory. We can predict more or less successfully because our predictions are attached to the world. By taking the challenges to the instrumentalist account seriously, and by carefully considering the way in which metaphor functions in biological explanation, we have been led, just a short way, toward a kind of realism.

7.11 Problems and Puzzles

One possible objection to the view presented above is that the antirealist conclusion I have drawn does not follow. It would seem that, in principle, nothing bars us from finding a pattern that captures all the complexity of some phenomenon, the process of selection or inheritance for example. I agree that in principle such a pattern might be found, but I think that the criterion of predictive success fosters a skeptical attitude about considering any pattern as final and unrevisable. Prediction is directed toward the future and we might never know what new cases might bring evidence that might cause us to revise the existing pattern or to search for a new one entirely. This attitude is just the sort of healthy skepticism that we expect good scientific work to embody.

Another objection, one not explicitly dealt with by Black, Boyd or Kuhn, is that successful metaphors are not dependent on actual properties of things. (Recall Searle's example "Sally is a dragon.") If metaphors in biology make use of such imaginary subjects there is the implication that either imaginary objects exist in some way because the patterns they embody are real, or that these patterns are unreal, a violation of the fact that patterns are by definition candidates for pattern recognition and independent of any particular cognizer. Though this is worrying, my answer is that these created patterns are amalgamations of other "real" patterns and have low if any predictive leverage. For this reason, such patterns are not ones that have been used successfully in prediction.

So, are the entities and processes that Dawkins' discusses really in nature or are they only metaphorical? The as-if attributions that Dawkins makes use of are not there, that is, cognitive attributes are not in genes, but the patterns of behaviour that make the as-if talk successful are. My basis for saying this, however, is just that we might predict behaviours more successfully if we do not view genes in terms of the implication complex given by cognitive agents. Cognitive metaphors can mislead us by making salient patterns that are easy to use that may not be as successfully predictive as other metaphors. Over sixty years ago, Arturo Rosenblueth and Norbert Wiener (1943; 1951) articulated one pattern underlying our attribution of teleological and cognitive capacities to non-cognitive systems. They recognized that many behaviours can be described very well as purposeful and goal directed, not as-if, but according to a pattern that they discerned in human and animal behaviour as well as in some machines and that singles out a class of predictive behaviour (1943, 22). Patterns of purposeful and non-purposeful behavior and the talk associated with the recognition of these patterns is, they thought, not only useful but essential to particular kinds of scientific investigation (1951, 321). "We believe that the only fruitful methods for the study of human and animal behavior are the methods applicable to the behavior of mechanical objects as well." (326)

What about replication? My thought here is that there may in fact be some kind of "bit map" representation of what is involved in such a process, but, if we want to make predictions and extend this model into unknown territory a metaphor will be essential for doing this. So, at least in Dawkins' case, the process he is talking about is metaphorical. He begins with the idea of a copying process, and then, by comparing different kinds of copying processes, he chooses one with which to see the processes of inheritance, selection and development. As we saw, however, because it is a metaphor, there are other metaphors we might invoke and different patterns that we might recognize.

This brings me to the final problem for the instrumentalist, one often touted as the stake-in-the-heart, namely, the difficulty of distinguishing between observational and non-observational entities. This is, I think, the most serious problem for the provisional view I have adopted. Though I have no solution to this problem, I will make two remarks.

First, it may be possible to modify the instrumentalist notion of observability to incorporate a measure of directness between our unaided sensory experience and that made possible by instruments. Stethoscopes, and microscopes and telescopes that use lenses, are instruments that serve to enhance our senses. Other instruments may be less direct in this sense and the products of these instruments might count as unobservable. Electron microscopes point to the trickiness in making a clear distinction based on directness because their products are in principle observable, but they achieve this in an indirect manner. Also, the construction and use of such a microscope requires the use of a number of theories that appeal to entities that, on the instrumentalist account, are closer to unobservables than the products of the microscope might be.

Second, the problem of making the observable/unobservable distinction, though difficult, is closely related to one that I pointed to in Chapter Four, but did not address – the distinction between metaphorical and non-metaphorical descriptions. There, I said only that we might make a prima facie distinction between metaphorical and non-metaphorical descriptions. I do not have a clear solution for this problem either, but I think that it is just as much a problem for the realist. The realist must adopt a very clear distinction between the metaphorical and non-metaphorical uses of language, provide an account of what it means for a metaphor to be fully explicated, and tell us how we might recognize when this has been achieved. As I have tried to argue, the substitution and comparison views are inadequate as accounts of metaphor and, even someone as adept at employing metaphor as Dawkins cannot clearly separate metaphorical language from non-metaphorical language. In both the observational/non-observational and the metaphorical/non-metaphorical cases we can usually make a prima facie distinction, but when pressed our divisions become more problematic. My point, however, is that if the observational/theoretical distinction counts against the instrumentalist, the related metaphorical/non-metaphorical distinction counts equally against the realist.

7.12 Conclusion

The overriding thought behind this account of the function of metaphor in explanations in evolutionary biology has been that metaphor provides us with a way of discerning real patterns in nature. By making salient real patterns, it allows us to make successful predictions. It helps us to predict what we can expect to see and find in the biological world and the sorts of changes we might observe.

In developing this view, instrumentalism was adopted as a provisional philosophy of science. It is commonly rejected because of the difficulty of distinguishing clearly between observational and non-observational entities and in explaining predictive success. Describing metaphor in terms of patterns went some way to giving an account of this success. The problem of the observational/non-observational distinction has not been resolved, but I have argued that if this is the main criticism of instrumentalism, the realist is beset with no less a difficulty in clearly distinguishing between metaphorical and non-metaphorical language and the conditions under which theory-constitutive metaphors might be said to be fully explicated. The difficulties inherent in drawing these lines are illustrated by Richard Dawkins gene's-eye view explanations. Prima facie metaphorical descriptions, such as attributing cognitive states to genes and non-human animals, provided predictive leverage at the level of phenotypic traits. Other putative literal descriptions such as the genetic replicator could plausibly be replaced by other metaphors and the limits placed on what is included in replication seemed either artificially constrained or to misconstrue biological facts.

Instrumentalism was an appealing provisional stance because it allowed for a symmetry between the way we talk about theories and metaphors –we traded evaluation of either in terms of truth or falsity for evaluation in terms of applicability or inapplicability, tractability or intractability. And it is here that the realist has another problem because he wants to make truth claims about theories. As we have seen, truth talk is absent in the most plausible account of metaphor - the interaction account. Adopting the view that metaphors are patterns required that we modify our instrumentalist account and accept the fact that patterns and similarities are "already there" prior to the construction of our metaphors. Accepting this implication has also allowed us to accommodate the fact that metaphors play a crucial role in our creation of causal stories in evolutionary biology and they are not simply predictive instruments.

Future investigations of metaphor might delve into just what is involved in our recognition of patterns. For example, one might explore the common sense idea that patterns are a function of perceived similarities and examine the nature of our capacity for recognizing similarities and kinds. A sizeable amount of work has been done in this area in philosophy, and there is an increasing amount of work being done here in cognitive and developmental psychology, so any investigation into our notions of similarity and kinds will most likely involve a detailed examination of the results in these three disciplines. Although I have tried to argue that the connection between metaphor and induction is made strong enough by an instrumentalist view of science, it may not be robust enough for people who hanker after a deeper connection between predictive power, similarity and truth about the world. A strong disagreement over the limits that instrumentalism places on explanations may ultimately make even this modified instrumentalism unappealing. Adopting realism, however, requires confronting all the difficulties with metaphor that I have raised in this thesis.

The fact that metaphorical talk in science can be unsettling for realists needs some sort of response. If realists are bothered by Dawkins' descriptions of selfish genes and programmed behaviours that, in some ways, can't be helped. As David Hull has noted in a related context,

> Many of the things which scientists say when they are reworking the foundations of their discipline sound peculiar, as peculiar as subatomic particles being able to move from one place to another without traversing the distance between or space being curved, but that cannot be helped. Inherent in the scientific enterprise is the need to go beyond ordinary usage and common conceptions. (1981, 25)

What I hope my investigation has helped to make clear is that there may be many cases in which we continue to change our metaphors in our attempt to construct different causal stories in evolutionary biology to achieve better predictive capabilities. Further, the important role that metaphor plays and our inability to draw a clear line between metaphorical and non-metaphorical descriptions suggests that biological explanations do not converge toward one stable description of how things are. "[N]o single system will ever explain the world in all its aspects and detail. The scientific approach has helped to destroy the idea of an intangible and eternal truth."(Jacob 1973, x)

Literature Cited

Allen C., Bekoff M. and G. Lauder. 1998. *Nature's purposes: analysis of function and design in biology*. Cambridge, MA: MIT Press.

Aristotle. 1941. *The basic works of Aristotle.* Edited by Richard McKeon. New York: Random House.

Axelrod, R. and W. Hamilton. 1981. The evolution of cooperation. *Science* 211: 1390-6.

Ayala, F.J. 1995. The distinctness of biology. In *Laws of nature: essays on the philosophical, scientific and historical dimension,* edited by F. Wienert. New York: Walter de Gruyter.

Bird, A. 1998. *Philosophy of science.* Montreal: McGill-Queen's University Press.

Black, M. 1955. Metaphor. *Proceedings of the Aristotelian Society* 55: 273-294.

------. 1962. Models and archetypes. In *Models and metaphors,* Max Black. Ithaca, NY: Cornell University Press.

------. 1993. More about metaphor. In *Metaphor and Thought*. 2nd ed., edited by A. Ortony. Cambridge: Cambridge University Press.

Boyd, R. 1993. Metaphor and theory change: What is 'metaphor' a metaphor for? In *Metaphor and thought*, 2nd ed., edited by A. Ortony. New York: Cambridge University Press.

Braithwaite, R.B. 1953. *Scientific explanations: a study of the function of theory, probability and law in science.* Cambridge: Cambridge University Press.

Cranor, C.F. 1994. Genetic causation. In *Are genes us? The social consequences of the new genetics,* edited by C.F. Cranor. Rutgers University Press.

Brooks, C. and R. Penn Warren. 1950. *Fundamentals of good writing: handbook of modern rhetoric.* New York: Harcourt, Brace.

Davidson, D. 1978. What metaphors mean. *Critical Inquiry* 5 (1): 31-47. Reprinted in *Philosophical perspectives on metaphor*.1981. Edited by Mark Johnson. Minneapolis, MN: University of Minnesota Press, 200-220.

Darwin, C. [1859] 1968. *The origin of species.* Toronto: Penguin Books.

Dawkins, R. 1982. Replicators and vehicles. In *Current problems in sociobiology*, edited by the King's College Sociobiology Group. New York: Cambridge University Press.

------. 1986. *The blind watchmaker*. Toronto: Penguin Books.

------. 1989. *The selfish gene.* 2nd ed. New York: Oxford University Press. Originally published in 1976.

------. 1999. *The extended phenotype.* 2nd ed. New York: Oxford University Press. Originally published in 1982.

Dennett, D.C. 1987. *The intentional stance*. Cambridge, MA: MIT Press.

------. 1991. Real patterns. *Journal of Philosophy* 88(1): 27-51.

Desmond, A. and J. Moore. 1991. *Darwin: The life of a tormented evolutionist.* New York: W.W. Norton and Company.

Dobzhansky, T. 1951. *Genetics and the origin of species.* 3rd ed. New York: Columbia University Press.

Feldman, M.W. 1992. Heritability: Some theoretical ambiguities. In *Keywords in evolutionary biology*, edited by E.F. Keller and E.A. Lloyd.

Fine, A. 1986. *The shaky game: Einstein, realism and the quantum theory.*

Fogle, T. 1990. Are genes units of inheritance? *Biology and Philosophy* 5: 349-371.

Futuyma, D. 1986. *Evolutionary biology.* 2nd ed. Sunderland, MA: Sinauer Associates, Inc.

Gallie, W.B. 1956. Essentially contested concepts. *Proceedings of the Aristotelian Society* 56: 167-198. Reprinted in The importance of language. 1962. Edited by Max Black. New Jersey: Prentice-Hall Inc., 121-146.

------. 1968. *Philosophy and the historical understanding.* 2nd ed. New York: Schocken Books.

Genter, D. and M. Jeziorski. 1993. The shift from metaphor to analogy in western science. In *Metaphor and Thought*, 2nd ed., edited by A. Ortony. Cambridge: Cambridge University Press.

Gibson, J. 1977. Review of *The selfish gene* by Richard Dawkins. *Aggressive Behavior* 3: 393-394.

Gould, S. J. 2002. *The structure of evolutionary theory.* Cambridge: Belknap Press.

Gould, S.J., and R. Lewontin. 1993. The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist paradigm. In *Conceptual issues in evolutionary biology*, 2nd ed., edited by E. Sober. MIT Press, 73-90. Originally published in 1978.

Grant, C. 1991. The gregarious metaphor of the selfish gene. *Religious Studies* 27: 431-450.

Green, T.F. 1993. Learning without metaphor. In *Metaphor and thought*, 2nd ed., edited by A. Ortony. Cambridge: Cambridge University Press.

Griffiths, P. 2001. Genetic information: a metaphor in search of a theory. *Philosophy of Science* 68: 394-412.

Gross, A. 1990. *The Rhetoric of science.* Cambridge, MA: Harvard University Press.

Hamilton, W. 1964. The genetical evolution of social behaviour I and II. *Journal of Theoretical Biology* 7: 1-52.

Hanson, N.R. 1958. *Patterns of discovery.* Cambridge, MA: Cambridge University Press.

Hartl, D.L. and E.W. Jones. 1999. Essential genetics. 2nd ed. Toronto: Jones and Bartlett Publishers.

Herbert, S. 1971. *Darwin, Malthus and selection*. Journal of the History of Biology 4(1): 209-217.

Hesse, M. 1966 *Models and analogy in science.* Notre Dame, IN: University of Notre Dame Press.

Hull, D. 1981. Units of evolution: A metaphysical essay in *The Philosophy of evolution,* edited by U. J. Jensen and R. Harre. The Harvester Press.

Jablonka, E. and M. Lamb. Forthcoming. *Evolution in four dimensions.*

Jacob, F. 1973. *The Logic of life*. Princeton University Press.

Johnson, M. ed. 1981. *Philosophical perspectives on metaphor*. Minneapolis, MN: University of Minnesota Press.

Johnson, M. and G. Lakoff. 1980. *Metaphors we live by*. Chicago: University of Chicago.

Keller, E.F. 1992. *Secrets of life, secrets of death: Essays on language, gender and science*. New York: Routledge.

------. 2000a. *The century of the gene.* Harvard University Press.

------. 2000b. Decoding the genetic program: Or, some circular logic in the logic of circularity. In *The Concept of the gene in development and evolution: Historical and epistemological perspectives*, edited by P. Beurton and R. Falk. Cambridge, MA: Cambridge University Press.

Keller, E.F. and E. Lloyd eds. 1992. *Keywords in evolutionary biology.* Cambridge, MA: Harvard University Press.

Keller, E.F. and H.E. Longino eds. 1996. *Feminism and science.* New York: Oxford University Press.

Kosso, P. 1992. *Reading the book of nature: An introduction to the philosophy of science.* Cambridge: Cambridge University Press.

Kourany, J. 2000. A successor to the realism/antirealism question. *Philosophy of Science* 67 (supplement): S87-S101.

Kuhn, T. 1993. Metaphor in science. In *Metaphor and thought*, 2nd ed. A. Ortony ed. New York: Cambridge University Press, 533-542.. Originally published 1962.

------. [1962]1996. *The structure of scientific revolutions.* 3rd ed. Chicago: University of Chicago.

Lennox, J.G. 1992. Philosophy of biology. In *Introduction to the philosophy of science: A text by the members of the department of the history and philosophy of science of the University of Pittsburgh*, by M.H. Salmon, J. Earman, C. Glymour, et al. New Jersey: Prentice Hall.

Lewontin, R. 2000. *The triple helix.* Cambridge, MA: Harvard University Press.

Lloyd, E.A. 1992. Unit of selection. In *Keywords in evolutionary biology*, edited by E.F. Keller and E.A. Lloyd. Cambridge, MA: Harvard University Press.

Marcos, A. 1997. The tension between Aristotle's theories and uses of metaphor. *Studies in the History of Philosophy* 28(1): 123-139.

Markovits, B. 2003. Review of *Motion studies*, by Rebecca Solnit. *Times Literary Supplement*, March 28: 36.

Maynard Smith, J. 1972. Game theory and the evolution of fighting. In *On Evolution*, J. Maynard Smith, 8-28. Edinburgh: Edinburgh University Press.

------. 2000a. The concept of information in biology. *Philosophy of Science* 67: 177-194.

------. 2000b. Reply to commentaries. *Philosophy of Science* 67: 214-218.

Midgely, M. 1979. Gene-juggling. *Philosophy* 54: 439-458.

Millikan, R.G. 1998. In defense of proper functions. In Allen C., Bekoff M., and Lauder, G. 1998. *Nature's purposes: Analysis of function and design in biology*. Cambridge, MA: MIT Press.

Musgrave, A. 1989. NOA's ark – Fine for realism. *Philosophical Quarterly* 39: 383-398.

Ortony, A., ed. 1993. *Metaphor and thought.* 2nd ed. Cambridge, MA: Cambridge University Press.

Oyama, S. 2000a. *Evolution's eye: A systems view of the biology-culture divide.* London: Duke University Press.

------.[1985] 2000b. *The ontogeny of information: Developmental systems and evolution.* 2nd ed. London: Duke University Press.

Pepper, S. 1942. *World hypothesis.* Berkeley: University of California.

Poundstone, W. 1992. *Prisoner's dilemma: John von Neumann, game theory and the puzzle of the bomb.* New York: Doubleday.

Ridley, M. 1996. *Evolution*. Cambridge, MA: Blackwell Science, Inc.

Richards, I.A. [1936] 1965. *The philosophy of rhetoric*. New York: Oxford University Press.

Robert. J.S., Hall, B.K. and W.M. Olsen. 2001. Bridging the gap between developmental systems theory and evolutionary developmental biology. *BioEssays* 23: 954-962.

Rosenblueth, A. and N. Wiener. 1943. Behavior, purpose and teleology. *Philosophy of Science* 10: 18-24.

-------. 1951. Purposeful and non-purposeful behaviour. *Philosophy of Science* 18: 318-332.

Rowland, M. 2004. Teleosemantics. In *A field guide to the philosophy of mind*, edited by M. Nani and M. Marraffa. http://host.uniroma3.it/progetti/kant/field/teleo.html.

Ruse, M. 1999. *Mystery of mysteries: Is evolution a social construction?* Cambridge, MA: Harvard University Press.

-------. 2000. Metaphor in evolutionary biology. *Revue International de Philosophie* (no. 214): 593-619.

Salmon, W. 1984. Scientific explanation and the causal structure of the world. Princeton, NJ: Princeton University Press.

Schaffner, K. F. 1993. *Discovery and explanation in biology and medicine.* Chicago, IL: University of Chicago Press.

Schön, D. Generative metaphor: A perspective on problem-setting in social policy. In *Metaphor and Thought*, 2nd ed., edited by A. Ortony. Cambridge: Cambridge University Press.

Sismondo, S. and N. Chrisman. 2001. Deflationary metaphysics and the natures of maps. *Philosophy of Science* 68 (supplement): S38-S49.

Sober, E. 2000. *Philosophy of biology.* Westview Press.

Sterelny, K. 2000. The 'genetic program' program: a commentary on Maynard Smith on information in biology. *Philosophy of Science* 67: 195-201.

Sterelny, K. & Griffiths, P. 1999. *Sex and death: an introduction to the philosophy of biology.* Chicago: University of Chicago Press.

Sterrett, Susan G. 2002. Darwin's analogy between artificial and natural selection: How does it go? *Studies in History and Philosophy of Biological and Biomedical Sciences* 33: 151-168.

Trivers, R. L. 1971. The evolution of reciprocal altruism. *Quarterly Review of Biology* 46: 35-57.

Trivers, R. L. 1974. Parent-offspring conflict. *American Zoologist* 14: 249-64.

Turbayne, C. 1962. *The myth of metaphor.* New Haven: Yale University Press.

Van Frassen, B. 1988. The pragmatic theory of explanation. In *Theories of explanation*, edited by Joseph C. Pitt. Oxford: Oxford University Press.

Wade, M.J. 1992. Heritability: Historical perspectives. In *Keywords in evolutionary biology*, edited by E.F. Keller and E.A. Lloyd.

Williams, G. C. 1966. Adaptation and natural selection: A critique of some current evolutionary thought. Princeton, New Jersey: Princeton University Press.

Wilson, C. 1996. Instruments and ideologies: The social construction of knowledge and its critics. *American Philosophical Quarterly* 33(2): 167-181.

Wilson, R.A. 2003. Pluralism, entwinement, and the levels of selection. *Philosophy of Science* 70: 531-552.

Wittgenstein, L. 1968. *Philosophical investigations*. Translated by G.E.M. Anscombe. Oxford: Basil Blackwell & Mott, Ltd.

Wray, K.B. 2001. Science, biases, and the threat of global pessimism. *Philosophy of Science* 68: 467-478.

Young, R.M. 1985. *Darwin's metaphor: nature's place in victorian culture*. New York: Cambridge University Press.