ARISTOTLE’S TELEOLOGY AND MODERN MECHANICS

A Dissertation

Submitted to the Graduate School

of the University of Notre Dame

in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

by

Christopher V. Mirus, B.A., M.A.

______________________________
Phillip R. Sloan, Director

______________________________
Michael J. Loux, Director

Graduate Program in History and Philosophy of Science

Notre Dame, Indiana

January 2004
This dissertation addresses teleology in the writings of Aristotle and in relation to modern mechanics. In chapters one through three, I argue that Aristotle’s teleology is theoretically grounded in the claim that both as an explanatory and as a causal factor, actuality is prior to potentiality. As actualities, therefore, both form and function are prior to the material and efficient causes that condition their occurrence in nature. In chapters four and five, I then consider the recent “systems” or “cybernetic” view of goal-directedness, along with some basic features of mechanical systems and laws more generally, in light of Aristotle’s teleology. I conclude that from an Aristotelian point of view there is no conflict between teleological and mechanical approaches to nature.
For my parents
## CONTENTS

Introduction: Teleology, Nature, and Aristotle ........................................................... 1

Chapter One: Teleology in *Generation and Corruption* and *Meteorology* ........... 20

Chapter Two: The Metaphysical Roots of Aristotle's Teleology ........................... 82

Chapter Three: Teleology and Continuous Change .............................................. 126

Chapter Four: Goal-directedness and Mechanical Systems .............................. 170

Chapter Five: Teleology and Laws ...................................................................... 209

Conclusion .............................................................................................................. 253

Appendix to Chapter One: Natural Substances in *Metaphysics* Ζ–Η .............. 258

Appendix to Chapter Four: Teleology and Backward Causation ....................... 271

References .............................................................................................................. 278
INTRODUCTION

TELEOLOGY, NATURE, AND ARISTOTLE

For the past fifty years philosophers have been trying to graft the teleological notions of function and goal-directedness onto an understanding of nature that arose largely from the rejection of teleology. The teleology rejected was that of Aristotle, and although it was not abandoned all at once, or in every respect, its absence, or at least its restriction to the biological sciences, is characteristic of what is surely the dominant trend in the history of modern natural philosophy. This trend has its roots in the mechanical philosophy of the sixteenth and seventeenth centuries. It continues and grows to the extent that wholes are explained in terms of their parts, and events in terms of their antecedent causes and conditions; whereas the parts, and the antecedent causes and conditions, are considered basic. If we appropriate the term “mechanical philosophy” for the view of nature and science that makes this sort of explanation primary, we can say that function and goal-directedness, as topics of philosophical discussion in the last half-century, are apparently teleological features of the natural world—
above all the world of living things—that must somehow be accounted for without modifying the basic commitments of the mechanical philosophy.

In the last two chapters of this dissertation, we shall consider the following question: Does the modern science of mechanics provide us with any reason to adopt the mechanical philosophy? Otherwise put, does the success of classical (non-quantum) mechanics provide any reason to exclude teleology from our conception of nature, or to restrict teleological concepts and explanations to a distinctive subset of mechanical phenomena? The most plausible answer to this question, I shall argue, is that it does not. On the contrary, mechanical phenomena themselves have an unmistakably teleological dimension.

Before asking about teleology and modern science, however, we need to have clearly in mind what we mean by teleology. Chapters one through three, therefore, address a historical question: What, in its core tenets, is the Aristotelian view that the mechanical philosophers rejected? For our purposes, this question is not about how the original mechanical philosophers understood Aristotle. This is an interesting and important issue, but if our goal is philosophical rather than historical, it is presumably more important to know what Aristotle himself thought, rather than how his late scholastic proponents interpreted him. My suggestion in chapters four and five, therefore, is not that the mechanical
philosophers were wrong to reject Aristotle as they understood him. Rather, I shall argue that when Aristotle’s teleology and modern mechanics are properly understood, there is no tension between the two. Even further, the proper analysis of basic mechanical phenomena reveals a teleological dimension that can be understood in Aristotelian terms.

The possibility of a reconciliation between teleological and mechanical approaches to nature is interesting for a number of reasons. In a moment, I shall mention some recent developments in science, primarily in biology, that may well tend in this direction. Apart from particular scientific questions, however, each of these two approaches tends to be associated with a far-ranging family of ideas concerning the natural world, human life and its meaning, and the existence of God. In public discourse, moreover, and to some extent also in the academy, the relative strength and weakness of these associations are not always carefully considered, and what ought to be a calm reflection on the causal structure of natural phenomena ends up as something quite different. One need only think, for example, of the extent to which thinkers like Richard Dawkins and Philip Johnson have gone to obscure the obvious purposefulness of living things, on the one hand, and their obvious connection with non-living nature, on the other. Though usually not quite so fierce, such battles go back at least to the
seventeenth century and the disassembly of the Aristotelian cosmos. It seems difficult to take nature as it comes, and to worry about the consequences later.

My claim that mechanistic and teleological approaches to nature are compatible is, we shall later see, more modest than it first seems. Precisely for this reason, however, it calls into question the immediate ideological significance of the whole question. At the least, it destroys any obvious connection between mechanism and a blind, autonomous nature, on the one hand, or between teleology and anti-naturalistic theories of design, on the other. In doing so, it leaves us with a teleology inseparable from mechanics, and therefore with a mechanics that shares the inherent riddle of teleology. It is not that mechanics becomes more puzzling than it has been in the past; rather, as we shall see, a teleological viewpoint simply highlights the puzzle inherent in the notion of a law of nature.

In the rest of this introduction, I shall first briefly review, in section one, the place of Aristotle’s ideas in the development of modern science, and suggest that recent developments call for a renewed study of his teleology. Then, section two will provide an overview of the recent “biological” interpretation of Aristotle’s teleology, with which I shall take issue in the following chapters. Section three concludes the introduction with an overview of the argument developed in the following chapters.
1. Aristotle’s teleology and the contemporary study of nature

In the sixteenth and seventeenth centuries, both the stagnation of Aristotelian philosophy and the rediscovery of other ancient approaches to the natural world contributed to a new conceptual and experimental search for adequate causal explanations of natural phenomena. The long-term success of the “mechanical philosophy” to which this search gave rise encouraged practitioners of natural philosophy—or at least those not engaged in the study of living things—to replace Aristotle’s abstract analysis of causal explanation with a simpler, more intuitive discussion of bodies and their powers. The simultaneous mathematization of science, moreover, and its increasing separation from pursuits that retained the title of philosophy, gradually allowed the study of inanimate nature to progress with less explicit reflection among scientists themselves concerning the nature and varieties of causality.

By the time science had fully achieved its institutional and conceptual independence, it was generally agreed, again outside biology, that of Aristotle’s four types of cause, the mature study of nature would retain only two. The agent, or in Aristotle’s careful phrase “that whence the beginning of a change,” now reserved for itself the title of cause. Matter, after suitable conceptual transformations, was now taken for granted in its pre- and post-Aristotelian role as that
which exists primarily, the modifications of which are the phenomena of nature. Form, Aristotle’s primary substance, quietly disappeared—at least as such. That for the sake of which and the end, however, which for Aristotle was the cause of causes and the key to the study of nature, was often retained in the stocks and pilloried, a much-valued negative example of the progress and fruits of scientific reason.

I shall not be arguing that early modern natural philosophers were wrong to reject early modern Aristotelianism. Indeed, it would be naïve to protest too much on Aristotle’s own behalf. The success of physics, chemistry, and their offspring since his cosmology was dissolved and his relatively a priori approach to the physical sciences replaced speaks for itself.¹ Nevertheless, as I have hinted already, Aristotle’s fate in early modern science is not all of a piece. Long after he had faded into the history of the physical sciences, his biological writings provided inspiration and guidance for those who undertook the more difficult study of living things. In the eighteenth century, Aristotelian ideas appear in the work of naturalists as diverse as Linnaeus, Buffon, Cuvier and Wolff. Even after the

¹ I do not mean to suggest that Aristotle’s considered methodology in the natural sciences was overly speculative. On the contrary, it is clear from both his practice and his explicit statements (cf. esp. GA II.8, 748a7–17; III.10, 760b28–33) that he preferred empirical arguments whenever he could get them. However, it is also clear that he was far from conceiving of the specifically experimental methodology that is especially required for progress in what we now call physics and chemistry.
cell and evolutionary theories of the nineteenth century provided new theoretical foundations for biology, Aristotle did not go away entirely. The early twentieth-century embryologist Hans Driesch drew explicitly on Aristotle’s concept of form, and it is no accident that Driesch’s contemporary D’Arcy Thompson published a translation of the *History of Animals* just seven years before the appearance of his most important biological work.\(^2\)

Despite his remarkable staying power, Aristotle’s direct contribution to contemporary biology is not increasing. Paradoxically, this is due in part to the fact that in the second half of the twentieth century a number of biologists and philosophers of biology began to question the reductive aspirations of their immediate forbears, and to suggest that unlike the sciences of matter, biology may be irreducibly teleological—or at least that at the level of biological practice, functional explanation will probably never go away. In analytic philosophy, discussions of goal-directedness and function began around mid-century, and the analysis of functional explanation has continued off and on since then. In biology itself, the resurgence of developmental biology in the 1980’s as a bridge between evolutionary theory and genetics has raised theoretical issues that were central to Aristotle’s account of life, and by no means unrelated to his teleology. The

\(^2\) Driesch 1914; Thompson 1910, 1917.
resulting new discipline of evolutionary-developmental biology, flanked on one side by the exploration in physics of self-organizing complex systems, and on the other by developmental systems theory in the philosophy of biology, has made the role of teleological notions in biology more interesting than it has been for over a century. In short, it seems that Aristotle himself is becoming less directly relevant to biology because—to take an Aristotelian point of view—biology is doing quite well without him.

The apparently increasing relevance of Aristotelian ideas in biology, however, only intensifies the problem with our understanding of nature as a whole to which I have already referred. Can we give an adequate account of teleological phenomena in biology, while holding on to the antipathy for final causes that characterizes our approach to nature as a whole? In chapters four and five I shall raise this question directly in relation to mechanics; for now, it will be easier to anticipate that later discussion from the point of view of biology.

The scientific developments just mentioned, and others related to them, are reminiscent of more than one Aristotelian concept. Renewed interest in development, the notion of a developmental system, theories of self-organization, and the recognition of hierarchical structures with multiple levels of explanation all suggest that the concept of form, as an explanatory category, may also be
increasing in importance. Now in light of this observation it is highly interesting that according to Aristotle, teleology cannot be separated from the concept of form. More generally—as critics such as Galileo, Bacon and Descartes were well aware—final causality does not appear in Aristotle’s natural philosophy as a module that can be detached and discarded without harm to the rest. It is, rather—and rather more so than is generally understood—the key to a network of concepts that stands or falls as a whole. The early modern rejection of final causality, accordingly, was the rejection of this whole network. Therefore, if we find that teleology has not gone away after all—that in some areas of natural science, moreover, it is getting harder rather than easier to dispense with—this fact may well have consequences for our understanding of nature as a whole.

Before moving on to the contemporary discussion of Aristotle’s teleology, a methodological note is in order. It is appropriate to ask what reason we have, aside from the historical importance of Aristotle’s teleology, to think that a lengthy discussion of his writings will bear fruit for contemporary discussions in the philosophy of science. In the end, of course, the results of such an investigation must speak for themselves. Nevertheless, there are several reasons for beginning the project. First, as an “analytic” philosopher Aristotle ranks, at least, with the best philosophical minds of our day. Second, without compromising
this analytic rigor, he achieved a synthetic view of the topics and problems of philosophy that no one has matched since the contemporary analytic project began. Third—as I shall argue—the core tenets of Aristotle’s teleology are independent of his conclusions in the various particular branches of natural science. They are certainly independent, for example, of his cosmology, and of his assumption that biological species are fixed. In short, if we want to know what are the elements and implications of a coherent teleology, we can hardly do better than to spend some time with Aristotle.

2. The contemporary study of Aristotle’s teleology

At about the time teleology became an issue in the philosophy of science, a new generation of Aristotle scholars began to pay serious attention to the details of his natural science. They did so not only as a contribution to the history of science, but even more because of their conviction that Aristotle’s philosophical works reflect his deep interest in the natural world, and therefore that works like Meteorology and Parts of Animals ought to shed light on better known works such as the Physics and the Metaphysics. Their scholarship, despite the unfortunate fragmentation that has tended to characterize Aristotle studies in recent years, has enriched the field immensely. Within this scholarship, moreover, work on
the philosophical issues in Aristotle’s biology has been, and continues to be, es-
pecially fruitful.

When we approach the literature that deals with Aristotle’s teleology,
however, we encounter a problem. Perhaps influenced by the concepts of func-
tion and goal-directedness prevalent in analytic philosophy of biology, scholars
working on Aristotle’s biology have concluded that his teleology is not, primar-
ily, a general feature of his analysis of nature. It belongs first of all, rather, to his
biology; and it appears in other areas of his science, if at all, only by extension.3
Now although this biological reading has yielded a number of insights, I do not
think it reflects Aristotle’s conception of nature. Moreover, it distracts us from
the central features of his teleology, which are not specifically biological. Before
saying more about the biological reading in general, however, I first want to re-
view three of its major variants. In addition to providing examples of the general
approach, this will allow me to note at the outset some of the contributions its
proponents have made. I shall comment, specifically, on the interpretations of

---

3 I discuss motivations for the biological reading in the introductory section of and the
appendix to chapter one.
action for the sake of an end proposed by Allan Gotthelf and Susan Sauvé Meyer, and on David Depew’s discussion of Aristotle on adaptation.⁴

According to Gotthelf, first, Aristotle’s teleology is a rejection of reductionism in biology. The central thesis of Aristotle’s anti-reductive view is that the coming to be of a living thing proceeds from an irreducible potency for form. In other words, the change by which the form of a mature animal or plant comes to be in appropriate matter cannot be understood solely in terms of the active and passive potentialities of the elements, of which both parent and offspring are made. There is, rather, a primitive or irreducible potency that bears precisely on the form produced, and so we can say that this form is that for the sake of which the change occurs. The conceptually basic teleological claim, therefore, is that the development of a living thing is for the sake of the mature organism. If development as a whole is for the sake of an end, however, then a living thing’s parts and its developmental stages also come to be for the sake of an end. And because they have come to be for an end, we can say that they exist for that end. The essence of Aristotle’s teleology is thus his rejection of a reductive understanding of living things.

Meyer’s interpretation, second, emerges against the background of Gotthelf’s and others like it. As she reads Aristotle, however, the key question is not about the reducibility of the reproductive faculty, but about the ontological ineliminability of the form that results. Aristotle’s teleology goes through, she argues, as soon as we recognize that there is such a thing as form, whether or not the form is reducible, on some plausible account of reduction, to the interactions of an organism’s parts. Meyer does not, however, contest Gotthelf’s view that Aristotle’s central teleological thesis concerns the development of living things.

Both Gotthelf’s and Meyer’s views capture important aspects of Aristotle’s understanding of living substances, aspects that Aristotle expresses in teleological terms. He does think that development proceeds from a specific, unified potentiality, and he does insist that the forms of living things cannot be explained away. As essays into Aristotle’s biology, therefore, Gotthelf’s and Meyer’s interpretations are largely successful. However, although they help us to understand the place of Aristotle’s teleology in his biology, they do not show, or even attempt to show, that the teleology is limited to or dependent on the biology. This would be a much larger project, and one that to the best of my knowledge has not been seriously attempted.5

---

5 I would like to thank James Lennox for pointing out that an interpretation of Aristotle’s teleology as it appears in his discussion of living things can be converted to a claim about his
Depew’s interpretation, unlike those of Gotthelf and Meyer, focuses on the concepts of function and adaptation. Depew argues that Aristotle’s teleology, like Darwin’s, is a version of the contemporary etiological account of function. In other words, teleological claims depend on the existence of a cycle in which parts, processes, and activities contribute to their own continued existence, through biological reproduction. The function or end of such a part, process, or activity is its contribution to the cycle, and this contribution counts as a function because it is an effect with which the very existence of its cause is bound up.

Depew’s discussion, like those of Meyer and Gotthelf, is an important contribution to our understanding of Aristotle’s biology. Aristotle does use reproduction to explain the ongoing existence of properly functioning, well-adapted animals and plants. Moreover, Depew is correct to point out that *Generation of Animals*, Aristotle’s treatise on reproduction and development, is part of his response to Empedocles’ claim that well-adapted living things occur by chance. As we shall see, however, Aristotle’s account of function in general does not depend on the notion of adaptedness, and so it is not etiological. Depew also, therefore, provides no reason to think that Aristotle’s teleology as a whole is biological.

---

teleology in general by the addition of a further premise: namely, that in Aristotle’s considered view, only living things are substances in the full sense. I sketch an initial response to this premise in the appendix to chapter one.
Let us return now to the biological reading in general. As all three of the above-mentioned commentators recognize, Aristotle’s teleology was forged in an intense debate with his philosophical predecessors. Moreover, his disagreements with the early naturalists are clearest in the fierce assault he mounts on their attempts to explain the phenomena of plants and animals reductively, or as a result of chance. However, Aristotle’s criticism of earlier theorists is much deeper and broader than the biological reading of his teleology would suggest. As *Physics* I and the opening chapter of book two make clear, at the heart of his natural philosophy is the ontology he develops to explain the most basic natural phenomenon of all: the phenomenon of ordered change, and especially the coming to be of substances. As he states repeatedly in *Metaphysics* B and elsewhere, Aristotle was convinced that his predecessors lacked the ontological resources to deal with the most basic features of the perceptible world. At best, the early naturalists could handle a limited number of relatively uninteresting changes. In the end, therefore, they left Parmenides’ riddle unanswered.

Aristotle’s teleology is already present in the basic elements of his account of change: matter, form, privation, agency. More specifically, given the basic structure of change as Aristotle understands it, the form that results from any natural change is necessarily an end; it is that for the sake of which the change
occurs. Moreover, Aristotle also believed that the causal structure of natural substances cannot be fully understood without reference to the metaphysical distinction between potentiality and actuality. In the last analysis, his natural teleology rests on this distinction and its implications, and so it applies wherever the distinction can be made.

3. Order of the discussion

The first part of the following dissertation deals almost exclusively with Aristotle’s teleology. My goal in this part is threefold. In the first chapter, I directly attack the biological reading of Aristotle’s teleology. Focusing on Aristotle’s discussions of coming to be, the elements, and the homogeneous bodies in Generation and Corruption and Meteorology, I show that Aristotle’s teleology is at work in what might seem an unlikely place: his account of sublunary, inanimate substances. Moreover, I explain the teleology’s relative lack of prominence in Generation and Corruption and Meteorology through reasons that are internal to the teleology itself and that are, in fact, available in Meteorology.

In the second chapter, I examine the metaphysical roots of Aristotle’s teleology in his concepts of potentiality and actuality. Along the way, I also show how the notion of an end in natural philosophy depends on the metaphysical concept of a limit, and how Aristotle’s conception of goodness enters into his
teleology. Finally, by showing that Aristotle’s teleology is ultimately grounded not in his natural philosophy, but in his metaphysics, this chapter clarifies how he can apply teleological concepts not only outside biology, but even apart from change for the sake of an end.

In chapter three, having already discussed the most basic and general concepts in Aristotle’s teleology, I take up the more specific concept of motion for the sake of an end. Drawing on the general account of motion in the *Physics*, and on the discussion of coming to be in *Generation and Corruption*, I show that for Aristotle every continuous motion—that is, every motion that is numerically one—is for the sake of its outcome. Because the strong concept of continuity required for this conclusion depends on the notion of a καθ’ αὑτό (per se) agent cause, in this chapter we shall also come to see why Aristotle thinks that every change takes place either by chance or for the sake of something.

In chapters four and five, we turn from the interpretation of Aristotle to contemporary philosophy of science. For reasons I explain at some length in the beginning of chapter four, I have chosen to focus on the concept of goal-directedness, and to begin with the relation between mechanical and teleological concepts in the “systems view” of goal-directedness. Thus in chapter four, after reviewing the core tenets of Aristotle’s teleology as it emerges in the previous
chapters, I present the systems view as it is found in Ernest Nagel’s classic discussion. In the rest of the chapter, drawing on more recent developments and criticisms of this view, I argue that its plausibility depends on its implicit use of Aristotelian concepts. By focusing, specifically, on the notion of a καθ’ αὑτό connection between a system’s structure and its behavior, we can see why the concept of a goal is appropriate to the study of nature. In addition, this concept turns out to have a much broader application than proponents of the systems view have envisioned.

In chapter five, finally, we shall turn to the relation between teleology and basic laws in mechanics. After reviewing some recent discussions of causal explanation, focusing on the explanatory role of general laws, I argue that the explanation of particular events in terms of a general law is in fact teleological. For example, the abstract form of motion that general relativity theory attributes to all objects in spacetime—motion along a geodesic—qualifies as an Aristotelian end. The various ways in which motion along a geodesic can be instantiated, moreover, are plausibly thought of as means. Finally, in the last part of chapter five, I suggest a flexible account of goal-directedness in general, and distinguish a number of ways in which this goal-directedness appears to be found in nature.
The topics addressed in these five chapters are, of course, only a few of the important issues surrounding Aristotle’s teleology and its present viability. In particular, I say very little in the following pages about teleological explanations in Aristotle’s biology, or about contemporary questions in biology and the philosophy of thereof. In the course of researching and writing this dissertation, however, I became convinced of two points that have determined the contours of the finished product. First, Aristotle’s teleology is much broader than his account of living substances. For this reason, I could not address the teleological dimensions of his biology, in particular, without first undertaking a much broader investigation. Second, the viability of this teleology in contemporary philosophy of science is an open question, whose investigation might bring some surprising results. The following chapters are the first results of that investigation.
CHAPTER ONE

TELEOLOGY IN GENERATION AND CORRUPTION AND METEOROLOGY

The biological reading of Aristotle’s teleology is mistaken. It nevertheless highlights some important aspects of his natural philosophy, and before attacking it we should understand its attraction. After all, one does not lightly propose an interpretation that seems to fly in the face of an author’s explicit claims; and Aristotle’s direct statements on “the end and that for the sake of which” suggest a universal or near-universal natural teleology. In Physics II.8, for example, arguing that nature acts for an end, he premises that whatever happens is “either from coincidence or for the sake of something” (199a3–4). In chapter five of the same book he remarks, without further comment: “for the sake of something are whatever might be done from thought, and whatever from nature” (196b21–2). In I.9, discussing change in general, he goes so far as to write that “there being something divine and good and desirable, we say that it has both a contrary, and

\[1\] All translations are mine. I have tried to be as literal as possible, and have occasionally consulted the Loeb and the revised Oxford translations of the various texts.
something that is such by its own nature as to desire and yearn for it ... but this is the matter” (192a16–22). Now although such claims might conceivably refer to the need for teleological explanation in biology, the words themselves do not suggest this interpretation. Why might one think, then, that Aristotle’s natural teleology applies, except perhaps in an extended sense, only to the organization, development, and activity of living substances?

One important motivation for the biological reading is an interpretation of Aristotle’s natural philosophy and metaphysics according to which, in the final analysis, only living things count as Aristotelian natural substances. Most recently, for example, James Lennox has suggested that the unity of Aristotle’s natural philosophy derives from the focal position of living things, which alone fully instantiate the concept of nature as Aristotle understands it.² If this claim about Aristotle’s understanding of nature is correct, then it is easy to see how his teleology can both play a central role in his account of nature and of natural substances—which it surely does—and apply in the first place to living substances only. Now it is impossible, in these pages, to respond adequately to this interpretation of Aristotle’s natural philosophy as a whole. It cannot, however, simply be left aside, for it is likely to have a strong influence on our interpretation of

² Lennox n.d.
particular texts, especially those we shall be considering in this first chapter. Thus, in addition to the positive account of Aristotle’s teleology found in chapters two and three, in an appendix to the present chapter I discuss the important claim that an identification between natural and living substances can be found in *Metaphysics* Z–H. Those who are inclined to accept such an identification will want to read the appendix before continuing.

Another aspect of the biological reading’s attractiveness appears if we turn from Aristotle’s general statements about teleological explanation to his particular examples, and to his explanatory practice in the various sciences. In discussing the “cause for the sake of which” in *Physics* II.5–9 and in *Parts of Animals* I.1, Aristotle draws his examples exclusively from human action and from the biological realm. The same is true of his comments on teleology in the *Metaphysics*. Moreover, in the nonbiological treatises that follow the *Physics*, teleological language is infrequent, and partly for that reason it is also usually ambiguous. The clearest teleological claims in these works concern the motion of heavenly bodies, but as Aristotle believed these bodies to be ensouled and therefore alive, their motion—goal-directed though it may be—is far from pointing to a universal teleology. In Aristotle’s scientific practice and in his examples, therefore, we
seem to find grounds for the biological reading, and so also reason to dispute the most obvious interpretation of the seemingly general statements quoted above.

Apart from this apparent textual support, proponents of the biological reading may also be influenced by the fact that on contemporary accounts of teleological phenomena, such as function and goal-directedness, it is difficult or impossible to make sense of any teleological claim that does not involve human intentions or living things. Indeed, there are strong terminological connections between the main variants of the biological reading and current philosophy of biology. For example, in each of the three interpretations that we considered in the introduction, the key concept (reduction, the elimination of form, or the etiological account of function) is one found under the same name in contemporary discussions, and in each case there is no corresponding term in Aristotle. Even without such connections, however, it would be surprising if the positions available in recent analytic literature were not partly responsible for the popularity of biological interpretations of Aristotle’s teleology.

These last two motivations for the biological reading suggest two lessons that we can take from it. First, it is certainly the case that Aristotle’s favorite examples of natural teleology are biological, and that teleological explanation is far more common in his biological works than in other areas of his natural science.
Any interpretation of his teleology, biological or not, should be able to account for these data. Second, whatever their deficiencies, the various biological interpretations and the arguments that support them do reveal some interesting connections between teleological explanation in Aristotle’s biology, and contemporary accounts of function and goal-directedness. If the former turns out to be embedded in a broader, teleological view of nature as a whole, this fact may well have implications for contemporary discussions of teleology.

In any case, the biological reading of Aristotle’s teleology depends on the claim that in practice, Aristotle restricts his teleological concepts to the study of living things, or at least that such concepts and explanations appear in other contexts only by an extension of their focal meaning. The foundation of any biological interpretation, therefore, takes the form of a universal negative: Except by relation to the study of living things, no Aristotelian science of non-living substances is teleological. If this claim stands, then there is good reason to think twice about abstract teleological statements that seem to be broader. If it falls, however, there is no reason to restrict the scope of such statements. My goal in this chapter is to head off any such restriction by establishing a clear counterexample to the biological reading’s key premise. I argue, specifically, that Aristotle applies a central concept of his teleology, “that for the sake of which,” to
inanimate bodies in the terrestrial sphere, quite apart from the role such bodies play in his biology. The texts and arguments needed to establish this claim, moreover, will also explain why that for the sake of which and related concepts seldom appear explicitly in Aristotle’s non-biological science.

The terrestrial or sublunary inanimate substances in Aristotle’s cosmos include the four elements, on the one hand, and the homogeneous or mixed bodies composed of these elements, on the other. Although teleological concepts and explanations are more prominent in Aristotle’s account of the heavens than in his discussions of the elements and of the mixed bodies, there are nevertheless several related reasons for focusing on the latter. First, as already mentioned, the fact that Aristotle considers the heavenly bodies alive makes them problematic for present purposes. Second, Aristotle’s account of the heavenly bodies is also closely linked to his cosmology, and although this cosmology is certainly teleological, the central concepts of his teleology do not in turn depend on the cosmology. Finally, focusing on sublunary rather than heavenly substances makes sense precisely because they constitute the more difficult case—the more unlikely the counterexample, the stronger the argument once it is established.

The main goal of this chapter, therefore, is negative, whereas the task of developing a positive account of Aristotle’s teleology belongs to chapters two
and three. For this reason, my discussion of the elements and the homogeneous bodies can be fairly limited. Although Aristotle discusses the elements in both *Generation and Corruption* and *De caelo*, I consider only the former work, which generally abstracts from their place in his cosmology. Likewise, although both *Meteorology* IV and *Parts of Animals* II.1–9 deal with the homogeneous bodies, I mention *Parts of Animals* only to distinguish its approach from that found in *Meteorology*. Thus, my discussion in the following pages is based almost entirely on *Generation and Corruption* and *Meteorology* IV, and involves no more analysis of the teleology found there than is necessary to establish its presence.

However, establishing the presence of teleological concepts in these two works will require making some initial points about the structure of teleological explanation as Aristotle understands it. I should make clear now, therefore, that in rejecting the restrictive, biological reading of Aristotle’s teleology, I do not intend to argue that this teleology is essentially cosmological, anthropocentric, or theological in character. In the last analysis, some or all of these dimensions may be present in Aristotle’s thought; in particular, I believe that he did favor a theological explanation for the teleological character of nature.3 Nevertheless, neither divine causality, nor cosmic order, nor human welfare enters into the basic

---

The universal structure of Aristotle’s natural teleology. Rather, the universality of this teleology lies in the fact that every natural substance in Aristotle’s cosmos acts for the sake of one or more ends determined by its own nature. It is teleology in this sense with which I am concerned in the present chapter and in those that follow.

To establish the presence of that for the sake of which (τὸ ὑὗ ἑνεκα) in Aristotle’s accounts of the elements and of the homogeneous bodies—apart, once again, from the role these bodies play in his biology— I shall consider five texts drawn respectively from Generation and Corruption I.7, II.6, and II.9, and Meteorology IV.2 and IV.12. Of these passages it is the last—the final chapter of Meteorology—that is both most interesting and most controversial. On the one hand, taken at face value this chapter provides more information than any other text concerning Aristotle’s teleological understanding of the elements and the homogeneous bodies. On the other hand, this very fact has prompted some to understand it as introducing the biological role of the homogeneous bodies that is later discussed in Parts of Animals, rather than as discussing these bodies in their own right. For this reason, although I argue separately for the presence of τὸ ὑὗ ἑνεκα

---

4 Although it might seem that homogeneous bodies such as flesh and bone are by definition parts of living substances, this is not in fact the case. Aristotle distinguishes between homogeneous or mixed bodies as such, and homogeneous parts of living bodies as such. Some of the former are the proximate matter of the latter, and go by the same name—for example, “flesh”—but such names are used homonymously for the living part and its matter. On this distinction see my 2001; earlier discussions of the same topic are Whiting 1992; Cohen 1992; and Williams 1986. The problem of the homonymous use of terms like “flesh” was posed by Ackrill (1978).
in each of the five texts just mentioned, most of what follows is also intended to lead up to my interpretation of *Meteorology* IV.12. By examining the other four passages first, I argue in effect that Aristotle’s account of the homogeneous bodies is teleological throughout; there is no reason, therefore, to impose a biological reading on *Meteorology* IV.12. In addition, before considering the text of IV.12 itself, I discuss at some length the topic and structure of *Meteorology* IV as a whole, to provide an immediate context for understanding its final chapter.

1. Teleology in *Generation and Corruption* and *Meteorology* IV.2

We begin, then, with *Generation and Corruption*. Of the three passages to be considered from this work, the clearest is an explicit statement in book two, chapter nine, that all natural coming to be occurs for the sake of an end. In fact, the passage in question is clearer on this point than any of the seemingly general claims about natural teleology in the *Physics*. It has the drawback, however—as a general claim found in one of the work’s concluding chapters—of being less clearly connected with inanimate substances in particular than are the texts from I.7 and II.6. It is with these texts, therefore, and in particular with the latter, that we shall begin.

*Generation and Corruption* I.7 is part of a long discussion covering various topics that Aristotle considers prerequisite for a discussion of the elements (see
I.6, 322b1–27). Within this discussion, it immediately follows his treatment of contact, in chapter six, and is the first of three chapters on action and being affected. These three chapters are followed in turn by a chapter on the last preliminary topic, mixing (chapter ten). The context, in short, indicates that Aristotle is concerned in chapters seven through nine primarily, if not exclusively, with the elements. Although this is especially clear in chapter eight, Aristotle indicates the unity of this chapter with chapter seven at the end of the latter: “So let what it is to act and to be affected, and to what things these belong, and on account of what, and in what manner, be defined in this way. But let us again discuss how it is possible for this to happen” (324b23–5). The discussion that follows explicitly concerns the reciprocal action and being affected of the elements and the homogeneous bodies, taking issue on this subject with various early naturalists, especially the atomists. In chapter seven itself, finally, Aristotle draws his examples primarily from the interaction of the elements (and, in order to distinguish proximate from remote agents, from the arts), as opposed to the action of plants or animals.5

5 It may be that Aristotle considers a discussion of the action and being affected proper to the elements and their mixtures an essential part of a discussion of action and being affected in general. The reason for this would be that all interaction among bodily substances depends on, without being exhausted by, that which is made possible by the active and passive qualities of bodies as such, namely, the hot, the cold, the wet, and the dry, when these bodies come in contact with each other.
The passage within chapter seven that we shall now consider occurs after Aristotle has explained that “whichever active things have the form in matter are themselves unaffected, but whichever are in matter, are such as to be affected” (324b5–6). He then continues:

The active (ποιητικόν) is a cause as that whence the beginning of the motion. But that for the sake of which (τὸ ὅου ἑνεκα) is not active, so that health is not something active, unless metaphorically: for with something acting, whenever it is present (τοῦ µὴν ποιοῦν· τος ὅταν ὑπάρχη), the patient becomes something, but when states are present (ἐξεσχα λαοῦσαν) it no longer becomes, but already is: and forms and ends are a kind of state (ἐξεις τινές) (324b14–19).

Given the last sentence just quoted, as well as the broader context discussed above, it would be hard not to conclude that in these lines, Aristotle is contrasting the immediate agent in each change with the state it brings about in the patient. Immediately above, in fact, he has referred to the agent of heating and its effect. Given his use of health as an example, however, and the fact that the action of heating may be part of an action such as making something healthy, it may be that only some forms are to count as ends, and as that for the sake of which.

On the one hand, Aristotle does not say in this passage that the state brought about by an agent is always an end; his claim is, conversely, that forms and ends are states. Thus, one might think that the passage has nothing to say about what counts as an end, or even that the use of health as an example points
toward a restricted use of “end.” On the other hand, the overall structure of the passage seems to involve a contrast between that from which a change begins and that to which it proceeds, both of which Aristotle considers causes, but only one of which is active. Read thus, it fits nicely into the chapter as a whole, which begins with an argument that agents and patients—leaving aside those agents that have no matter—are alike with respect to their matter but unlike in form, and that every agent therefore assimilates the patient to itself. Given this context, it seems arbitrary to restrict the terms “end” and “that for the sake of which” to a particular subset of the states that agents bring about. Otherwise put, if Aristotle’s basic point is that what is hot makes what is cold like itself, it makes perfect sense to say that heat in the patient is an end, that to which the agent as such is ordered or for the sake of which it acts.

Our second passage from *Generation and Corruption* occurs in II.6, where Aristotle is criticizing various aspects of Empedocles' account of nature. One difficulty he raises is that Empedocles cannot account for natural coming to be. His argument for this claim reads as follows:

The things that come to be by nature all come to be in a given way either always or for the most part, whereas things beyond what is always and for the most part are from chance and from luck. What, then, is the cause of man coming from man either always or for the part, and from wheat, wheat, rather than an olive? Or also, if things be put together (συντεθῇ) in a given way, bone?
For nothing comes to be because things come together (συνελθόντων) however it may chance, as he himself says, but through a certain proportion (λόγος). What, then, is the cause of these things? For it is surely not fire or earth. But neither is it love and strife: for the former is a cause of association only, the latter of dissociation. The cause is the substance (οὐσία) of each thing, and not only “a mixing (μίξις) and a separating of things mixed,” as he says—these are called chance, not proportion (λόγος), for they are a being-mixed as it may chance. The cause, in fact, of things that exist by nature is that they are in a certain state (τὸ οὕτως ἔχειν), and this is the nature of each, concerning which he says nothing. Therefore he says nothing “about nature.” And indeed this is both the well and the good, yet he credits the mixing alone (333b5–20).

This acerbic passage partly parallels the better known dispute with Empedocles in Physics II.8, where Aristotle’s argument based on the generation of like from like is spelled out more fully. By contrast with that earlier discussion, what is striking about the present passage is that after referring to the regular generation of men from men and of wheat from wheat, Aristotle adds a third example: the coming to be of bone when certain things are put together in a particular way. Because bone is part of a living body, one might think that in mentioning it, Aristotle is simply giving a third biological example. If so, the οὐσία in question would be either the form of the animal as a whole, or the function of bone within that whole. However, there are three reasons for rejecting this interpretation.

First, Generation and Corruption serves both as a treatise on coming to be and passing away in general, and a treatise about the elements, their
transformation, and their mixtures. Nowhere in its two books, accordingly, does Aristotle show any interest in the peculiarities of living things. When he discusses growth, for example (I.5), he treats living bodies simply as bodies, focusing on the distinction between the homogeneous and the heterogeneous. He mentions nutrition, at the end of this discussion, only to distinguish it from growth. At the end of the treatise (II.10), moreover, he mentions the perpetual coming to be of living things only as a cyclical process, on a par with the continual coming to be of the elements.

In II.6 itself, Aristotle sets out to criticize Empedocles’ theory that all bodies are composed of four elements, which are not transformed into one another, but are combined and dissociated by two generic influences, love and strife. He begins by attacking the view that the elements are equal to each other but not transformable; he then goes on to attack Empedocles’ accounts of growth, of coming to be, and of local motion. The chapter as a whole, therefore, constitutes an argument that Empedocles’ basic ontology is wholly inadequate to account for any of the basic types of natural change. The passage we are now considering is devoted to coming to be, and addresses natural coming to be as such. Thus, only if we identify natural coming to be with that of living substances and their parts can this passage plausibly be read as a discussion of living things as such.
In the context of *Generation and Corruption* as a whole, such a restriction of natural coming to be seems tendentious. It is therefore reasonable to think that Aristotle’s examples of man, olive and bone do not indicate an excursion into biology, but are mentioned as three distinct examples of substances for whose coming to be Empedocles cannot account.

A second reason to reject a biological interpretation of this passage is that throughout *Generation and Corruption* (as also in *Meteorology* IV), Aristotle treats flesh and bone simply as mixed bodies, no different for his present purposes from stone and bronze. The first major discussion of mixing, in I.10, includes few examples of mixed bodies, mentioning only the mixing of tin with bronze (328b). The second discussion, however, in II.7, repeatedly treats flesh and bone as mixed bodies that exist by a determinate proportion of the hot and the cold. It is true that in the passage on growth, he considers flesh and bone as having forms over and above their forms as homogeneous bodies (321b). However, as already stated, his interest here is in the fact that within living bodies, homogeneous solids like flesh and bone also have a determinate shape, rather than in any physiological function that they may also have.

---

6 For a more general discussion of the possible identification between natural and living substances, see the appendix to this chapter.
Finally, Aristotle’s account of the mixed bodies is borrowed in large part from Empedocles himself, who held that homogeneous bodies other than the four elements are mixtures of these elements in definite proportions. The Greek term for mixing is μίξις, and the term for proportion is λόγος. Earlier in Generation and Corruption (I.10), moreover, Aristotle has developed a highly technical sense of μίξις, having to do with the quantitatively precise action of the hot and the cold on the wet and the dry. Thus, although Aristotle is also criticizing Empedocles’ account of the coming to be of living things—through a random assembly of parts that might well be called μίξις, resulting in what Empedocles would probably have referred to as a λόγος—it would be tendentious to read his criticism in a way that excludes the more technical uses these terms acquire when used to discuss the mixed bodies.7

On the most likely interpretation of this passage, therefore, the οὐσία that is a cause of bone’s coming to be is the οὐσία of bone considered as a

7 Because Aristotle is taking advantage of Empedocles’ use of λόγος, the correct translation is almost certainly “proportion,” rather than “account” or “definition.” It should be noted that both Webster (1984) and Lee (1962) take the elements to be the understood subject not only of the passive verb συντεθη, but also of the participle συνελθόντων. If this translation is correct, then the whole sentence in which συνελθόντων probably refers to bone and other mixed bodies, and λόγος would then have to be taken to mean the proportion of the mixed elements, rather than the arrangement of any parts whatsoever. It seems to me likely that at some point, Aristotle moves back to a more general consideration of coming to be, and I am inclined to take the subject of συνελθόντων to be whatever can come together. However, the more restrictive rendering would make the passage an even more pointed attack on Empedocles, because by dwelling on the mixed bodies in particular, Aristotle would be emphasizing that chance mixings cannot account even for these, and much less for the coming to be of animals and plants.
homogeneous body, rather than either its function within a living body, or the form of that body as a whole. This interpretation, moreover, yields a very nice rendering of Aristotle’s argument as it would apply to mixed bodies. The argument goes as follows:

1. As Empedocles himself admits, a mixed body like bone comes to be from its constituents, say fire and earth, only if they are combined in just the right proportion.

2. The cause of this coming to be cannot, therefore, be the constituents themselves, which can be combined in any number of proportions.

3. Neither can Empedocles’ preferred efficient causes, love and strife, account for the coming to be of bone, because these are characterized as forces of association and dissociation in general, and are not determined to the production of bone in particular.8

4. If the material and efficient causes Empedocles recognizes did produce bone, therefore, it would be only occasionally and by chance; hence the coming to be of bone would have nothing to do with λόγος.

5. The cause of bone’s coming to be, rather, is its οὐσία, for bone comes to be regularly not when its constituents are mixed by chance, but when they are mixed in just such a proportion as to give rise to the οὐσία of bone.

6. In general, the cause of things that exist by nature is that they are in a certain state.9

---

8 This interpretation of the remark about love and strife finds support later in the chapter (333b23–334a9), where Aristotle criticizes these factors as supposed efficient causes by showing that they are too broad and indeterminate to account for the various particular motions that we observe.

9 We might paraphrase the construction τὸ οὕτως ἔχειν by saying that the matter is actualized “in just this way.”
7. This οὐσία is what we most properly call the nature of each thing.

8. It is also, finally, the good (ἀγαθόν), or good condition (τὸ εὖ), of the thing in question.

On this reading, Aristotle is criticizing Empedocles not for failing to notice a peculiarity of living things, but for failing to explain the coming to be of all natural substances in terms of the form, οὐσία or nature that results.¹⁰ In the case of mixed bodies, this form involves a λόγος in the technical sense of arithmetical proportion. As we shall see, moreover, it is precisely this neglect of form for which Aristotle faults his predecessors when he sets out to give the causes of coming to be in general, in *Generation and Corruption* II.9.

To be sure, Aristotle does not explicitly say, in the passage we have been examining, that bone comes to be for the sake of its form; nor does he say that the as yet unspecified efficient causes of its production act for the sake of that form. He does, however, clearly state that the nature or substance of mixed bodies is a cause both of their being and of their coming to be, and that it is also their good. As we shall see in chapter two, moreover, he also recognizes a general equivalence between “the good” and “that for the sake of which.”¹¹ Finally, as I have

---

¹⁰ Neither the present passage nor any in GC provides a reason to think that the terms οὐσία, φύσις, and τὸ οὕτως ἔχειν here refer to the forms, or actualities, of living things and their parts. For the view that only the forms of living things count as οὐσίαι, see the appendix.

¹¹ On this topic see also my n.d.
been arguing, there is no reason to assume that the nature, substance and good to which he refers in this passage is always that of a living substance.

Before we turn to our third passage, it is worth noting a point of contact between the present text from II.6 and the text from I.7 that we considered above. In both places, that is, Aristotle makes use of the notion of a state (ἕξις, τὸ ἔχειν). In doing so, he draws our attention to the stable actuality that results from a given change, and attributes to this actuality a causal role. In II.6, he explains that natural coming to be cannot be understood without reference to the resulting οὐσία. In I.7, considering change more generally, he contrasts the cause from which a change begins, namely the agent, with the cause for the sake of which, that is, the form or state in which the change ends. In chapter three, I shall argue at some length that Aristotle’s cause for the sake of which is, in changes that involve a process, nothing other than the stable actuality toward which the change is directed, and in terms of which it must be understood.12

For our last passage from Generation and Corruption, we turn to II.9, in which Aristotle begins his concluding discussion of “how many and what are the principles of all coming to be alike (πάσης γενέσεως ὁµοίως)” (335a26–7). This

12 In contemporary discussion, a teleological explanation is one that invokes the outcome of a given process, or the effect of a given item, to account for the process or item itself. Thus it is worth noting that if we knew nothing at all about Aristotle’s teleology in particular, we should characterize his reference to “states” in GC I.7 and II.6 as a teleological mode of explanation, regardless of the biological or nonbiological status of these states.
phrase recalls the opening sentence of the work, in which he states: “Concerning the coming to be and passing away of the things that come to be and pass away by nature, we must distinguish alike with respect to them all (ὁµοίως κατὰ πάντων) both their causes and their accounts (λόγους)” (314a1–3). In II.9, fulfilling this promise, he distinguishes three causes of coming to be in general. The first and the third, which he discusses at length, are the matter, or “that which can be and not be” (335a34), and the agent, to which it belongs “to move and to make” (335b30–1). The second cause is mentioned briefly between the other two: the cause “as that for the sake of which (ὡς δὲ τὸ οὗ ἑνεκεν),” Aristotle says, “is the figure (µορφή) and the form (εἶδος)” (335b6–7).

In II.9, therefore, we have a statement clearer than any found in the Physics that whenever something comes to be by nature, the form of the generated substance serves as a cause for the sake of which. Later in the chapter, moreover, Aristotle criticizes his predecessors for overlooking this cause. His objection is similar to the criticism of Empedocles that we have already seen in II.6: earlier writers, he says, “leave out the more controlling cause (τὴν κυριωτέραν αἰτίαν): for they remove what it is to be and the form (τὸ τί ἦν εἶναι καὶ τὴν µορφήν)” (335b34–336a1). Furthermore, just as he wrote in II.6 that the elements themselves cannot account for that which comes to be naturally from them, if we fail
to consider the οὐσία, so here he links the omission of form to the failure to specify an adequate agent: “Moreover, removing the cause according to the form, they grant bodies capacities, through which they bring things to be, that are too tool-like (ἀποδιδόασι ... λίαν ὀργανικῶς)” (336a1–3).13

In *Generation and Corruption* II.9, more so than in the previous two passages that we have considered, the burden of proof lies squarely on proponents of the biological reading. Aristotle begins the chapter by reminding us that things near the center of the cosmos come to be and pass away, and then proposes to state the principles of all such coming to be. He briefly mentions matter and form, stating that the latter is the cause as that for the sake of which. Then, without further discussion, he begins a controversy about agency with the Platonists and the early naturalists, criticizing the latter for conflating the agent of a change with the matter that undergoes it. As usual, his examples of causes for which the naturalists cannot account are living things and artifacts, to which he gives special consideration. The scope of the chapter as a whole, however, is clear enough, and so this choice of examples serves to strengthen the nonbiological reading of other such passages, rather than to weaken it in this case. Moreover, if the

13 It is worth noting that at the end of the chapter, Aristotle draws a connection between the account of matter and form he has just given and his initial discussion of the causes in *Physics* II: “We have spoken generally of the causes earlier, and now we have delimited the matter and the form (μορφή)” (336a13–14).
general scope of chapter nine were not clear enough from the text itself, it is even clearer in what follows. In the opening lines of chapter ten, Aristotle states that his discussion of matter and form is now complete, and goes on to consider the conditions that make coming to be eternal. In this and the following chapter, he considers the transformation of the elements as well as the reproduction of living things. In chapters ten and eleven, therefore, the general scope of the discussion that begins in chapter nine and continues through the end of the work is even clearer.

The one possible objection to this reading of *Generation and Corruption* II.9–11 rests on an alternate interpretation of the phrase with which Aristotle explicitly states the scope of the discussion: “all coming to be alike.” In his above-mentioned essay on Aristotle’s biology, James Lennox suggests that when Aristotle sets out to discuss the principles and causes of natural phenomena in general, we should not infer that all the causes he discusses occur in every natural phenomenon. Rather, Aristotle’s approach is to outline the principles that will enable us to understand all natural phenomena, including the phenomena of life; and it is because living substances are teleological that he includes that for the sake of which among the principles of nature in general.
Although this view of Aristotle’s natural philosophy certainly deserves further study, I do not think it fits the texts we are now considering. In both I.1 and II.9 of *Generation and Corruption*, Aristotle uses not only the adjective πᾶς, “all,” but also the adverb ὁμοίως, which I have translated “alike.” Now although the former term seems compatible with Lennox’s view, the latter much more clearly suggests that Aristotle is undertaking a uniform treatment of all coming to be. This helps explain why, after the many digressions into various preparatory topics that characterize the work as a whole, Aristotle repeats his opening statement in II.9: we are to understand that unlike much of what preceded, which is proper to the elements and the homogeneous bodies, what follows is intended to apply generally. Moreover, in this final reintroduction of the topic, Aristotle continues thus: “For thus we shall more easily understand (θεωρήσο-μεν) the particulars, when we have first grasped the universals” (335a27–8). It would be odd indeed if the following universals turned out to apply only to living things.14

14 There is a more general response to Lennox’s view that at least deserves mention, although it cannot be developed here. The question of where and how to deal with principles of relatively general application is one that Aristotle himself explicitly raises in the context of his biology. In *Parts of Animals* I, one of the methodological issues he raises is whether it is better to give a complete, continuous treatment of each species, or to treat common characteristics first, before proceeding to more specific traits (I.1, 639a16–b6; I.4, 644a24–b8; I.5, 645b1–14, 21–8). The method he adopts is the latter, and although this choice may at first seem to be only a matter of convenience, Lennox (1987) and Allan Gotthelf (1997b) have convincingly argued that Aristotle adopts it for substantive epistemic reasons. If, he thinks, we consider common features only in
In these three passages from *Generation and Corruption*, and in the connections between them, a pattern has begun to emerge. In each case we find Aristotle appealing to three causes to account for coming to be: first is the matter in which a given form is produced; second is the form, which is also the cause for the sake of which; and third is the agent that produces the form in the matter. We find the same three causes in the last short passage we shall consider, from *Meteorology* IV.2. In this chapter, having just discussed the coming to be and passing away of the homogeneous bodies in general, Aristotle gives an overview of one process by which they come to be, namely, concoction (πέψις). He writes:

Concoction, then, is a perfecting (τελείωσις) by the natural and proper heat from the contrary passive qualities, these being the matter proper to each. For when it has been concocted, it has been perfected and has come to be. And the source of the perfecting happens by the heat that is proper, although it may be assisted through the help of something external; ... but the source itself is the heat in it. The end, for some things, is the nature, taken in the sense of form and substance (τὸ δὲ τέλος τοῖς µὲν ἡ φύσις ἐστίν, φύσις δὲ λέγοµεν ὡς εἶδος καὶ οὐσίαν); for others the end of the concoction is some subordinate15 form (εἰς ὑποκειµένην τινὰ  

connection with individual species, we risk missing the real explanation of these features, which will presumably also be general rather than specific. On the requirement of a “primitive and universal” connection between explanans and explanandum, see *An. post.* I.4–5; cf. also *Phys.* II.3, 195b21–8. If Aristotle follows this methodological principle consistently, then his discussions of τὸ οὗ ἑνεκα in *Phys.* II and *GC* are evidence that he considers this cause to apply primitively to natural substances and changes as such. On the claim that only living substances are natural substances in the full sense, see above, and also the appendix.

15 Both Lee (1962) and Webster (1984) have “latent.” Assuming this is the same distinction as Aristotle makes elsewhere between essential and accidental heat (*PA* II.2, 648b35–649a9, II.3, 649b21–28), I have used “subordinate,” which has the further advantage of being a sense of ὑπόκειµαι acknowledged by LSJ.
μορφήν).... This being-affected (τοῦτο πᾶσχειν) happens to all, when the matter and the wetness has been mastered: for this is what is determined by the heat in the nature. For as long as the proportion is in the matter, this is the nature (ἐως γὰρ ἄν ἐνή ἐν αὐτῇ ὁ λόγος, φύσις τοῦτ' ἐστίν) (379b18–380a1).

In this passage we once again find three principles: matter as passive, a corresponding active power, and a form or end. Likewise, we have a familiar cluster of terms: τέλος, εἴδος, οὐσία, φύσις, and λόγος, along with a new term, τελείωσις. Although the phrase τὸ ὑ ἐνεκά does not occur, τελείωσις may reasonably be considered stronger than τέλος: it captures the connotations of both τέλος, the end of a process, and τέλειος, that which is complete or perfect. An even more distinctive feature of this passage, moreover, is Aristotle’s repeated insistence that the heat that is the ἁρχή of concoction is not that of an external agent, but that which is proper and natural to the substance coming to be. It is within this nascent substance that active and passive qualities are being distinguished, and it is the “mastery” of the latter by the former that accounts for both the coming to be and the being of the homogeneous body in question. Thus the heat is “a source and cause of being moved and of being at rest in that to which it belongs primarily” (Physics II.1, 192b21–2), and concoction, like the development of an organism, is an internally directed process aimed at the perfection or mature form of the developing substance.
It should be noted, finally, that in this passage Aristotle is considering concoction as a kind of paradigm of those processes by which the homogeneous bodies come to be. It would thus be seriously misleading if important parts of his discussion applied only when concoction takes place in a living body, acting for its own purposes. In fact, as I have just pointed out, he is careful to emphasize the autonomous character of the process by distinguishing the immediate agent—the heat internal to the body being formed—from an external source of heat that may also contribute to the process. Finally, he explicitly states that the nature that serves as τέλος of the process consists in the proportion of heat, cold, wetness and dryness that constitutes the concocted body as a stably existing substance.16 This nature, which in some cases comes to be within a living substance, can thus be defined independently of any such substance.

Keeping in mind what we have learned from the preceding four passages, we are now ready to consider Aristotle’s longest and most explicit discussion of final causality among inanimate substances. Carefully considering the text and context of *Meteorology* IV.12 will both support the interpretation of the passages already considered, and enable us to find in their central concepts a general

---

16 As appears in the text, Aristotle also uses the term “concoction” for changes that result in an accidental unity. In this case, the result of the change lacks the stability characteristic of substances. For more on this distinction and on the distinction between internal and external heat, see *PA* II.2–3, 648a20–649b35.
analysis of matter-form composites, whether living or not, that gives Aristotle a highly unified, yet highly flexible ontology for his natural philosophy. It is in this ontology, as we shall see in the following two chapters, that his natural teleology is rooted.

2. The purpose and structure of Meteorology IV

The central claim of Meteorology’s final chapter, as I understand it, is that “all things are defined by the function: for the things able to perform their own function are truly each thing” (390a10–12). The chapter as a whole is programmatic and transitional: Aristotle begins it by stating that the preceding account of the homogeneous bodies in Meteorology IV is incomplete, and that what remains is to provide a separate, functional account for each of these substances. In the rest of the chapter, he discusses functional accounts in general and in relation to different sorts of substance. There is no doubt, then, that Aristotle thinks the homogeneous bodies are defined functionally, and thus teleologically. This means that proponents of the biological reading must also defend, more specifically, a biological interpretation of the functional accounts discussed in Meteorology IV.12. Such an interpretation is provided by Mary Louise Gill.17 According to Gill, the function of a homogeneous body is not something it has in itself, but is

rather the role it plays within a larger whole, such as a living body or an artifact. Flesh and bone, for example, are functionally defined only within a biological context, as in *Parts of Animals*. A complete discussion of homogeneous bodies like flesh and bone therefore necessarily includes an account of their biological functions, and the purpose of *Meteorology* IV.12 is to provide a transition between Aristotle’s science of inanimate bodies and his biology—in particular, between *Meteorology* IV, and the discussion of the homogeneous parts of living bodies in *Parts of Animals* II.1–9.

As we shall see, there are serious textual difficulties with Gill’s interpretation of *Meteorology* IV.12. Moreover, in claiming that what remains at the end of *Meteorology* is to discuss the role of homogeneous bodies in functionally organized heterogeneous bodies, Gill misunderstands the purpose of *Meteorology* IV as a whole. Finally, as I have tried to show in the first part of this chapter, there is strong, independent evidence that Aristotle understands the homogeneous bodies teleologically. Thus, there is little motivation to force a biological

---

18 In fact, there is a strong *a priori* argument against a biological interpretation of this chapter. In both *Mete.* IV.12 (389b27–8) and *PA* (II.1, 646a13–24) Aristotle states that the homogeneous parts of a living body serve as matter for the heterogeneous parts. Moreover, he clearly holds that form is predicated accidentally of that which serves as its matter, so that the matter (or what serves as matter) cannot be defined in terms of the form. Hence flesh and bone, as homogeneous bodies, cannot be defined in terms of the living substance as a whole. I have elsewhere addressed the claim that there is a tension between this account of matter and form and Aristotle’s explicit statements that flesh and bone (taken, I argue, in another sense) exist only when they are alive: see my 2001.
interpretation on Aristotle’s use of the term ἔργον (function) in the last chapter of Meteorology. As I argue below, Aristotle’s understanding of what we call functions makes it much more reasonable to think that elements and mixed bodies have ἔργα in their own right, and are defined in terms of these.

Gill is certainly correct, however, that our understanding of the task Aristotle sketches in Meteorology IV.12 will depend on what we think has been accomplished, with regard to the homogeneous bodies, in the earlier chapters of Meteorology IV. Before turning to the final chapter itself, therefore, we need to consider the work it concludes. In doing so, I shall try to establish the following two points. First, Meteorology as a whole is not about a particular class of substances, but rather about the production of various sublunary phenomena through the interaction of the elements. In keeping with this purpose, the first seven chapters of book four are directly concerned with a set of processes rather than with a set of substances—those processes, it so happens, that are involved in the coming to be of mixed bodies. Therefore, only chapters eight through eleven—and if we want to be even more careful, chapters ten and eleven—consider the homogenous bodies in their own right. In these chapters, Aristotle tells us what a complete account of the homogeneous bodies would look like. He lists and briefly describes their typical differentiae, and divides them into three
broad classes based on their matter. Although he deliberately stops short of considering their formal nature, he makes it quite clear that this consists in the active and passive qualities that mixed bodies have in themselves, not of their roles in functionally organized composites. My second point is that not only in IV.12, but elsewhere in book four as well, there are indications that Aristotle envisages a separate work (or works: see n. 42) covering the homogeneous bodies with respect to their specific forms. In the first eleven chapters of *Meteorology* IV, however, it is clear that the forms later to be discussed consist in qualities proper to the mixed bodies in themselves.

The first point to be established, then, is that *Meteorology* IV does not provide a full discussion of the homogeneous bodies even considered in themselves, for the simple reason that such a discussion is not appropriate to *Meteorology* as a whole. Here I am taking issue with the common view that Aristotle’s work on “meteorology” consists properly of books one through three only, whereas book four is Aristotle’s treatise on the homogeneous bodies. Although this last book does provide us with a great deal of information about the homogeneous bodies, it discusses them mainly insofar as their production and characteristics pertain to the science of meteorology, rather than as a distinct class of substances that can be investigated in their own right. More particularly, book four as a whole
considers the effects (ἔργα) produced by exhalations below the earth’s surface, coming immediately after an extended discussion of effects—such as the rainbow—produced by similar exhalations above the surface of the earth (III.6, 378a13–17). Like the previous three books, in other words, Meteorology IV has to do with the various natural effects produced by the interaction of the four sublunary elements within the basic structure of the cosmos described in De caelo and Generation and Corruption. In fact, book three ends with an overview of the bodies produced by these subterranean exhalations, which include stones, metals, and other such substances. Aristotle concludes thus: “We have, then, spoken generally about all these, and we must take each kind (γένος) separately and examine it in detail” (378b5–6).

What happens, more specifically, in book four? In the first seven chapters, Aristotle discusses the basic processes by which various homogeneous bodies come to be, focusing on the active and passive δυνάμεις involved. One important difference we do notice between these chapters and the preceding discussions is that Aristotle takes many examples, both of processes and of qualities,

19 Contrast Gill’s account, according to which Mete. IV.1–11 gives a complete account of homogeneous bodies as such (149), and IV.12 introduces the notion that they must also be understood in terms of their functional roles in various “complex autonomous wholes” (153).

20 The phrase “take each kind … in detail” is Lee’s (1962). The “kinds” of mixed body are mentioned again in the opening lines of IV.12, where the consideration of γένη provided in IV.1–11 is contrasted with a discussion still to come, which will deal with these same bodies καθ’ ἐκατότον (see below for further discussion of the γένη).
from the bodies of animals and plants. This is because the same sort of processes that produce stones and metals beneath the surface of the earth also produce the homogeneous parts of animals and plants. As Aristotle writes in *De sensu*, “Whatever external heat does in external bodies, that in the nature of animals and plants does the same” (4, 442a6–7). It is worth noting, however, that in these chapters the term “homogeneous body” does not even occur. It first appears, rather, in chapter eight, where Aristotle begins to consider the differentiae of the bodies produced by the processes described in the previous chapters.

In the first seven chapters of *Meteorology IV*, then, Aristotle discusses certain sets of processes: namely, those that produce homogeneous bodies within the earth, and that also, it so happens, produce other homogeneous bodies within animals and plants. In chapters eight and nine he goes on to discuss the most common differentiae of these bodies—that is, the qualities directly produced by the processes he has just considered. Although these qualities do constitute the forms of the bodies that have them, there are relatively clear indications that Aristotle is still working within the confines of *Meteorology*. Consider how, in the general remarks that open chapter eight, he introduces the differentiae that he is about to discuss as differentiae of the bodies that are constituted from water, earth, and “the enclosed exhalation of each.”
From these things it is clear that bodies are composed by the hot and the cold, which carry out their activity by thickening and solidifying (παχύνοντα καὶ πηγνύντα ποιεῖται τὴν ἐργασίαν αὐτῶν). Because they are produced (δημουγεῖσθαι) by these, heat is present in all of them, and in some cold as well, insofar as [heat] is absent. So, since these things belong due to the active, and the wet and the dry due to the passive (διὰ τὸ ποιεῖν ... διὰ τὸ πάσχειν), the mixed bodies (τὰ κοινὰ) share in them all. The homogeneous bodies, both in plants and in animals, as well as the metals such as gold and silver and all other such, are constituted from water and earth—from them and from the enclosed exhalation of each, as has been said elsewhere. They all differ from each other by the properties correlative to the faculties of sense (τοῖς τε πρὸς τὰς αἰσθήσεις ἰδίως), by being able to act in some way (τῷ ποιεῖν τι)—for a thing is white or fragrant or noisy or sweet or hot or cold by being able to act in some way with respect to sensation—and by other, more proper affections (οἰκειοτέροις πάθεσιν), whatever is spoken of as being affected—I mean like the soluble, the solid, the flexible, and any others of this sort: for all such things are passive (παθητικά), just like the wet and the dry. These are the immediate differences (τούτοις ἤδη διαφέρει) among flesh, bone, sinew, wood, bark, stone, and each of the other homogeneous and natural bodies (384b24–385a11).22

---

21 The term οἰκειοτέροις may mean either that these πάθη are more proper to bodies than is their ability to affect the senses, or that they are more properly called πάθη, due to their passive character. The former interpretation is weakened by the fact that the sensible πάθη include those relative to touch, the sense that Aristotle thinks apprehends bodies in their most basic characteristics (DA II.11). (The second group of differences also includes qualities relative to touch: see Mete. IV.4, 382a17–20 on touch as the standard for application of the terms “hard” and “soft.”)

22 Looking ahead for a moment to Meteorology IV.12, it is worth noting that although animals and plants are living bodies, and not simply bodies (see DA II.1, 412a12–20), their homogeneous parts must be differentiated, as homogeneous, by just the sort of πάθη that differentiate the homogeneous bodies in general. Flesh and bone, whatever else they may be, are composed of the elements in various proportions, and are differentiated as mixed or homogeneous bodies by the same sort of qualities that differentiate copper and tin. (For the distinction between somatic and psychic properties in Aristotle see my 2001, 359–60.)
In the following discussion, moreover, rather than considering the qualities just introduced as constituting the forms of specifically or generically similar homogeneous bodies, organizing his discussion according to the main classes or genera as he does in the biological works, he simply lists the eighteen common pairs of passive qualities (πάθη), states that most bodies (τὰ πλείστα σχεδὸν τῶν σωμάτων) differ by these affections, and goes on to give a brief account of the “capacity” (δύναμιν) of each differentia (see 385a19–21). In short, chapters eight and nine discuss a set of qualities that are relevant to Meteorology IV primarily insofar as they result from a certain kind of process, and not as the differentiae of a genus of substances that is being discussed in its own right.

It is, therefore, not until chapter ten that Aristotle turns his attention directly to the class of substances that has emerged from the foregoing discussion.23 Indeed, it is at the beginning of this chapter that Aristotle formally introduces, so to speak, the homogeneous bodies as a class:

By “homogeneous” I mean such as the metals—copper, gold, silver, tin, iron, stone, and the other such things, as well as whatever come to be by being extracted from these—and those in animals and

---

23 Additional evidence for my division of Mete. IV is its conformity with Aristotle’s methodological prescription (see e.g. An. post. I.1, II.7, 9–10) that in the case of complex objects of knowledge, one ought not to assume the existence of the subject matter (ὅτι, τὸ εἶναι) but to demonstrate it. It is easy to see the existence of the homogeneous bodies as a distinct class of substances emerging in Mete. IV.1–9. Then, in chapters 10–12, Aristotle provides a preliminary classification of these substances and a brief discussion of what will be involved in discussing their essence (διότι, τὸ τί ἦν εἶναι).
plants, such as meats (σάρκες), bones, sinew, skin, gut, hairs, fibres, veins, which are the immediate constituents of (ἐξ ὧν ἢδη συνέστηκε) the nonhomogeneous, such as face, hand, foot, and the other such things, and in plants wood, bark, leaf, root, and all other such (388a13–20).

Immediately afterward, he states that the goal of the following discussion is to provide a rough classification of these bodies on the basis of their material composition:

Since these things are constituted by another cause, and their matter from which is the dry and the wet, and hence water and earth, which respectively have these two capacities most clearly, and the agents are the hot and the cold, for these constitute and solidify from those, let us consider what sort of homogeneous bodies are forms of earth (τῶν ὁμοομερῶν ποῖα γῆς εἴδη), and which of water, and which are common (388a20–27).24

This is, in fact, what we find in the rest of chapter ten. Aristotle divides the homogeneous bodies into three main classes: those composed mainly of earth, those composed mainly of water, and those of a mixed composition.25 In doing so, he assigns the passive qualities discussed in chapters eight and nine, and thus also the bodies characterized by them, to one or another of these classes. Finally, in chapter eleven he completes this sketch by briefly considering the relation of heat and cold, the primary active qualities of the homogeneous bodies, to the three classes just distinguished.

24 Although the use may be non-technical, it is interesting that Aristotle refers to those things constituted from water and earth as the εἴδη, forms, of these elements.

25 Cf. GC I.7, 324b6–8: τὴν μὲν γὰρ ὄλην ... ὠσπερ γένος ὁν.
To summarize my first point about *Meteorology* IV, then, it turns out that the book is not a treatise on the homogeneous bodies as a distinct class of substances. It is, rather, the completion of *Meteorology* as a whole, insofar as it deals with the phenomena produced by exhalations enclosed beneath the surface of the earth. On the other hand, unlike other phenomena discussed in *Meteorology*, the processes discussed in book four do in fact give rise to a distinct class of substances. Thus in chapters ten and eleven, Aristotle introduces these substances as a subject for investigation in their own right, distinguishing their main genera and the differentiae typical of each. He does not, however, undertake a systematic discussion of the homogeneous bodies in their specific natures.

One might argue, of course, that the sketch found in the final chapters of *Meteorology* is all that Aristotle thought needed to be said about the homogeneous bodies as such. This leads to my second point: *Meteorology* IV as a whole suggests that there is, in fact, more to be said. Moreover, this is quite apart from any role these bodies may have in larger, functionally organized wholes. As early as chapter five, in fact, Aristotle lets us know that to understand the mixed bodies in their actuality, we need not refer to any such role. In discussing the process of solidification, he writes: “Now besides the matter there are two causes, the agent (τὸ ποιοῦν) and the affection (τὸ πάθος)—the agent as that whence the
motion, the affection as form (εἶδος)” (382a28–30). In other words, the form or actuality of bodies that come to be by processes like solidification is the passive quality (πάθος) or qualities, such as hardness, softness, and liquidity, that come to be in the matter as a result of these processes.

In chapters eight and nine, as already mentioned, a number of these qualities receive brief descriptions: specifically, the passive qualities perceptible by touch that Aristotle refers to in the introduction to chapter eight as “more proper” (see above). These two chapters, and especially the introduction thereto, make it clear that the formal nature of any homogeneous body consists primarily in its tangible qualities, which are the affections immediately produced by the action of the hot and the cold on the wet and the dry. In fact, chapter ten begins by summarizing the foregoing discussion thus: “It is by these affections and these differences that the homogeneous bodies, as has been said, differ from each other with respect to touch, and [they differ] further by taste, smell, and color” (388a10–13). In chapter eleven, finally, Aristotle adds to this picture the active tangible qualities, heat and cold.26 In short, to give an account of the formal

---

26 These primary active qualities, along with the primary passive qualities, the wet and the dry, have already been discussed at length in GC II.2, and the non-primary active qualities are simply the degrees of heat that lie between extreme heat and extreme cold. These two facts, perhaps, explain why in Mete. IV, the only differentiae of the homogeneous bodies that Aristotle discusses are the most common non-primary passive qualities.
nature of a mixed body is to discuss its specific perceptible qualities, and above all those qualities perceived by the sense of touch.\textsuperscript{27}

In addition to providing an overview of the formal differentiae of the homogeneous bodies, \textit{Meteorology} IV.8–9 also makes clear that the same sort of causal account applies to all such bodies, regardless of whether they are found in animals and plants or in inanimate nature. Flesh and bone, for example, like copper and tin, are produced by the hot and the cold from the wet and the dry,\textsuperscript{28} and they are differentiated by their sensible qualities and their capacities for

\footnotesize
\textsuperscript{27} The focus on tangible qualities should not be surprising, because Aristotle’s discussions of perception make it quite clear that in his view, it is the sense of touch that apprehends bodies at the most basic level. It may be helpful at this point to recall the relation between Aristotle’s account of perception and his understanding of bodies as essentially perceptible, and in particular, tangible.

First, Aristotle holds that to study the objects of the various senses belongs to the science of the senses. Thus in \textit{De sensu} 3 he states that whereas \textit{De anima} dealt with the actuality of sensible things, and with their identity as actual with the act of sensation, the present treatise must discuss what is potentially sensed; that is, it must say “what each of them is so as to produce sensation and actuality (τι δὲ ἐκαστὸν αὐτῶν ὁν ποιήσει τὴν αἴσθησιν καὶ τὴν ἐνέργειαν)” (439a17–18; cf. also \textit{PA} I.1, 641a36–8; cf. \textit{DA} I.1, 402b12–17; II.6, 418a7–8). Later (4, 440b27–8), he notes that the object of hearing was discussed sufficiently in \textit{De anima} (II.8). In \textit{De sensu} itself, he discusses at length the objects of sight (ch. 3), taste (ch. 4), and smell (ch. 5). Nowhere in either treatise, however, is there a corresponding discussion of the objects of touch, although in \textit{De sensu} 3 Aristotle refers to such a discussion (439a12). The passage in question is usually understood to say that such a discussion is required. However, the phrase ὁμοίως δὲ καὶ περὶ ἀφῆς, immediately after a list of the objects of the other senses, should actually be taken to mean that sound, color, smell, and taste must be discussed just as touch has been. In \textit{De anima} II.11, Aristotle specifies which discussion he has in mind: “What are tangible, then, are the differentiae of body as body: I mean the differentiae that distinguish the elements, hot, cold, dry, wet, concerning which we have spoken in the writings on the elements” (423b27–9). \textit{Generation and Corruption}, that is, and by extension \textit{Meteorology} IV, have already covered the object of touch.

\textsuperscript{28} Cf. \textit{De sensu} 4, 442a6–7: “In fact, whatever external heat does in external bodies, that in the nature of animals and plants does the same.”
being acted upon. The unity of this class of substances is clearest, however, in what I have called Aristotle’s formal introduction of the homogeneous bodies. The only differences he recognizes, for present purposes at least, have to do with their context of origin: beneath the earth, in animals, in plants.

If homogeneous bodies are to be defined in terms of the tangible qualities, however, what indication is there outside chapter twelve that Aristotle envisions a more detailed discussion of these bodies than he provides in the final chapters of Meteorology IV? Let us return to an interesting feature of the introduction to chapter ten (see 388a20–27, above). In this passage, having just formally introduced the homogeneous bodies, Aristotle uses the causal schema that we have seen in Generation and Corruption II.9 to characterize these bodies in terms of three causes. These are the matter (ὕλη), the agent (τὰ ποιοῦντα), and “another cause” (ἄλλης αἰτίας, l. 21) by which, as opposed to the matter and the agent, the homogeneous bodies are “constituted” (συνέστηκεν). The deliberately vague reference to form in this text is striking, especially by contrast with the explicit mention of the matter and the agent. In fact, it is these latter two causes that structure the rest of chapter ten and chapter eleven: as mentioned briefly already, chapter ten provides a broad classification of the homogeneous bodies based on their matter, and chapter eleven sketches how the bodies in these classes will tend to
be hot or cold, depending on the role of heat and cold as agents of their coming to be. It is only in chapter twelve, as we shall see, that Aristotle returns to the “other cause.”

Textual evidence aside, it should be clear that without a systematic discussion of the homogeneous bodies with respect to their specific forms, Aristotle’s natural philosophy would be seriously incomplete. It is not surprising, therefore, that in the last chapter of *Meteorology* he clearly states the need to deal with this topic in another work. What sort of work this should be will depend on how we read the chapter as a whole.\(^2\) Insofar as the results can be predicted from the rest of *Meteorology* IV, however, what we are looking for is something like what we find for living things in Aristotle’s biological works: a further division of the homogeneous bodies into intuitive lesser genera and species, each characterized by a coherent set of differentiae, and perhaps also a discussion of how the material composition and coming to be of these bodies can and must be understood in relation to their forms.

\(^2\) Introducing her interpretation of *Mete.* IV.12, Gill states that “If IV 12 is an introduction to Aristotle’s biology, we need not take its claims about teleology to apply to all materials. My main suggestion is that the chapter confines its teleological proposal to materials within a biological context” (148). Later, Aristotle’s attribution of functional definitions to inanimate bodies like copper and silver force her to retreat to the claim that homogeneous bodies are defined functionally in terms of “the form of some complex autonomous whole, such as an animal or an artifact” (153). It would be helpful to have some further indication of how this latter thesis relates to her reading of *Mete.* IV.12 as an introduction to the biological works.
3. Teleology in *Meteorology* IV.12

Let us proceed, then, to *Meteorology* IV.12. Given the importance of this chapter, it will be worthwhile to reproduce it as a whole. Rather than dividing the text and commentary line by line, I shall simply give the text and state what I think Aristotle is doing, arguing for particular points as necessary.

Since we have distinguished concerning these things, let us say with respect to each what is flesh, or bone, or any of the other homogeneous: for we have the things from which the nature of the homogeneous bodies is constituted, their kinds (γένη), of what kind each is, through coming to be: for the homogeneous are from the elements, and all the works of nature from the homogeneous as matter (389b23–9).

Whereas all are from the things mentioned as from matter, with respect to substance [they exist] by the account (λόγος) (389b28–9).

It is always clearer regarding the later things, and generally whatever are like tools and for the sake of something. For it is clearer that a dead man [is a man] homonymously. Thus also the hand of one who has died [is a hand] homonymously, just as also stone flutes would be spoken of. For these also are as tools of a sort. Such things are less clear regarding flesh and bone. Still less regarding fire and water (389b29–390a3).

For that for the sake of which is least clear here, where indeed there is mostly matter: for just as if also the last points should be taken, the matter is nothing other besides itself, and the substance is nothing other than account, the intermediate is proportionate by being near to each, since any of these whatsoever is also for the sake of something, and is not fire or water in any state (πάντως ἔχων), just as neither is flesh or gut. Even more than these, face and hand (390a4–10).

All are defined by the function: for the things able to perform their own function truly are each thing, as an eye if it sees (what cannot do so being so called homonymously, like a dead or stone [eye]: for neither is a wooden saw [a saw], except as an image. Thus
also flesh: but its function is less clear that that of the tongue. Likewise also fire: but still less clear [is its function] to natural science (φυσικῶς) than the function of flesh. Likewise also the homogeneous bodies in plants, and those that are inanimate, as copper and silver: for all exist by a certain power either of acting or of being affected, just as do flesh and sinew: but their accounts are imprecise (390a10–20).

Hence when they are present (ὑπάρχει) and when not, is not easy to tell, unless they are thoroughly decayed and only the shapes are left, as also the bodies of old dead persons suddenly become dust in the tombs; also fruits in shape only, not appearing so to perception, exceedingly old; and the solids from milk (390a20–b2).

Such parts are brought to be by heat and cold and by the motions that are [caused] by these, being solidified by the hot and the cold—I mean as many as are homogeneous, like flesh, bone, hair, sinew, and however many are such. For they all differ by the differences mentioned before, tension, ductility, fragmentability, hardness, softness, and the other such: these come to be by the hot and the cold and the mixed motions (390b2–10).

Those constituted from these, the non-homogeneous, like head or hand or foot, would no longer seem [such] to anyone, but just as also cold and heat and motion are the cause of copper and silver coming to be, but not, further, of a saw or a cup or a box, but rather in this case art, and in the other nature or some other cause.

If, then, we know of what kind is each of the homogeneous bodies, it is left to say what each of them is—what is blood, or flesh, or semen, and each of the others. For thus we will know for each why it is, and what it is, if we have either the matter or the account (λόγος) of the coming to be and passing away, and especially when we have both, and also whence is the source of the motion. When these things have been clarified we must likewise inquire about non-homogeneous, and finally about the things constituted from these, like man, plant, and the other such.

In the opening lines of this chapter (389b23–9), Aristotle states that by considering the coming to be of the homogeneous bodies, we have come to
understand (a) the things from which they are constituted (i.e., their matter), and (b) the kinds into which they fall. This is, of course, precisely what has taken place in the previous two chapters: Aristotle has discussed water and earth as the matter of the homogeneous bodies, has divided these bodies into three broad classes on the basis of their matter, and has further indicated how these three classes are differentiated with respect to heat and cold. Nevertheless, he continues, it is still necessary to deal with the nature of each individually. Now unless Meteorology IV.12 is completely disconnected from the previous chapters, what in fact remains is a systematic discussion of the various homogeneous bodies in terms of the specific tangible qualities that constitute the form of each. Finally, Aristotle concludes his opening statement by pointing out the importance of the discussion he has in mind. The homogeneous bodies, he reminds us, exist at a level of composition intermediate between that of the elements and that of the

30 The last paragraph of the chapter may seem to indicate that Aristotle is aiming immediately at a discussion of the role played by certain homogeneous bodies in living substances. On my reading this paragraph looks even farther ahead: after the discussion of the homogeneous bodies is complete, further progress through the field of natural philosophy leads us to a study of the living substances formed from them and their heterogeneous compounds. It is important to keep in mind, here, the distinction between homogeneous bodies and the homogeneous parts of living bodies. Just above (390b2–10), Aristotle has told us that the homogeneous bodies, including flesh, bone, and so forth, differ by the tangible qualities discussed in chapters 8–9. This is true, however, only of homogeneous bodies considered in themselves. The homogeneous parts of living substances are defined explicitly in terms of these substances, and cease to exist when the latter pass away. In 390a20–b2, Aristotle tells us that this is not the case for flesh and sinew considered as homogeneous bodies: only when the remains of a living substance are thoroughly corrupted can we say for sure that the homogeneous bodies that served a matter for that substance are no longer present.
nonhomogeneous works of nature. In other words, investigating the specific natures of the various homogeneous bodies will reveal (a) what things, specifically and actually, are formed directly from the elements, and (b) from what things, in virtue of their own natural qualities, nature’s more complex works can be formed.

Immediately after his opening statement (389b28–9), Aristotle brings us back to the “other cause” of chapter ten. At every level of composition, he writes, it is not sufficient to specify a thing’s matter. If we want to get at the οὐσία of each thing, we must realize that it exists not only from matter, but by a λόγος, an account or definition. This brings us to the first reason found in the text of Meteorology IV.12 itself for rejecting the biological interpretation of this chapter. A thing’s λόγος, as we know from the Metaphysics and elsewhere, is what corresponds most properly to the question τί ἐστί, which Aristotle now wants to answer for the homogeneous bodies. He has already told us in chapters eight and ten, however, that the differentiae of the homogeneous bodies are their active and passive tangible qualities. In IV.2, moreover we have seen a correlation of the terms οὐσία and λόγος with the εἶδος or φύσις that results from
concoction.\textsuperscript{31} Therefore, we can reasonably expect the \textit{λόγος} of the homogeneous bodies to be given in terms of the differentiae that have already been mentioned.

Nevertheless, Aristotle is not content with simply stating his point. He seems to realize that as a universal claim, the necessity of form and its distinction from matter require considerable explanation, and it is this explanation that occupies the bulk of \textit{Meteorology} IV.12. As usual, his procedure is to explain what is less clear in terms of what is clearer,\textsuperscript{32} and in this case, he states, the point “is always clearer regarding the later things, and generally whatever are like tools and for the sake of something” (389b29–30). The “later things” are those whose existence presupposes the earlier,\textsuperscript{33} that is, those existing at a higher level of composition. The clarity they exhibit comes from the fact that they resemble tools (\textit{ὄργα}-\textit{να}), and are therefore obviously for the sake of something.\textsuperscript{34} In these cases, that is, it is clear that what looks like a man, a hand, or a flute need not actually be

\textsuperscript{31} Gill tries cautiously (153–4) to draw a connection between Aristotle’s use of the term \textit{λόγος} and his supposed transition, in this sentence, to the biological context, drawing support from \textit{GA} II.1, 734b28–34, in which the \textit{λόγος} of flesh and bone is clearly biological. However, in addition to the use of \textit{λόγος} with respect to concoction in \textit{Mete.} IV.2, and with respect to the mixed bodies more generally in \textit{GC} II.6, Aristotle generally uses \textit{λόγος} to stress that the account of a given substance must be given primarily in terms of its form, as opposed to its matter. Thus the connection in the biological works between the phrase \textit{ὁ λόγος τῆς οὐσίας} and the form of an organic whole is not due to any connection between \textit{λόγος} and organic wholes, but to the fact that the phrase is occurring in works dealing with living substances.

\textsuperscript{32} Cf. \textit{Phys.} I.1, 184a16–21.

\textsuperscript{33} Cf. \textit{Cat.} 12, 14a30–5.

\textsuperscript{34} The qualification “obviously” is mine, but seems justified by the fact that Aristotle refers twice, immediately below, to “that for the sake of which” in the case even of the elements.
one. Without the appropriate form, Aristotle implies, it will be called “man,” or “hand,” or “flute” only homonymously, receiving the name but not the λόγος τῆς οὐσίας (389b31–390a2).35 By contrast, he continues, the distinction between matter and λόγος is less clear in the case of the homogeneous bodies, and it is even more obscure when we come to the elements (390a2–4).

Not surprisingly, this confusion regarding the λόγος of the elements stems from the fact that in them, as Aristotle puts it, that for the sake of which is “least clear (ἥκιστα δῆλον).” And indeed it is at this point that we may be tempted to balk. We have, to be sure, seen in Generation and Corruption I.7 that the interaction of the elements can give rise to states that are causes as “that for the sake of which.” Moreover, the elements themselves come to be, and Aristotle says in Generation and Corruption II.9 that in all cases of coming to be alike, the formal element is the cause “as that for the sake of which.” Finally, Meteorology IV.2 refers to the nature that results from concoction as an end (τέλος). Nevertheless, we may well find it intuitively implausible that the elements, or even the homogenous bodies, exist for the sake of something. We may, therefore, be tempted to explain away these texts, or to relate them somehow to Gill’s

35 Cf. Cat. 1, 1a1–2.
“biological context.” Either way, we can follow Gill in relativizing to this biological context Aristotle’s claims about function in IV.12.

Aristotle himself, however, seems to have been just as concerned about the reactions of his contemporaries to the claim that there is something “for the sake of which” in the elements and the homogeneous bodies. In any case, he proceeds to explain why τὸ οὗ ἔνεκα is not obvious in these cases (390a4–7). What he seems to mean is something like this: in “later things,” the various layers of form obviously determine the substance in question to a very particular activity, its ἔργον. Where there are fewer layers of form, however, the determination is less complete: “there is mostly matter.”36 Whenever there is any composition, however, there is at least some determination, and so a characteristic activity can

36 The place of the elements in this progression raises a controversial question. Aristotle’s inclusion of the elements among those things that must be understood καθ’ οὐσίαν, and not just with respect to their matter, only makes sense if we hold that there is a “first” or “prime” matter underlying the basic qualities that constitute the elements. If, therefore, Gill (1989, ch. 2) were right that the traditional doctrine of prime matter is not, in fact, found in Aristotle, and that the elements are therefore “pure matter” (1997, 145), then we might be pushed toward her interpretation of Mete. IV.12. I do not, however, think that her argument against prime matter succeeds. Without attempting to defend this claim in detail here, I will simply suggest that for Aristotle, every change requires a substratum, and that, pace Gill, the stable quality in an elemental transformation cannot serve as substratum for the quality that comes to be and for that which this quality replaces. To explain elemental transformation, Aristotle appeals to a distinction between the four elements, and body in general. Body in general never exists apart; it is rather a determinable (we might prefer to call it “corporeality”) that is at every moment determined by two perceptible qualities. (Gill may well be right, however, that these qualities can change only one at a time, because there must be some persistent actual determination that corporeality underlies while it serves as substratum for a transformation.) The texts that I would use to support this interpretation include GC I.3–4, 319a17–b25; I.5, 320a26–b25; I.6, 322b18–21; II.10, 329a9–14, 329a24–b4; II.5, 332a27–b1.
be specified, for the sake of which the composite in question exists (390a7–10).

Thus it is that Aristotle can say, with respect to any composite whatsoever, that it is not itself in just any state (πάντως ἔχον).\(^{37}\)

The phrase πάντως ἔχον is an important one, because it establishes a clear relation between Meteorology IV.12 and Generation and Corruption II.6. There, as we have seen, Aristotle states that the cause of natural coming to be is τὸ οὕτως ἔχον, which he equates with having a definite λόγος or formal nature. To say that fire is not fire, or water water, if it is πάντως ἔχον, is simply another way of stating the same thing. What Aristotle means, of course, is that a thing called “water” is not actually water unless it has the form of water, namely, the basic

---

\(^{37}\) This statement, which Aristotle explicitly applies to the elements, suggests another problem for Gill’s account: there is no suggestion anywhere in the corpus that Aristotle thinks the elements are defined even partly in terms of their roles in “complex autonomous wholes” — within which, moreover, as Gill admits (n. 27), he does not think they are actually present. It is unlikely that in Mete. IV.12 Aristotle is sanctioning a sense of the term “fire,” for example, in which it can be applied literally only to fire that is a constituent of a living being. The same problem applies with respect to the homogeneous bodies that occur outside living things. Gill suggests that “The functional-form of copper or silver is imprecise simply because the function of these materials within a higher complex is hard to recognize and characterize.” It is unlikely, however, that Aristotle would allow such a functional definition of copper or silver — which, unless it simply involved the tangible qualities of these substances, would presumably have to be different for copper in a hairpin than for copper in a cup (Gill’s examples). In such cases there is, in fact, no analogue to the definition of homogeneous parts such as flesh and bone. As Gill herself would agree (159–60), the definition of flesh, for example, in terms of its biological function, is not a definition of the homogeneous body flesh, which happens to serve a function in a complex whole. Rather, it is a definition of flesh in another sense of that term, according to which it designates the organ of touch that has the homogeneous body flesh for its immediate matter. There is no reason to suppose that Aristotle would countenance a corresponding “organic” sense of “copper” or “silver.” For more on the two definitions of homogeneous parts/bodies, see my 2001. I hope the present discussion will confirm my contention there that Mete. IV.12 deals only with homogeneous bodies, and not with the homogeneous parts of living bodies.
contraries by which water is defined. To be more precise, he might have used the language of *Generation and Corruption* II.9: water does not actually exist when that in it which “can be and not be” is in just any state. What is required is not merely τὸ πάντως ἔχειν, but τὸ οὕτως ἔχειν—in other words, that state which is the formal nature of the composite.

In the course of explaining how it is that every composite exists for the sake of something, Aristotle adds yet another locution for form to an arsenal that already includes µορφή, εἶδος, οὐσία, λόγος, and τὸ οὕτως ἔχειν. All things, he says, are defined by their function (ἀπάντα δ᾽ ἐστὶν ὡρισμένα τῷ ἔργῳ) (390a10–11). To make this point clear, he runs once more—for the third time in fact—through the various levels of composition, this time in the language of ἔργον, and repeats the point about homonymy (390a11–9). Moreover, he carefully gives examples of homogeneous bodies not only from among the parts of animals, but also from those of plants and from the inanimate realm. In all three cases, they are defined in terms of an ἔργον. Finally, he is willing to admit that

---

38 *PA* II.1 contains a similar review of these levels of composition (646a12–24). In that chapter, however, Aristotle is primarily concerned with flesh, bone, and so forth as homogeneous parts of animals, rather than as homogeneous bodies that can be considered in their own right.
where there are no ὀργανα, accounts are less precise, and so it is hard to tell just by looking whether a homogenous body really exists or not (390a19–b2).³⁹

I hope the preceding paragraphs have already shown that in light of the passages discussed above, the difficult central lines of Meteorology IV.12 are quite plausibly read as a general discussion of composite substances, the basic point of which can be applied at any level of composition. There is, however, more to be said about Aristotle’s use of ἐργον, not all of which is a direct commentary on Meteorology IV.12. Therefore, I will comment briefly on the rest of the chapter, and then return to the meaning of ἐργον.

In the lines of IV.12 that remain (390b3–23), Aristotle simply reiterates some points concerning the homogeneous bodies that we have already made, stating these points now from the point of view of οὐσία. He wants to stress that the agents discussed in Meteorology, namely, the hot and the cold, are sufficient to account for the coming to be of the defining capacities of these bodies in their matter.⁴⁰ This point further solidifies my interpretation of the preceding lines,

³⁹ It is unclear what it means for definitions to be οὐκ ἀκριβεῖς; it may have to do with the fact that the ultimate matter is less fully determined at lower levels of composition.

⁴⁰ Gill takes Aristotle to be saying that in theory, the homogeneous bodies flesh and bone (as opposed to the living parts by the same name) could be generated by heat and cold outside any living animal. Although she is correct that nothing in Aristotle’s natural philosophy excludes this possibility, he actually means, as in the passage cited above from GA II.1, that the production of these parts of animals, considered precisely as homogeneous bodies, can be explained in terms of the heat and cold present within the animal.
because these lines closely parallel the opening passage of IV.8 (see above).

There, as here, Aristotle connects the differentiae of the homogeneous bodies with the agents that account for their coming to be. These agents are, of course, the hot and the cold. When it comes to the nonhomogeneous bodies, however, a new cause must be introduced, one correlative not to properly tangible qualities, but to particular shapes and arrangements of parts. By contrasting the sort of agent required to produce a homogeneous body, as such, from the sort required for various nonhomogeneous bodies, Aristotle separates his account of the former from his discussions of biological reproduction and of art, to whose products heat and cold in themselves are not adequate.41 With this final distinction in hand, Aristotle can now make more explicit the next topics to be covered in natural science: first the λόγοι of the various homogeneous bodies, and then the treatment of the nonhomogeneous bodies and of the living bodies composed from them.42

41 Aristotle makes the same point in GA, stating explicitly that heat and cold are sufficient to produce the qualities that in Mete. he identifies as the differentiae of homogeneous bodies: see II.1, 734b20–735a5.

42 The former discussion may well be found in three separate works. For the homogeneous bodies found in animals, the obvious reference is to PA. It is fairly certain, however, that Aristotle never intended to write a similar work on plants; the two botanical works (Historia plantarum and De causis plantarum) of his collaborator and successor in the Lyceum, Theophrastus, seem intended to fill this lacuna. It is noteworthy that Theophrastus also wrote a work De lapidibus. In any case, it seems clear that in the outline of Aristotle’s natural investigations, Mete. IV.12 is not immediately followed by biological inquiry strictly speaking, but by the various
Let us return, now, to Aristotle’s use of the term ἔργον, and its implications for Meteorology IV.12. To limit the length of this discussion, I consider for the most part only Meteorology, adding a few brief additional considerations at the end.

It should be obvious that in reading Meteorology IV.12, a great deal turns on how we are predisposed to interpret the term ἔργον. As I have stated elsewhere, I believe that the usual translation, “function,” is often seriously misleading. Thus, explaining his use of function in IV.12, Aristotle writes, “for all exist by a certain power of acting or of being effected” (390a18–9). Whether or not he intends this as a formal account of ἔργον, this gloss surely does not line up with our use of the term “function,” unless we think of an action that is performed in the context of a larger whole and somehow contributes to it. This larger whole, whether it be a complex human activity, an organized group of persons, a machine, or an organism, is always present in contemporary philosophical discussions of function, but it is by no means always present in Aristotle’s use of the term ἔργον.

---

43 See my 2001, esp. 363–4. The following remarks support the arguments found there, which are based on Mete. IV.12 alone.

44 Ironically, the term “function” is misleading even in the famous “function argument” of NE I.7, as anyone who has taught that text knows.
Consider the rest of *Meteorology*. Throughout books I–III, ἔργον is used to mean simply an effect, something accomplished or caused by something else. Typical phrases are “this [theoretical conclusion] is clear from the effects (ἐκ τοῖς ἔργοις),” and “the effect (or work: ἔργον) of such and such a cause.” The causes are, in general, the various exhalations whose activity Aristotle is investigating. Although it would perhaps be unwise to put too much weight on these occurrences taken by themselves, there does not seem to be a clear line between this use of the term—for which translators never use the English “function”—and those for which the English term is typically used.

For example, because Aristotle most often applies the homonymy principle to living things and their parts, translators typically render ἔργον as “function” when the homonymy principle is clearly in use. This is the case, for example, in *Meteorology* IV.12. Earlier in the same book, however, in IV.1, there is an interesting passage in which Aristotle uses ἔργον to denote the effects of precisely those qualities in virtue of which the elements actually exist:

First, then, simple coming to be and natural change are universally the work of these capacities, as well as the contrary passing away according to nature (πρῶτον μὲν οὖν καθόλου ή ἀπλῇ γένεσις καὶ ή φυσικὴ μεταβολὴ τούτων τῶν δυνάμεων ἐστιν ἔργον, καὶ ή ἀντικειμένη φθορὰ κατὰ φύσιν) (378b28–30).
The capacities in question are, as the context makes clear, the hot, the cold, the wet and the dry, and so in this passage we have an ἔργον, in the sense of effect, which is an effect of precisely those capacities that make the agent what it is.45

This passage, together with the similar uses of ἔργον in the first three books of Meteorology, suggests that to translate ἔργον as “function,” in contexts where the English term is appropriate, may well be to exaggerate the distinctions between various senses of this term that are present in Aristotle’s Greek. It seems likely that in its primary sense, ἔργον signifies a thing’s typical activity: that which it is suited to do or to bring about. When the object in question is part of a larger whole (in other words, is an ὀργανον) then the English “function” comes naturally. Nevertheless, it might be better to use a term, such as “work,” “job,” or “task,” that is less likely to impose contemporary philosophical distinctions on Aristotle’s text.

This initial conclusion is strongly supported by another passage from Meteorology IV, this time in chapter nine. Here, moreover, we find not only support for a general thesis about the term ἔργον, but a clear use of that term in the sense required to support a non-biological interpretation of IV.12. In concluding his discussion of bodies that give off fumes, Aristotle writes: “Sweet wine gives off

45 Cf. IV.3, 381a30, where roasting and boiling are referred to as ἔγγυα.
fumes, for it is fat, and does (ποιεῖ) the same things as oil, for it is not solidified
by cold, and will burn. It is wine in name, but not in effect (ἐστιν δὲ ὄνόματι οἶνος, ἔργῳ δ' οὐκ ἐστιν), for the taste is not winelike, and so it does not intoxi-
cate” (387b9–13). Although the point is typically missed by translators,46 these
lines contain a clear application of the homonymy principle. Aristotle states that
what is called “sweet wine,” because it does not have the ἔργον (or ἔργα) typical
of wine, is wine in name only. The ἔργα to which he refers, however, are nothing
other than differentiae of the sort that he has already said characterize homoge-
neous bodies.47 As it turns out, then, IV.12 does not include the only statements
in Meteorology, nor those that are clearest of interpretation, to the effect that the
homogeneous bodies are defined in terms of their ἔργα.

The foregoing considerations make it clear that the use of ἔργον in Meteor-
ology IV.12 should not, by itself, incline us to the biological interpretation of that
chapter. If, moreover, to specify the λόγος of a composite is always to specify
that for the sake of which it exists, and to specify that for the sake of which it

46 Lee has, “And though called wine, it has not the effect of wine”; Webster: “Really it is
not wine at all in spite of its name.” Compare again the clear contrast in the Greek: ἔστιν δὲ ὄνό-
ματι οἶνος, ἔργῳ δ' οὐκ ἐστιν.

47 The reference to intoxication might be interpreted as introducing the effect of sweet
wine on its drinker, thus justifying the use of the term ἔργον. However, these comments are part
of a discussion of the qualities that differentiate homogeneous bodies in themselves. In this con-
text, the presence or absence of a capacity to intoxicate serves as an indicator that two substances
differ in themselves.
exists is to specify its ἔργον, then these considerations concerning Aristotle’s use of ἔργον positively support a non-biological, and indeed a non-functional interpretation—in the English sense of “function”—of τὸ οὗ ἔνεκα as well. The way is clear for a non-biological interpretation of the chapter as a whole.

4. Limitations on the use of teleological concepts and explanations

I have argued that according to Aristotle in Meteorology IV.12, the ἔργα of homogeneous bodies, for the sake of which they exist, are nothing other than those characteristic actions and affects, the capacities for which constitute their substance, nature, and form. These capacities exist in the wet and the dry as their proximate matter, and heat and cold, the elemental agents of all inanimate change, are sufficient to produce them. In short, the ἔργον and οὗ ἔνεκα of each homogeneous body exists in a causal nexus, complete at its own level, that includes no reference to living bodies. This causal nexus is ontologically presupposed by, and therefore ontologically and definitionally independent of, the analogous nexuses found at higher levels of composition.

Moreover, the four texts examined in section one all point toward the same causal nexus that has emerged from our analysis of Meteorology IV.12. This nexus, which we found clearly and explicitly set forth in Generation and Corruption II.9 as providing a causal analysis of all natural coming to be, consists of
matter, a form that is also that for the sake of which, and an agent that brings about the form in the matter. *Meteorology* IV.12 confirms that this analysis is intended to apply not only to the homogeneous bodies, but also to living bodies and their parts, and even to the elements.

There remains, finally, the question of why teleological concepts and explanations appear so seldom in non-biological works such as *Generation and Corruption* and *Meteorology*. Much of the answer to this question is implicit in what has already been said—specifically, in our discussion of *Meteorology* IV.12. Aristotle’s natural teleology, as this chapter makes clear, is strongly connected with his hierarchical account of the structure of natural substances. In particular, he holds that form and function are prominent in a given substance to the extent that it is not merely a body, but rather, through various layers of composition, exists at a higher level of determination and actuality. An obvious corollary of this observation, however, is that the role of teleological concepts and explanations in the science of inanimate nature, though analogous to their role in the life sciences, is nevertheless much reduced. In a word, the simpler the substance, the less teleology has to do. Thus although teleological concepts are theoretically appropriate in discussions of these relatively simple substances, they are likely to
appear less often in explanatory practice. It is only in biology that getting one’s theory straight becomes essential for the actual practice of science.

Moreover, it is worth noting that if my interpretation of Meteorology IV.12 is correct, then Aristotle did not actually write a complete treatise on the homogeneous bodies as such. If he had written such a work, he presumably would have tried to meet the demands both of his scientific method in general and of Meteorology IV.12 in particular. This would mean considering the forms, or combinations of active and passive qualities, by which the various homogeneous bodies are constituted, and indicating how the account including these active and passive qualities would, in each case, be an account of a single thing—i.e., a definition (see An. post. II.10; Metaph. Z.12, H.6). It would also mean considering more carefully the relation between a homogeneous body’s form, its matter, and its coming to be, with a view to showing that the matter and coming to be are hypothetically necessary for the form (see An. post. II.11, Phys. II.9, PA I.1). In short, it is not at all clear that such a treatise would be completely different in explanatory style from Aristotle’s biology.

Indeed, it may well be that teleological concepts and explanations are present in Aristotle’s non-biological science exactly where they are called for. For example, many of the arguments in De caelo depend on the implicit premise that
the cosmos is neatly, rationally ordered—perhaps even organized in the best possible way. This is, of course, a teleological premise, a premise to which *Metaphysics* Α at least suggests Aristotle is committed. At times, moreover, the arguments in *De caelo* become quite explicitly teleological (see II.3–5, 12; also the end of III.8). However, as stated at the beginning, I do not want to hang too much on *De caelo* here.

Moving on, then, to *Generation and Corruption*, it is worth noting that for most of this work, Aristotle is not discussing the general causal structure of coming to be. Rather, he has first to establish the existence of coming to be and passing away, next to distinguish it from other sorts of change, and then to deal with combination, action, being affected, and other related topics. Within this sequence of topics, form considered as end does in fact appear in the first place where one might reasonably expect it to be mentioned: the account of action (I.7). After this long preparatory discussion, Aristotle next considers the elements and the manner of their coming to be and passing away, and here—the next good place to mention that for the sake of which—he criticizes Empedocles for failing to consider form as the nature and good of that which comes to be (II.6). Next, he provides a brief preview of how mixed bodies come to be from the elements. It is not, therefore, until II.9 that Aristotle finally considers himself in a position to
discuss the principles of all coming to be alike. Here, as in *Physics* II, form as end holds a prominent place.

Turning to *Meteorology*, we should not be surprised if, despite the fact that Aristotle’s homogeneous bodies have a teleological structure, teleological forms of explanation are scarce in this work as a whole. As we have seen, *Meteorology* is not about a particular class of substances, their coming to be, and their proper activity. Rather, it is about phenomena resulting from the activity and interaction of the sublunary elements within the overall structure of the cosmos. These phenomena occur with some degree of regularity, and are therefore open to investigation. They are not, however, proper effects or functions of particular substances, and their occurrence is subject to a good deal of chance. This is why Aristotle begins the work by stating that the events to be studied, although they are natural, have an order that is less perfect than that of the primary bodies. In short, the subject matter of *Meteorology* is such as to exclude the use of teleological explanation for most of the phenomena it studies. When the investigation does turn to a distinct class of substances, however, Aristotle carefully points out that the complete account of these substances will be teleological (IV.12).

The exclusion of teleological explanations from most of *Meteorology*, therefore, does not involve their exclusion from a particular class of substances.
Rather, it means only that the phenomena themselves are not ends at which substances aim by their nature. They do, however, take place because nonliving substances are aiming at their own ends, and, in so doing, actually bring about different effects in different circumstances. Meteorology is thus an excellent showcase for Aristotle’s account of chance (Phys. II.4–6). His universal teleology not only does not exclude, but actually entails, the existence of secondary natural phenomena that do not take place for the sake of an end. Thus in his zoology, likewise, he devotes the last book of Generation of Animals to the nonteleological explanation of analogous phenomena.

Moreover, to claim that the homogeneous bodies have a teleological structure does not entail that they always or even generally come to be for the sake of an end, at least in any strong sense. Aristotle’s teleology, properly understood, applies to matter-form composites in their being, and not only in their coming to be, and the teleological character of the former is not derived from that of the latter. This claim requires defense, of course, but the independence of Aristotle’s teleology as a whole from the notion of goal-directed change is an important theme in chapter two. For now, I shall only suggest that the concept of hypothetical necessity, which is inseparable from Aristotle’s teleology, applies to the production of composite substances because it applies first to their actual being.
To conclude, although teleological concepts and explanation are not prominent in *Generation and Corruption* and *Meteorology* by comparison with Aristotle’s biological works, there are several passages in which it becomes clear that his accounts of the elements and of the homogeneous bodies do incorporate such concepts. Moreover, *Generation and Corruption* II.9 and *Meteorology* IV.12 indicate that far from being essentially biological, Aristotle’s teleology is an essential feature of his accounts of natural coming to be (*GC* II.9) and of natural substances themselves (*Mete.* IV.12). Finally, the discussion of definition in *Meteorology* IV.12 also provides resources for explaining why teleological explanation is so much more prominent in Aristotle’s biological works than it is elsewhere in his natural science—if, indeed there is a significant contrast to explain.
CHAPTER TWO

THE METAPHYSICAL ROOTS OF ARISTOTLE’S TELEOLOGY

Besides establishing the presence of teleology in Generation and Corruption and in the Meteorology, our discussion thus far has located τὸ οὗ ἐνέκα within a causal nexus that Aristotle sketches in II.9 of the former work in order to account for “all coming to be alike.” This nexus is constituted by the end—that is, the form produced—by the matter in which it is produced, and by the agent that produces it. According to Meteorology IV.12, moreover, form itself must be understood in terms of the species-typical activities that follow upon its presence, and for the sake of which each composite substance exists. Of the two sorts of thing that have emerged as ends, therefore, form and activity, the latter seems to be ultimate. Although form is the immediate end of coming to be, a composite substance exists in the last analysis for the sake of its activity.

In this chapter and the next, I shall argue that the foregoing conclusions implicitly contain a simple yet complete account of Aristotle’s teleology. In De anima II.1 (412a20–9), Aristotle states that the term “actuality” (ἐντελέχεια)
signifies both form, or first actuality, and activity, or second actuality. Form is the actuality of a natural body, in other words, but this actuality itself involves a capacity for further actuality—that is, for activity of a certain kind. Because we have already identified both form and activity as ends, it thus appears that for Aristotle, that for the sake of which coincides perfectly with actuality. Clearly, however, this claim goes well beyond the negative conclusion of the previous chapter, for it entails that far from being bound up with his biology, the roots of Aristotle’s teleology are not even to be found in his understanding of nature in general, but rather in his first philosophy or metaphysics. It is the tension between these two disciplines, natural philosophy and metaphysics, that determines the basic motion of the following argument.

In referring to a tension between Aristotle’s first and second philosophy, I wish to highlight the fact that if his teleology is ultimately grounded in his understanding of being, it does most of its work in his science of nature. In the following discussion, therefore, I begin with an account of τὸ οὗ ἐνέκα taken from Physics II, where Aristotle lays out his general understanding of nature. Throughout the rest of the chapter, however, I emphasize how a careful analysis of this account leads us to the Metaphysics, and in particular to the related concepts of actuality (ἐνέργεια) and limit (πέρας). By contrast, the movement in
chapter three is from metaphysics back to the philosophy of nature. There, hav-
ing already explored the metaphysical dimensions of that for the sake of which, I
examine the criteria by which Aristotle identifies particular changes, whether of
substance or in one of the other categories, as being for the sake of an end.

In the present chapter, more specifically, I press the need for a metaphysi-
cal analysis by arguing that Aristotle’s concept of that for the sake of which
needs to be considered apart from his account of motion or change. As we shall
see, Aristotle himself highlights the problematic relation between these two no-
tions when he asks, in the *Metaphysics*, how teleological causality can pertain to
unchanging substances such as God. Although I am not primarily concerned
with Aristotle’s theology, the question of God will be an important subtheme in
the following pages, reminding us periodically of the need for a properly meta-
physical approach. For, as *Metaphysics* E.1 (1026a27–31) tells us, it is precisely the
existence of an immovable substance or substances that distinguishes the science
of nature, for Aristotle, from first philosophy.

In section one, therefore, I begin by introducing the starting point of my
analysis in this chapter: the characterization of that for the sake of which found
in *Physics* II.2. In this text, Aristotle states that every outcome of a continuous
change, provided that it be “what is best,” is an end or that for the sake of which.
Before proceeding to the analysis itself, however, I lay out a serious challenge that any interpretation of this and similar passages must face: Aristotle’s suggestion, in the theological aporia just mentioned and in its later resolution, that there cannot be a strong, general connection between τὸ οὖ ἔνεκα and motion or change. I also briefly clarify Aristotle’s use of the term “motion” (κίνησις), as opposed to “change” (μεταβολή), in the text from *Physics* II.2.

In section two, I begin to address the separation between that for the sake of which and motion by showing that, theology aside, Aristotle’s account of natural substances also precludes an account of τὸ οὖ ἔνεκα in terms of motion, insofar as the matter of every composite substance is, as such, for the sake of its form. The inadequacy of motion as a context for understanding Aristotle’s teleology points us to the more basic concept of actuality, in terms of which Aristotle provides a unified account of both being and coming to be for the sake of an end.

Having considered the relation between that for the sake of which and change, I return in section three to Aristotle’s seemingly unremarkable characterization of the end as an outcome: τὸ ἐσχάτον. Locating the metaphysical roots of this temporal “last” in the prior notion of a limit, I go on to show how the concept of a limit applies both to being and to coming to be or change, and to
associate it with the concept of actuality. For Aristotle, limit and actuality are distinct aspects of the same metaphysical principle.

Finally, returning in section four to the condition that an end or that for the sake of which be “best,” I briefly point out that goodness, as Aristotle understands it, also coincides with actuality. In short, the complementary concepts of actuality, limit and goodness provide us with an adequate understanding of “that for the sake of which,” the central concept in Aristotle’s teleology.

1. Introduction to the problem

To set the stage for the metaphysical analysis to which most of this chapter is devoted, I want to introduce two important texts and to make a preliminary clarification regarding Aristotle’s use of the term κίνησις. The first text is his introduction of that for the sake of which, in the general discussion of natural science that occupies book two of the Physics. The second is his first statement, in Metaphysics B.1, of the theological aporia mentioned above.

The passage from Physics II.2 is one of Aristotle’s most helpful and concise statements concerning τὸ οὗ ἐνεκα, and it will serve as a seminal text for our discussion in this chapter and the next. We should not, however, mistake it for a definition of that for the sake of which, or even for a statement of necessary and sufficient conditions. Although Aristotle’s programmatic discussions of natural
science both in *Physics* II and elsewhere contain a number of statements characterizing the “cause as that for the sake of which,” none of these statements is intended as a definition in the strict sense. This is not surprising: to expect such a definition would be to assume that Aristotle’s teleology can be explicated in terms of more basic, nonteleological notions. However, it seems further that nowhere in Aristotle’s writings does he give even an informal general explanation of what he means by τέλος or τὸ ὥν ἐνέκα. In *Physics* II and *Parts of Animals* I.1, the texts most likely to contain such an explanation, he is fully engaged in a prior debate concerning the very existence of action for an end in nature, and the consequent necessity of explaining natural phenomena teleologically. He therefore tends to focus on the most obvious cases of goal-directed action, and seems content to provide only sufficient conditions for calling something an end.¹

Turning now to the text itself, we find that in *Physics* II.2 Aristotle introduces τὸ ὥν ἐνέκα in terms of three criteria:

For with respect to things of which—the motion being continuous—there is some end, this last is also that for the sake of which. For this reason also the poet was carried away absurdly in saying, “he has death (τελευτήν), the very thing for the sake of which he

¹ For this reason—and also because *Physics* II has been mined so thoroughly by previous commentators—I have tried to avoid relying too heavily on this book alone for an understanding of Aristotle’s teleology. Rather, we shall begin with one of the most helpful texts from *Physics* II, and then follow the conceptual leads this statement provides to other texts, throughout the corpus, that seem likely to deepen our understanding.
came to be.” For not everything that is last claims to be an end, but
the best (194a29–33).²

In other words, something is an end if (a) it is last (ἐσχάτον)—in other words, is
an outcome of some motion; (b) the motion in question is continuous; and (c) the
outcome is, as he puts it, the best. In short, an end or that for the sake of which is
the good outcome of a continuous motion. Each of the criteria involved in this
description calls for an extended discussion; as already mentioned, moreover,
our concern in this chapter is with conditions (a) and (c). With respect to (a), we
shall focus in section two on motion, moving toward the concept of actuality,
and in section three on the notion of an outcome or “last,” moving toward the
concept of a limit. Section four, which is much briefer, deals with (c), while the
whole of chapter three is an exploration of (b).

If this text from Physics II.2 is the proper starting point of our discussion,
the metaphysical issues with which we shall have to deal are raised forcefully by
an aporia that Aristotle poses in Metaphysics B.2, and again in K.1. The puzzle,
which has to do with the very possibility of first philosophy, is this: How can
there be a general science of first principles if some types of principle are not

² Parallel texts are Phys. II.8, 199a8–9: “Further, in as many things as have some end, the
first and the successive are done for the sake of this”; and PA I.1, 641b24–6: “We always say that
this is for the sake of that, whenever there is clearly some end at which the motion concludes,
should nothing stand in the way.” Similar statements can be found elsewhere (in Phys. II.3 and
II.7, for example), but these three are the most complete.
found in every realm of being? The offending principle that Aristotle has in mind is the good—which, according to book A, is one of the central topics of first philosophy. As it seems, however, the good cannot be found among unchanging beings:

Everything that is good in itself and of its own nature is an end, and thus a cause for the sake of which other things come to be and exist, and the end and that for the sake of which is an end of some action, and all actions involve motion (B.2, 996a23–7; cf. K.1, 1059a34–8).

Consider then, in light of this argument, Aristotle’s first mover or God: on the one hand, God is completely unchanging; he neither comes to be nor passes away. He is not the outcome of any change, and so cannot be an end. On the other hand, precisely as eternal and necessary, God is supremely good, so that according to the aporia he must be an end, an outcome of change.3

It is important to clarify the exact nature of this difficulty. The problem is not that God is the end or good of other substances: to begin with, Aristotle has a perfectly coherent way of asserting this without making God an outcome of change. He makes a first distinction relative to this point in Metaphysics A.7 (1072b1–3), where the question is, precisely, how God can be “something for the sake of which.” An unchanging being cannot, he writes, be an end in the sense of

---

3 For God’s necessity and goodness see Metaph. A.7, 1072b10–30; N.4, 191b16–21.
a subject or beneficiary of change (τινὶ τὸ οὗ ἔνεκα), but it can be that toward which the change is directed (τὸ οὗ ἔνεκα τινὸς). As is clear from other passages, however, it is not at God in himself that other substances aim, strictly speaking, but at whatever share in the divine their nature permits (see De caelo II.12, 292a19–292b25; DA II.4, 415a27–b8; Metaph. Λ.10). After all, the alternative would be to say that natural substances aim at some sort of union with God, or even at becoming God. Rather, the immediate end for the sake of which (τινὸς) things that aim at God come to be and act is either a form or an activity, by which they themselves are constituted or of which they are the subject (τινὶ).4

There is another reason, as well, that the aporia of B.2 cannot be about God as the good of things other than himself. Insofar as God is the end of natural substances, his goodness is goodness for something else. Aristotle’s God, however, is supremely self-sufficient (αὐτάρκης); his goodness does not depend on anything outside himself, but is due simply to his own necessary and eternal

---

4 It is not clear whether Aristotle can give any further account of this unique way in which God is an end, or whether he even thinks such an account is required. The notion of sharing in or imitating the divine is involved in a conclusion about natural substances that Aristotle considers necessary on other grounds, and he seems to consider this sufficient justification for it. However, it may be helpful briefly to review this justification. In Metaphysics Θ.8 and Λ, Aristotle argues that absolutely speaking and in the cosmos as a whole, actuality must be prior to potentiality. He takes this conclusion to entail the existence of a substance completely free of potentiality, and therefore not subject to change. Without such a substance, he thinks, one cannot account for the motions (incomplete actualities) by which potentialities in the cosmos are constantly being actualized. In his works of natural science he invokes these conclusions as principles, most explicitly in De caelo II.12, 292a19–292b25; GC II.10, 336b25–337a8, and DA II.4, 415a27–b8.
being (*Metaph. A*.7, 1072b10–30; N.4, 191b16–21). Thus, according to the aporia God should also be an end simply in himself, without reference to the motions and activities of natural substances; nor can the motion involved in being an end be located outside him.

Finally, it is important to note that Aristotle himself is clearly committed to the minor premise of the aporetic syllogism, that every good is necessarily an end. In fact, he considers something an end or that for the sake of which if and only if it is good. The precise problem at hand, therefore, is that if this identification of the good with that for the sake of which is to stand, there must be some sense in which, precisely as an eternal and unchanging being, God is an end and that for the sake of which. Whether these terms apply to God in just the way they apply to the forms and activities of changeable beings is another matter; at least, however, their theological usage must be recognizably related to the sense they have in discourse about nature.

Before beginning to see how Aristotle resolves this aporia, as a final preliminary we need to look briefly at the term “motion” itself. In the text quoted above from *Physics* II.2, Aristotle speaks of ends as outcomes of motion (κίνησις). In his technical usage, however, κίνησις applies only to changes in quality,

---

5 See my n.d., summarized below in section four. Note especially the following passages, to the effect that (a) every end is good: *Phys. II*.2, 194a32–3; cf. *Pol. I*.2, 1253a1–2; and (b) every good is, as such (ἅπλως), a cause for the sake of which: *Metaph. A*.7, 988b6–16; cf. α.2, 994b9–13.
quantity, and place: it cannot signify the coming to be or passing away of substances (*Phys. V.1, 225a34–b3*). Nevertheless, he also clearly holds that the coming to be of living things, for example, takes place for the sake of its outcome.

How, then, should we understand his use of κίνησις in *Physics* II.2? A first point is simply that Aristotle sometimes uses κίνησις in a broader sense, as equivalent to μεταβολή, his term for change in general (see e.g. *Phys. II.7*). Moreover, he does not even introduce the distinction between these two terms until *Physics V.1*, well after both the passage we are now considering, and his well-known definition of motion in *Physics* III.1. The use of κίνησις in the present passage, therefore, may not be significant.⁶

At the same time, Aristotle also seems to allow a sense of κίνησις in which, although not equivalent to μεταβολή, it nevertheless includes the coming to be of natural substances. According to Sarah Broadie, the significance of the distinction between κίνησις and μεταβολή emerges when we consider that whereas the former signifies change as involving a temporally extended process, the latter signifies only the emergence of a new state: change, pure and simple.⁷

Now although I do not think this is what Aristotle had in mind when he

---

⁶ Because “motion” does not quite work in English, I shall use “change” interchangeably with “motion” whenever it is not necessary to make a clear distinction between Aristotle’s κίνησις and μεταβολή; when this distinction is relevant, I shall use the Greek to clarify my English usage.

formulated the distinction, Broadie’s claim does point to the fact that the Greek word for coming to be, \(\gamma\varepsilon\nu\varepsilon\sigma\varsigma\), also has two senses in Aristotle. The sense operative in *Physics* V.1–2 is metaphysical: it reflects Aristotle’s assertion, against Parmenides and the early naturalists, that coming to be and passing away are not illusory. That is, in addition to the motions (\(\kappa\iota\nu\iota\sigma\varsigma\epsilon\iota\varsigma\)) undergone by existing substances, there is also real change (\(\mu\epsilon\tau\alpha\beta\omega\lambda\iota\iota\)) from nonbeing to being (\(\gamma\varepsilon\nu\varepsilon\sigma\varsigma\)) and from being to nonbeing. Because it is impossible to be more or less a substance, however, this sort of change is necessarily instantaneous, and so is not a motion. The second sense of \(\gamma\varepsilon\nu\varepsilon\sigma\varsigma\) appears often in Aristotle’s biological works—in the title of *Generation of Animals*, for example. Here it refers to the extended process by which a new organism comes to be and grows to maturity. In many important respects, at least, this process is surely a motion in the sense that is at work in *Physics* II.2, and is later discussed in III.1.⁸

⁸ This process certainly involves a coming to be in the metaphysical sense, but it is not so clear how we should map the passage from nonbeing to being onto the developmental process described in *Generation of Animals*. My own suspicion is that for Aristotle, an animal’s unqualified coming to be coincides with the formation of the embryo. At the least, he is quite clear that at this point a new substance exists, which although it depends on the mother as an external source of nourishment, strictly speaking nourishes itself and directs its own development to maturity. Although many of its faculties have not yet “come to be,” at no later stage does development involve an unqualified coming to be. It is, rather, a passage from one contrary (immaturity) to another (maturity) through a series coordinated motions. On the ontological status of the embryo after conception see *GA* II.4, 739b20–740a30, especially 740a24–27. In the latter passage Aristotle is generally translated as saying that the embryo is an animal only potentially (see Peck 1942, 197; Platt 1984, 1148). However, the context suggests that he is actually characterizing the embryo as an animal in capacity: “Since in capacity it is already an animal, though incomplete (\(\delta\nu\nu\alpha\mu\epsilon\iota\mu\epsilon\nu\ \eta\delta\eta\ \zeta\iota\omicron\omicron\nu\ \alpha\tau\tau\lambda\epsilon\lambda\epsilon\varsigma\ \delta\epsilon\)) it must receive nourishment from elsewhere; and so it uses the womb and
2. That for the sake of which and change

We have seen that according to Physics II.2, every good outcome of a continuous motion is an end or that for the sake of which. However, the aporia of Metaphysics B.1 brings to light a serious problem with this formulation, and in particular with its reliance on the concept of motion. A first step in resolving this problem, providing a broader context within which to understand the claim of Physics II.2 and similar statements, is to point out that there are other, nontheological reasons for questioning any strong connection between that for the sake of which and change.

Aside from the theological case, Aristotle clearly thinks that teleological explanation can and should be invoked not only in cases of goal-directed change, but also for certain cases of stable existence. Consider, for example, Metaphysics Ζ.17 (1041a28–32). In this passage, he gives a general argument that in composite substances, matter as such is for the sake of form: First, that for the sake of which a composite substance comes to be is its form or essence (τὸ τί ἦν εἶναι). This essence, however, does not lose its causal role when the process of coming to be is finished, but continues in the composite as the ongoing cause of its being. It

the mother to take nourishment, as a plant does the earth, until it is completed to the point of being now an animal with the capacity for locomotion (ἡδὴ ὑπὸν δυνάμει πορευτικόν)."
follows that the cause for the sake of which—unlike, Aristotle adds, the efficient cause—is sought not only for coming to be but also in the case of being.

In other texts, as well, it is clear that form remains that for the sake of which in an already existing substance. The most obvious case is that of living things, whose parts, as organs or tools, both come to be and exist for the sake of the whole (PA II.1, 646a8 ff.; on tools see also Phys. II.3, 194b32–195a3). Moreover, it is not only the organs that exist for an end: according to Parts of Animals II.1, within every animal the elements exist for the sake of the homogeneous parts, and these exist for the sake of the heterogeneous (646a35–b13). In other words, as we have already seen in Meteorology IV.12, Aristotle takes organs to be only the most obvious instance of the relation between matter and form in general, so that in every actually existing composite substance, matter exists for the sake of form. In other words, Aristotle’s teleology must not only account for the coming to be of composite substances, but must also contribute to explaining their ongoing existence.

We have, therefore, a tension that needs to be resolved. On the one hand, in his theoretical discussions of natural science Aristotle introduces that for the sake of which in relation to motion or change. On the other hand, he also thinks

---

9 Or, ultimately, between matter and the substance’s ἔργον. However, the intermediate position of form between matter and ἔργον insures that matter will also be for the sake of form.
in teleological terms about the being—not just about the coming to be—both of natural bodies and, as we saw above, of God. According to some authors, this apparent tension can be resolved by observing that for Aristotle, \( x \) exists for the sake of \( y \) if and only if \( x \) has come to be for the sake of \( y \).¹⁰ Thus “coming to be for the sake of” is conceptually prior to “being for the sake of,” and all teleological claims depend on claims about goal-directed processes or activities. Even at first glance, however, there are serious difficulties with this solution. For one thing, it cannot handle the case of God, and for another, it violates Aristotle’s general claim that being is conceptually and ontologically prior to coming to be.¹¹

In fact, Aristotle himself clearly holds that the teleological character of coming to be is dependent on, rather than determining, the relation between a composite substance’s matter and its form. To see this, we must first remember that in both Physics II.9 and Parts of Animals II.1, the claim that matter is for the sake of form captures a relation that Aristotle also signifies by the phrase “hypothetical necessity,” or “necessity on a supposition.” As he concludes in the Physics, “the necessary is in the matter, that for the sake of which in the account” (II.9, 200a14–5). By extending an example that Aristotle himself uses, we can see that


¹¹ Aristotle holds that coming to be is for the sake of being (see Metaph. Θ.8, 1050a8–10), and this suggests that coming to be for an end is for the sake of being for the sake of that end. Because the end is conceptually prior to that which is for the sake of it, being for the sake of an end turns out to be conceptually prior to coming to be for the sake of an end.
this claim is true of coming to be precisely because it is true of being (200a10–13). The iron in a saw does not exist for the sake of the blade because it came to be as part of the saw; in fact, the smith probably bought it from a miner whose only intention in producing it was to be paid. Iron exists for the sake of the blade only insofar as it is actually part of a saw—or at least is becoming such—because only then does the relation signified by the preposition ἓνεκα obtain. In the terms of *Metaphysics* Z.17, the primary cause of a thing’s being is its form or essence; it is for this—for the sake of this, we may safely gloss—that the matter is hypothetically necessary. Whether its coming to be is also for the sake of this form, however, is quite another question.

It seems, then, that our opening aporia is of considerably greater scope than it might first have seemed. Like God, natural substances and artifacts involve a ὦ ἓνεκα, namely their form, even if we abstract from their coming to be. We therefore need a way to extend Aristotle’s teleology from becoming to being, and to show how the two cases are related. Not surprisingly, since Aristotle states his aporiai with a view to solving (or dissolving) them, the theological conundrum with which we opened contains a clue. In the aporetic argument,

---

12 The homogeneous bodies provide a similar example in the order of nature as opposed to art: although their matter exists in them for the sake of their form (cf. the discussion in chapter one of *GC* II.6), we may presume that this matter, namely the elements, usually does not come to be for the sake of this form, but rather through an independent causal process occurring on some prior occasion.
Aristotle connects that for the sake of which with change through the concept of πρᾶξις ("action" or "activity"). In *Metaphysics* Θ, however, speaking in his own person, he simply rejects the claim that every activity involves motion. With this claim and its consequences in hand, the aporia in all its forms disappears. Our next stop, therefore, is the ninth book of the *Metaphysics*.

In *Metaphysics* Θ.6, having completed his discussion of potentiality (δύναμις) in the sense of a capacity for change, Aristotle goes on to introduce, as he says, the more properly metaphysical concepts of actuality and potentiality. These are, of course, the concepts at work whenever he states that being, in whatever category, may be either potential or actual (e.g., *DA* I.1, 402a23–b1; *Metaph.* Δ.7, 1017a35–b2; Θ.1, 1045b33–4; Θ.10, 1051a34–b1). The term that he uses for actuality is ἐνέργεια, which despite its narrower use in other places is obviously meant in this context to include everything that, elsewhere, he sometimes calls ἐντελέχεια. At the end of the chapter (1048b18–30), turning briefly

---

13 In fact, ἐντελέχεια appears in several of the passages just cited in the text as referring to the distinction between actuality and potentiality. Its occurrence is not surprising in *DA* I.1, given that the being in question is the soul. In *Metaph.* Δ.7, Aristotle uses ἐντελέχεια with examples for which ἐνέργεια would have been quite appropriate. In *Metaph.* Θ.1, he observes that being is spoken of both as something and of such a sort and of such a quantity, and "in potentiality and actuality and in work (κατὰ δύναμιν καὶ ἐντελέχειαν καὶ κατὰ τὸ ἐργον)" (1045b33–4). Because he has just been focusing on the fact that all categories other than substance are referred to substance, this last phrase may actually signify a threefold distinction between a substance existing in potentiality, the same substance existing in (first) actuality through its form, and that substance existing in full (second) actuality through the exercise of its typical activity. As soon as he begins to consider potentiality in the sense of an active power, however, he switches to
to the ontological status of πρᾶξις, he concludes that this class includes both ἐν- ἕργειαι, which count as activities or πρᾶξεις in the full sense of that term, and motions (κινήσεις), which are πρᾶξεις in an extended sense. Motions are those activities, such as making something thin, that involve a process; as Aristotle puts it, they aim at an end outside themselves. By contrast, ἐνέργειαι are activities, such as seeing, in which the end is already present (cf. DA II.5; NE X.7).14

The distinction between πρᾶξεις that are motions (and thus πρᾶξεις in an extended sense) and those that are not is already significant for our purposes. Even further, as the definition of motion developed in Physics III.1–3 makes clear, the distinction between ἐνέργειαι and κινήσεις must be taken as applying more broadly than the term πρᾶξεις strictly allows. First, we have already seen that there is a sense of γένεσις, coming to be, in which coming to be counts as a motion, and thus as an incomplete actuality. Second, although the corresponding complete actuality—the new substance’s first actuality (ἐντελέχεια) or form—

14 Aristotle defines motion, in Physics III.1–3, as an incomplete ἐνέργεια, so that the strict and extended uses of this term parallel those of πρᾶξις. However, as will be noted in a moment, in both senses the former term is notably broader than the latter.
does not count as a πράξις, the term ἐνέργεια can be taken in a broad sense to include all actuality (cf. Metaph. Θ.6). Thus the distinction between complete and incomplete actualities applies in the category of substance as well, and so in all the categories in which change is possible.

In Metaphysics Θ.6, then, Aristotle includes both κίνησις and ἐνέργεια under the general concept of πράξις, the concept used in the aporia of B.2 to connect that for the sake of which with motion. This means that whereas every πράξις does involve an end, not every πράξις—despite the claim made in the aporia—implies motion. Seeing, for example, does not of itself tend to any further end; it simply is the end. Moreover—returning to the priority of being over coming to be—it cannot be said that seeing is an end because it is the outcome of a change from not seeing to seeing,15 or even because animals come to be for the sake of seeing. This can easily be seen from another example. For Aristotle, the activity of the heavenly bodies is circular motion. This motion is their ἔργον, and they exist for the sake of it. However, they do not come to be for the sake of their motion, nor do they pass from not moving to moving; rather, they both exist and move eternally (De caelo II.3, 286a8–13). Their activity—and thus also seeing, in the case of a sighted animal—counts as an end simply because it is an ἐνέργεια.

15 In DA II.5 Aristotle states that this change, not being a change between contraries, is not an alteration (and therefore not a motion) in the ordinary sense of this term. It is, however obviously a change.
Aristotle’s discussion of concepts like πρᾶξις, κίνησις, and ἐνέργεια makes it clear that the concept of that for the sake of which can be separated from that of motion. Whereas the end of a motion is that for the sake of which the motion occurs, a complete activity constitutes an end for its subject simply as such, and apart from any motion with which it might be connected. Moreover, because the distinction between ἐνέργεια and κίνησις can be applied not only to the activity (ἐνέργεια) of a substance, but also to its first actuality (ἐντελέχεια), we can safely conclude that, as Metaphysics Ζ.17 suggests, a substance’s form counts as that for the sake of which not only as the outcome of its coming to be—of a κίνησις, broadly speaking—but also, and in fact primarily, because it is the form or actuality (ἐντελέχεια) by which the substance actually (ἐνέργεια) exists.\footnote{Another interesting text from the Metaphysics occurs in M.3, where Aristotle is also clearly responding to our initial aporia. He argues that goodness, or more properly speaking, beauty, is a cause even among the objects of mathematics. He distinguishes goodness from beauty in that the former always involves πρᾶξις, whereas the latter occurs also in motionless things (1078a31–2). Given that πρᾶξις do not always involve motion, I take the distinction to be between those things to which the concept of actuality can properly be applied, whether or not any motion or change is involved, and mathematical objects, with regard to which we cannot properly speak of potentiality and actuality.}

We began this section by describing that for the sake of which as the good outcome of a continuous change. It now seems, however, that Aristotle’s τὸ οὗ ἐνέκα may have more to do with actuality in general than with changes and their outcomes in particular. Given that the initial connection between that for
the sake of which and change comes from *Physics II*, which is generally considered the most important locus for Aristotle’s teleology, this is hardly a comfortable conclusion. Whatever else may be said for “actuality” as a replacement for “outcome of change,” it leaves us with a good deal of explaining to do. I hope this can be accomplished in the next section, where we shall discuss Aristotle’s concept of limit and its place in his teleology.

Before moving on, however, it is interesting to see exactly how Aristotle’s solution to our opening aporia applies to the theological version of the problem. First—and this is a point worth keeping in mind—in *Metaphysics* Θ.9 Aristotle argues that actuality is prior to potentiality both in account (λόγος) and in substance (οὐσία). This conclusion has a number of consequences, of which the most important for present purposes is that to say God exists actually, or even that he simply is actuality (see Λ.6, 1071b12–23), does not involve even an implicit reference to potentiality of any sort. God does not exist through the actualization of any potentiality; he simply exists. Specifically, the divine substance is an act of understanding that does not differ from its object, so that, because the usual distinctions between substance, activity, and object do not apply, there is simply no room for potentiality in God.
By denying any distinction between the divine substance and the activity by which that substance possesses complete actuality, and between the activity of knowing and its object, Aristotle ensures that God does not exist for any end outside himself. This way of being contrasts sharply even with that of the heavenly bodies, whose activity is distinct from their substance, and which exist for the sake of that activity. Nevertheless, it would be odd to say that God is thereby excluded from the sphere of the good and of that for the sake of which. For Aristotle, rather, he exhibits both characteristics in a far more perfect manner than do natural substances: he simply is his own good and his own end; he exists for his own sake, resting in himself and transcending any need for change. At the same time, as we have already noted, he serves as that for the sake of which other things come to be, exist, and act.

We now have, therefore, a solution to our theological aporia and its variants. As I have already remarked, though, it is a rather uncomfortable solution. In the case of God—but in the case of natural substances as well—we seem to have a striking separation of teleology from motion, and it would be helpful to have at least an initial indication of how this alleged separation fits the many texts in which these concepts are so closely joined. How, in particular, does it affect our understanding of the claim that every good outcome of a continuous...
motion is an end? For now I will suggest only that our investigation so far con- 

firms a claim I have already made about the passage from Physics II.2 and others 

like it: namely, that it should be understood not as an adequate general account 

of τὸ οὗ ἐνεκα, but rather as a set of sufficient conditions for the existence of an 

end. As such, it is all Aristotle requires to establish the need for teleological ex- 

planations in natural science.

In any case, the problems we have encountered thus far make it even 

clearer that we cannot leave formulations like that of Physics II.2 as they stand. 

Adequate as they may be for the Aristotelian naturalist, they cannot be fully un-

derstood without a much deeper analysis, one that takes place at the metaphysi-

cal level. We have already begun such an analysis by relating Aristotle’s concept 

of κίνησις to the related concepts of πρᾶξις, ἐνέργεια, and ἐντελέχεια. In the 

following section, returning to the conditions laid down in Physics II.2, we shall 

consider his use of the term “last” (ἔσχατον), tracing the meaning of this term to 

its metaphysical root in the concept, so important to Greek philosophy, of a limit 

(πέρας).

3. That for the sake of which as limit

To get an initial sense of how important Aristotle’s notion of a limit is for 

our study, let us return to a passage we have already mentioned: the last section
of *Metaphysics* Θ.6, in which Aristotle distinguishes actualities from motions. It reads as follows:

Since of the actions of which there is a limit none is an end, but [they are] of the things that respect the end (like making thin or thinning: the things themselves, when one is making them thin, are therefore in motion, those things not being present for the sake of which the motion is), this is not an action (πρᾶξις), or at least not a complete one (τέλεια), but it is not an end (τέλος). But that in which the end is present is indeed an action.... But if [it is] not [present], it would have had to stop at some point, just as when it makes thin, but as a matter of fact, it does not, but it lives and has lived. So of these we must call the ones motions, and the others activities (ἐνεργείας). For every motion is incomplete (ἀτελής): thinning, learning, walking, house-building—these indeed are motions, and incomplete (*Metaph. Θ.6*, 1048b18–30).

Although the term “limit” appears only once in this text, the corresponding concept obviously plays an important role. Moreover, the definition of motion as an incomplete action—one that has a limit distinct from itself—already suggests that the ἔσχατον of *Physics* II.2 is an end because it constitutes the limit of the preceding motion. Finally, in line with what I have already claimed concerning motion and that for the sake of which, Aristotle recognizes certain actions, which he calls complete, in which there is an end without any preceding motion.17

---

17 On the complete as that to which the end belongs see *Metaph. Δ.16*, 1021b23–30. That Aristotle associates the complete with the bounded or limited, as opposed to the unlimited (ἀπέρον), appears from the fact that he deals with the complete and with limit in consecutive chapters (Δ.16–7), and characterizes them in similar terms.
The connection in *Metaphysics* Θ.6 between limits and ends suggests that a further examination of what Aristotle means by “limit” may not only explain the role of the ἔσχατον in *Physics* II.2, but also shed further light on how there can be a οὗ ἕνεκα that is not, temporally, something last. For if it is now clear that τὸ οὗ ἕνεκα is not necessarily an outcome of change, we are still looking for an explanation of what, in fact, it is, of its importance in Aristotle’s philosophy, and of the relation between goal-directed change and Aristotelian teleology in general. We can find the initial leads for our investigation in the chapter from Aristotle’s metaphysical lexicon devoted to the concept of limit, *Metaphysics* Δ.17. Because the chapter is brief and already quite compressed, I reproduce it here in full:

We call a limit the last of each, and the first outside of which is found nothing, and the first within which all, and whatever may be the form of a magnitude or of what has magnitude, and the end of each, such being that to which is the motion and the action, and not that from which, but sometimes both, namely that from which and that to which and for the sake of which, and the substance of each and what it was to be for each, for this is the limit of knowledge, and if of knowledge, also of the thing. So it is clear that limit is spoken of in as many ways as is principle, and in more still: for the principle is a sort of limit, but not every limit is a principle (1022a4–13).

For present purposes, two points stand out in this chapter. The first is that Aristotle refers to “that to which” not only as a limit, but as that for the sake of which a motion occurs. Even more interesting, however, is his characterization of substance (οὐσία) and what it was to be (τὸ τί ἦν εἶναι) as a limit.
In section two, I concluded that a substance’s form counts as an end not only as the outcome of its coming to be, but also, and even primarily, as the actuality by which it exists. In *Metaphysics* Δ.17, we find a striking parallel. Aristotle not only calls the endpoint of a motion its limit; he also characterizes as a limit—the limit of a composite substance—the form or “what it was to be” of that substance. To understand how substance (οὐσία) and what it was to be (τὸ τί ἦν εἶναι) enter his discussion of limits, and how such limits are related to the endpoints of motions, we must draw on his account of substance in *Metaphysics* Z–H.

I shall not, of course, propose an overall interpretation of these books and their conclusions, but shall take from them a few helpful points, which I hope are sufficiently uncontroversial not to require an extended defense.

It is well known that Aristotle uses οὐσία in a number of senses, and commentators are divided about exactly where the distinctions and arguments of *Metaphysics* Z–H leave us. Clearly, however, the sense at work in Aristotle’s discussion of limit is that in which οὐσία signifies a thing’s essence, or in Aristotle’s roundabout phrase, what it was to be (τὸ τί ἦν εἶναι) for that thing. Moreover, it is also clear that in the case of matter–form composites, with which Aristotle is directly concerned in *Metaphysics* Z–H, substance in this sense is the composite’s form or actuality (see e.g. Z.6, 10–11, 15–17), which he elsewhere says is the cause
of its being (Δ.8, 1017b14–16; DA II.4, 415b10–22). Next, the phrase τὸ τί ἦν εἶναι
has a number of derivative senses as well: it can be applied to qualities, to quan-
tities, and indeed to anything predicatable of a substance, insofar as these are, sec-
ondarily, what it was to be (τὸ τί ἦν εἶναι) for their subjects (Z.1, 4). A final
relevant point that becomes clear in Metaphysics Z is that substance, in the above
sense, is what we are trying to capture when we define something (see e.g. Z.4).

After all, a definition is meant to answer the question “What is it?” (τί ἐστι), and
substance is what it was to be (τὸ τί ἦν εἶναι) for the object in question. To know
something simply speaking, therefore, is to know its substance, which thus con-
stitutes the primary object of knowledge.

In Δ.17, then, it is presumably this last point, concerning substance as an
object of knowledge, that Aristotle has in mind when he says that a thing’s οὐσία
or τὸ τί ἦν εἶναι is the limit of knowledge. It is, after all, the object of knowledge
that makes an act of knowledge be knowledge of a particular kind, and therefore
something definite and limited. Aristotle goes on to say, however, that if sub-
stance is the limit of knowledge (γνῶσις), this is only because it plays an analo-
gous role in the thing (πρᾶγμα) itself. Thus the concept of a limit (πέρας) seems

18 Correspondingly, Aristotle occasionally uses a broader sense of οὐσία that enables him
to ask about the “substance,” or essence, of an item from any of the categories (A.3, 983a27; 7,
988a34–b6).

19 Although it would be dangerous to rely on this text, note that in Cat. 2, being knowl-
edge of grammar limits knowledge in the same way that being a horse or a man limits animal.
to provide him with another perspective on the causal role of a composite’s οὐ-
σία, or form. Substance is the cause of a thing’s being insofar as—by contrast
with matter, the principle of potentiality—it determines or limits what the com-
posite actually is. This same point, in fact, appears in other terms in H.2, where
Aristotle characterizes substance, in the sense of actuality, as a differentia or dif-
ferentiae of substance in the sense of matter. Finally, if the terms “being” (εἶναι)
and “what it was to be” apply in a derivative sense to nonsubstances, it is just as
clear that quantities, qualities, and so forth are secondary limits on, or determina-
tions of, their subjects’ being.

Interesting as it may be, however, the foregoing account of substance as
limit does not yet shed much light on Aristotle’s teleology. What we need now is
to connect this notion of a limit of being with the notion of a limit or endpoint of
motion, which we found in Metaphysics Θ.6. As we shall see, the relation between
limits of motion and limits of being turns out to correspond perfectly with the
relation, which we have already considered, between coming to be for the sake of
something and being for the sake of something. This correspondence points to
the conclusion that in Aristotle’s usage, the terms τὸ οὗ ἐνέκα, ἐνέργεια, and πέ-
ρας designate complementary concepts, each of which reveals a different aspect
of the same metaphysical principle.
Given what we already know about the place of form in Aristotle’s teleology, it should come as no surprise that limit as end and limit as form are indeed closely related. The necessary link appears in any number of passages in which he discusses the various types of cause, both in *Physics* II and elsewhere. He repeatedly states that in the central cases of goal-directed change, namely, the coming to be of natural substances, that for the sake of which is nothing other than the form or substance—τὸ τί ἦν εἶναι—of what comes to be (see e.g. *Phys.* II.2, 194a27–b9; II.7, 198a24–7; *GC* I.7, 324b14–20; II.6, 333b5–20; II.9, 335b6–7; *Mete.* IV.2, 379b18–380a1 *DA* II.4, 415b10–22; *PA* I.1, 639b15–7). Moreover, this is no accidental connection. In *Parts of Animals* I.1, when he explains why his predecessors failed to make use of teleological explanation, the reason he cites is that they had no clear notion of “what it was to be, of definition with respect to substance” (642a25–7). The terms he uses are, of course, τὸ τί ἦν εἶναι and οὐσία.

Of the various texts in which Aristotle identifies form (the limit of being) with that for the sake of which (the limit of coming to be), it is this remark about the failure of the early naturalists that best reveals the nature of the connection. To explain how a substance is generated, he thinks, is to explain how it comes to be what it is, or in other words, how it comes to have the form that determines or
limits it. By failing to understand that form, rather than matter, is the cause of a substance’s actual being, the naturalists missed not just the explanation, but the explanandum itself. In other words, mistaking the ontological structure of the outcome, they also mistook the structure of the process. Aristotle summarizes the order of inference nicely when he says, in *Generation of Animals* V.1, that coming to be (γένεσις) follows upon and takes place for the sake of being (οὐσία), rather than vice versa (778b6–7; cf. *Metaph.* Θ.7). Thus we must conclude that οὐσία is the defining limit of coming to be because it is the defining limit of that which is.

Once we understand form as the limit of a composite substance, and therefore as the cause of its being, we can see clearly one reason that Aristotle considers τὸ οὗ ἑνεκα to be a cause of being, and not only of motion or change.

By contrast with processes and tools, which are for the sake of an end but cease to be relevant to that end once it is achieved, a substance’s matter is necessary for the finished product as well. This is why in *Physics* II.9, to which I alluded in the previous section, Aristotle can speak indifferently of the relation between matter and form in coming to be, and their relation in being:

---

20 At one level of analysis it may be preferable to say that form determines or limits matter—but in so doing, it is the internal principle by which the concrete substance exists actually in a determinate way. Thus either way of speaking is correct.

21 See Meyer 1992, 820.
For example, why is the saw such? So that this and for the sake of this. It is impossible, however, that this end come to be unless it be iron; therefore it is necessary that it be iron, if a saw and its task shall be. It is necessary, indeed, on a supposition and not as an end: for the necessary is in the matter, and that for the sake of which in the account (λόγος) (200a10–15).

The account, form, or substance is that for the sake of which not only as cause of the saw’s coming to be, but also as cause of its being.

The identity between limit as essence and limit as end in the case of natural substances and their coming to be brings us an important step closer to understanding Aristotle’s teleology. To take full advantage of this step, however, we need to place it in the context of his general understanding of change. In Physics III.1–3, when Aristotle is developing his definition of motion in the broad sense, or change, his central claim is that change is not a reality over and above the various categories of being. It is, rather, an incomplete actuality corresponding to the potentialities and complete actualities found in the most basic of these categories. To understand the significance of this claim, however, and the role the concept of limit plays in explicating it, we need a minimum of historical context.

One of Aristotle’s most important contributions to Greek philosophy is his proposal of a strong competitor to the Parmenidean and Heraclitean understanding of change, which proceeded exclusively by contrasting change with stable
being. To many previous thinkers, including Plato, change seemed to be something so indeterminate, so close to nonbeing, that it was almost (or even in fact) illusory—although of course the nature of the “illusion” differed from one thinker to the next. Aristotle, however, by insisting that change be understood in relation to being, concludes that it is not in fact something completely indefinite (ἀόριστον), as these earlier thinkers had alleged (cf. Phys. III.2, 201b24–33).

Rather, every motion is, as such, defined by certain limits. Thus at the end of his lengthy discussion of change in Physics V–VI, he writes:

> No change is unlimited (ἀπειρος), for each was from something to something, both that in a contradiction and that in contraries, so that of those according to contradiction, the affirmation and the negation are a limit, such as being of coming to be, and nonbeing of passing away, and of those in contraries, the contraries.... That change, therefore, is not unlimited in such a way that it is not defined by limits, is clear (Phys. VI.10, 241a26–241b12).

Change between contraries is, as Aristotle has already explained (Phys. V.1), motion in the strict sense, in quality, quantity, or place. Change between contradictions includes the coming to be and passing away of substances, and probably also the transition between activity and inactivity in the case of sight, understanding, and so forth (see DA II.5, 417b2–9, where Aristotle denies that the passage from habitual knowledge to reflection is an alteration in the normal sense of
that term). In each case, there are limits from which and to which, and these limits account for the definite character of the change.  

In many respects, Aristotle’s claim that every change is defined by limits is relatively straightforward. However, there are two cases that require special consideration. The first is passing away, the limit of which Aristotle explicitly says is nonbeing (τὸ μὴ ὄν). The difficulty here is that as we have seen, Aristotle’s concept of limit seems intended, in most cases at least, to capture the well-defined or bounded character of actual being—and, derivatively, of change. How, then, can nonbeing serve as a limit? It certainly does not seem to provide the same kind of definiteness that Aristotle wants from limits in general. In response to this difficulty, I think Aristotle would simply admit that the limit of passing away is a limit or boundary in an extended sense, which we should think of in the same way as we sometimes think of nonbeing as “something.” Moreover, he would be quick to point out that the nonbeing in question is by no means absolute. Rather, each thing passes away into its own proper matter,

---

22 Although he admits elsewhere that circular locomotion has no beginning or limit or middle simply speaking (De caelo II.6, 288a22–4), Aristotle does not allow in Physics VI.10 that it is an exception to his claim that no motion is indefinite in the sense of not being defined by limits. Rather, he contrasts “infinite motion,” which has no definite goal, to circular motion, which is quite well defined (241b2–11, 18–20). In De caelo, moreover, he states that such motion is complete in a way that rectilinear motion cannot be (I.1, 269a19–24). The possible tension among these claims seems to be resolved by the statement of Physics VIII.8 that in circular motion alone, the starting-point (ἀρχή) and the limit (πέρας) coincide.
which is nonbeing only insofar as it now lacks the form it once had. As he puts it in *Metaphysics* Z.10, “some things are from, as principles, these [parts] into which they pass away … and so the clay statue passes away into clay, the sphere into bronze, and Kallias into flesh and bones” (1035a24–5, 31–3).\textsuperscript{23}

The second difficult case is one that I have already mentioned, namely, the passage from inactivity to activity. The problem here is that whereas it is not difficult to think of a substance’s form as its defining limit, and its nonessential attributes such as place, size and color as secondary limits, it seems odd to call the activity of seeing, for example, or of building a house, a limit. These activities seem, rather, to be unlimited in precisely the way that motion and coming to be are not. For although both seeing and house-building have their own limits in the object of sight and in the house, they are not clearly secondary limits on the agent. At the least, it seems odd to characterize as a limit the exercise of a thing’s natural powers, especially insofar as this exercise appears to be relatively open-ended. This difficulty brings us back to the passage with which we began our study of limits, in which Aristotle distinguishes activity from motion; its solution is implicit in the same text.

\textsuperscript{23} In *GC* II.3, Aristotle articulates the general point that the passing away of one thing is always the coming to be of another (318a24–8). The comment in *Metaph.* Z.10 reflects his specific account of unqualified passing away.
In *Metaphysics* Θ.6, Aristotle says that ἐνέργειαι do not have ends: in such an activity or actuality, the end is already present. The question we must ask ourselves is, of what is such an actuality the end? Aristotle provides the answer two chapters later, when in discussing the ontological priority of actuality to potentiality he states that in general, “actuality is an end, and it is for the sake of this that potentiality is acquired. For animals do not see so as to have sight, but have sight so as to see, and likewise the art of building so as to build, and the habit of contemplation so as to contemplate” (Θ.8, 1050a9–12). Activity, therefore—as the function argument of *Nicomachean Ethics* I.7 also makes clear—is the end of a substance’s existence; it is for the sake of its complete activity that each thing comes to be and exists. If this is so, however, then clearly this activity is the limit of a substance’s being, perhaps even in a stronger sense than is its form. Correspondingly, we now have another sense in which that for the sake of which—in this case the activity—is a cause not only of coming to be, but also of the substance’s being.

In the foregoing discussion, we have seen how Aristotle’s concept of limit informs and enriches his concept of actuality. By way of concluding this section, I now want to make a similar point about the effect of the latter concept on the former, emphasizing the positive character of Aristotle’s πέρας. To this end, a
final brief return to his theology may be helpful. If the concept of limit is truly coextensive with that of actuality, it should apply to God as well. As with the concepts of actuality, goodness, and that for the sake of which, however, this application requires a bit of work: in a substance not composed of matter and form, there would seem to be nothing for a limit to limit.

Aristotle’s God, we know, is pure intellectual activity (ἐνέργεια), and so the remarks we have just made about activity as limit would seem to be relevant to determining how the concept of limit might apply to him. On the other hand, God is a substance, and this point returns us to our previous discussion of οὐσία as limit. Rather than beginning immediately from either of these discussions, however, I want to bring in another text, one that highlights the importance of the concept of limit for Aristotle’s metaphysics as a whole. It appears in *Metaphysics* α.2, which argues “that there is a principle (ἀρχή τις), and that the causes of beings are unlimited (ἄπειρα) neither in a straight line nor in kind” (994a1–2). Among the various sorts of indefiniteness or infinity—the appropriate translation depends on the subject matter—that he rules out, it is his discussion of essence, or “what it was to be,” that most interests us at the moment.24

---

24 His comments on that for the sake of which are also relevant, although they add nothing new to our discussion: “Further, that for the sake of which is an end, and one such that it is not for the sake of another, but the others for that, so that if something such will be last, it will not
But indeed neither does what it was to be admit of being reduced to another definition fuller\textsuperscript{25} in account: for it is always rather the prior [that is the definition], whereas the subsequent is not; and where there is no first, there is no next. Moreover, those who speak thus destroy science (ἐπίστασθαι), for to know (εἴδεναι) is nothing but to come to the indivisible; and there is no acquaintance (τὸ γνῶσκειν) with anything, for how can one think (νοεῖν) things that are thus indefinite (994b16–23)?

In these lines, Aristotle argues that in order for τὸ τί ἦν εἶναι to serve as a limit of knowledge, it must have a finite number of basic parts or aspects that one can keep in mind and state in a definition. Presumably, he wants to impose the same requirement on τὸ τί ἦν εἶναι as a limit of being. Every form, simply as such, must confer a definite character on the substance whose form it is. Otherwise, what we have is not a substance, but an indefinite who-knows-what, which requires further determination in order to count as “something.” In short, when Aristotle requires that “what a thing is” not be indefinite (ἄπειρον), he is simply requiring that it really constitute a determinate way of being, providing a complete answer to the question “what.” Only a definite (non-ἄπειρον) essence can be, in turn, a limit for the composite whose essence it is, and for our knowledge of that composite.

\textsuperscript{25} The correct translation of πλεονάζονται is probably more like “going overboard.”
Considered in this way, the ontological role of form or actuality as the source of definiteness in a composite—and, conversely, the interpretation of limit as actuality—suggest an important clarification of the concept of limit. In our previous discussion of substance as limit, I focused on the limiting role of form with respect to matter: matter as such is something indefinite, and must be further determined to a specific, actual way of being. From this perspective, limit may appear most clearly as a boundary: a limit excludes ways of being that do not pertain to the composite whose limit it is. The text just quoted from the *Metaphysics*, however, reminds us that this exclusion is something secondary. The primary significance of actuality as limit is not the exclusion of other ways of being, but the positive and definite character of what truly is. This positive character of actuality as limit appears even more clearly if we recall that a substance’s limit in the most basic sense is not its form, but the complete activity that the form makes possible and in terms of which it must be understood.

Focusing thus on the positive character of limit is helpful in several respects. First, it helps us understand the priority of actuality to potentiality in

---

26 Aristotle’s most complete discussion of these issues appears, of course, in *Metaph.* Z.12–H.6. Although I have deliberately passed over the nuances of that discussion, it seems to me that many of Aristotle’s concerns in these chapters and in related texts (e.g., *An. Post.* II.7–10, 13; *PA* I.2–3) reflect his awareness that to be fully actual (which is clearly something positive) and to be fully determined are correlative, rather than conflicting ways of being. I am thinking, for example, of his rejection of universals as substances, his concern with the unity of definitions, and his rejection of negative differentiae.
Aristotle’s thought, making clear that this priority includes a priority of the determinate to the indeterminate, of that which has its own definite way of being and so can serve as a principle of being for other things, to that which is undetermined and so stands in need of something outside itself. Second, the positive character of limit sharpens our understanding of the relation between first and second actuality, revealing that first actuality determines a substance’s way of being by directing it to the activity that constitutes the ultimate limit, the defining actuality, of its being. Finally—to conclude our theological excursion—it reveals how the concept of limit applies in the case of God, for it is now clear that for Aristotle, what is fully in act is necessarily also fully determinate. It is not that God is limited, but rather that he is fully determinate without need of anything outside himself. Thus also his immutability and necessity, his self-sufficiency and, for Aristotle, his goodness.

The concept of limit will appear again in chapter three, especially when we consider how limits are involved in the teleological structure of changes outside the category of substance. Already, however, we have discovered a great deal. Through a series of conceptual simplifications, moving from τέλος to ἔσχα-τον and from ἔσχατον to πέρας, we have arrived at a concept that is central to Aristotle’s understanding of the being and activity of natural substances. The
concept of a limit is not just an extra, a useful idea for expressing what Aristotle has other perfectly good ways to say. It is, rather, at the heart of his metaphysics. By directing us from change to the limits of change, from becoming to being, and by expressing the well-defined, bounded character in virtue of which that being is a limit not just of change, but of the actually existing bodily substance and of our knowledge thereof, the notion of a limit reveals the entire natural world, the Heraclitean and Platonic world of becoming and opinion, as a world of being and understanding.

4. That for the sake of which as good

In section one, considering Aristotle’s introduction of that for the sake of which in *Physics* II.2, we distinguished three conditions for counting something as an end: (a) that it be an outcome of motion; (b) that the motion be continuous; and (c) that the outcome be good. In section two, the first of these criteria proved problematic, given that Aristotle considers that for the sake of which to be a cause of being as well as of becoming, and also that he considers God good, and therefore something for the sake of which. The need to resolve these difficulties led us to the tentative conclusion that in general, it is actuality or actual being that serves as that for the sake of which in Aristotle’s teleology.
In order to check this conclusion against our original passage from *Physics* II.2, we turned in section three from the connection between that for the sake of which and motion to the closely related characterization of that for the sake of which in terms of what is last (ἔσχατον). This term pointed us toward Aristotle’s concept of a limit, and we found, first of all, that a substance’s form is the limit or outcome of its coming to be precisely because it is the limit of the existing substance itself. Moreover, I have suggested that a substance’s complete activity, as the fulfillment of its natural potentialities, constitutes the ultimate end or limit of its being. In short, we have reached the same conclusion as in section one, but under a new aspect contributed by the concept of a πέρας.

So far, then, we have been probing in various ways the relation between Aristotle’s τὸ οὗ ἔνεκα and outcomes of motion; in doing so, our attention has shifted away from change to the metaphysical concepts of limit and actuality. The final task of this chapter is to see how goodness fits into the understanding of τὸ οὗ ἔνεκα that has emerged in the previous sections. We thus pass from condition (a) to condition (c): the requirement that an end or that for the sake of which be good. Our consideration of this point can be much briefer, for the pieces fall into place as soon as we understand what Aristotle means by “good”
(ἀγαθοῦ). I have analyzed his conception of goodness elsewhere,27 moreover, and so will limit myself here to summarizing the main points of that discussion.

First, then, as mentioned in section one, Aristotle’s claim in *Physics* II.2 that the end is something good is by no means an isolated statement. It is repeated, together with its converse, on a number of occasions. That for the sake of which is always some good, and the good as such is that for the sake of which.28

Second, there are two types of things that Aristotle typically identifies as ends. These are, as our discussion of limits has already suggested, the form of a substance and its activity. Form and activity are in turn, however, precisely the two things that Aristotle elsewhere directly identifies as good. On the one hand, within a substance’s essential constitution—that is, considering those things that are required for its very existence—it is form rather than matter that qualifies as good.29 On the other hand, a substance is not fully good simply by existing, but

---


28 In addition to *Phys*. II.2, 194a32–3, see II.7, 198b5–9; *Metaph.* A.7, 998b6–16; a.2, 994b9–13; B.2, 996a23–7; *Pol.* I.2, 1253a1–2. Given the arguments of chapter one, Aristotle’s identification of end and good means that the good will be found throughout nature. In fact, as noted in section one, Aristotle thinks that the good is a universally applicable causal principle: see *Metaph.* A.2, 982a19–b10; A.3, 984b8–22; a.2, 994b9–16; A.7, 10; M.3, 1078a31–b5.

29 See GC II.6, 333b16–9, a passage that we have discussed at length in chapter one, as well as *Phys*. I.9, 192a16–25, in which Aristotle states that form as such is “divine, good, and desirable.”
only by its activity, through which it fully actualizes the potentialities it has by nature.30

Third—as we have already seen—form and activity are not unrelated; in fact, Aristotle recognizes that they are both actualities and even makes an explicit distinction between actuality of the first type, or form, and actuality of the second type, or activity.31 Thus the connections between goodness, on the one hand, and both form and activity, on the other, add up to an identification between goodness and actuality.

Finally, this identification of goodness and actuality is independently supported by Aristotle’s frequent claims to the effect that being is better than nonbeing—and thus also that necessity, eternity, causal agency, and so forth are good-making attributes32—and by the claim that actuality, as such, has the character of an end and that for the sake of which.33

Aristotle’s identification of goodness with actuality obviously has important consequences for the normative value of his teleology—or perhaps better, of

30 That complete goodness is found in activity is a premise of the well-known “function argument” of the Nicomachean Ethics (I.7, 1097b22–1098a18); other relevant passages include NE I.1, 1094a3; I.7, 1098a5–7, 1098b32–1099a3; VI.2, 1139a16–7; De caelo II.12, 292a19–292b25; Mete. IV.12; DA II.4, 415a27–b8; De somno 3, 455b13–28.

31 See DA II.1, 412a20–9.

32 GC II.10, 336b25–337a1; PA I.5, 644b22–6; GA II.1, 731b24–732a10; Metaph. Α.7, 1072b10–30; Ν.4, 191b16–21. See also, on the connection between goodness and being, NE I.6.

33 Metaph. Θ.8, 1050a4–10.
the entire metaphysics in which this teleology is embedded. For our purposes, however, little more need be said. Considering the criterion of goodness supports and enriches the conclusions we have already reached: Goodness is actuality; actuality is a limit; and all three are that for the sake of which.

In the preceding pages, I have argued that the teleology pervading Aristotle’s natural science must ultimately be understood in metaphysical terms. This is because the incomplete actuality that is change, and that forms the subject matter of natural philosophy, must ultimately be understood in terms of complete actuality in the various categories, and especially in the category of substance. What a teleological account of natural change captures, therefore, is the ontological role of being, or οὐσία, in the ordered transformations that make nature what it is. Moreover, recognizing the fundamental importance of being and its relation to becoming in Aristotle’s teleology has enabled us to see how that teleology applies to a necessarily unchanging substance, God, and also to natural substances in their being as well as in their coming to be.
CHAPTER THREE

TELEOLOGY AND CONTINUOUS CHANGE

In chapter two, we examined at length the metaphysical foundations of Aristotle’s teleology, focusing on the concepts of actuality and limit. In the present chapter we move back from metaphysics to natural philosophy, turning our attention to goal-directed change. As before, however, I shall proceed by gradually clarifying the concepts underlying Aristotle’s teleological claims. Our starting point this time is condition (b) of Physics II.2: that to be for the sake of an end, a change must be continuous. In the pages that follow, I argue that this simple condition is in fact, by itself, a sufficient condition for goal-directedness. Striking as it may seem, this conclusion emerges from a careful analysis of Aristotle’s concept of continuity, particularly continuity in motion. In the course of this analysis we shall see that for Aristotle, a change is continuous if and only if it has, as a whole, a καθ’ αὐτό agent cause. A καθ’ αὐτό agent is one that tends by its nature to produce an outcome of a particular sort; when it does so, therefore, this outcome is an end.
My argument proceeds in two stages. In the first, I establish the connection between continuity and goal-directedness in the case of motion in the strictest sense: that is, for simple changes in quantity, quality, or place. In the second, I show how continuity, and so also goal-directedness, are found in complex changes such the coming to be of a living thing.

With regard to the first point, I begin in section one by considering Aristotle’s discussion of continuous motion in *Physics* V. By the end of this section we shall have an initial argument that continuity is a sufficient condition for goal-directedness. In section two, I develop this connection between goal-directedness and continuity by considering Aristotle’s claim that motions are individuated in terms of continuity: that is, to be a single motion is nothing other than to be continuous. In this section, also, the concept of καθ’ αὑτό agency emerges as the essential link between continuity and goal-directedness.

However, because Aristotle’s discussion of motion in *Physics* V–VI is concerned only with simple motions, the scope of our conclusions in sections one and two will be somewhat limited. In section three, therefore, moving on to texts from *Generation and Corruption* and *Parts of Animals*, I show how the notion of continuity can be extended to include complex changes such as the coming to be of a living substance. In section four, I go on to examine the unity of these
complex changes, and to show how this unity, established by the presence of a καθ’ αὐτό agent cause, is the condition of their goal-directedness.

Throughout the chapter, I gradually lay the foundations for a sound interpretation of Aristotle’s two best known arguments concerning natural teleology, namely, the first two arguments of Physics II.8. I consider the argument based on the occurrence of natural phenomena “always or for the most part” in section two and again in section four, and the analogy between nature and art in section four. We shall thus see that on the interpretation of goal-directed motion developed in this chapter, the arguments of Physics II.8 can be read in a straightforward and highly plausible manner.

1. Continuous motion

To understand what Aristotle means by continuity in motion, we must return to his claim that every change is defined by limits. More precisely, as he explains in Physics V.1, the beginning and end of a given change may be opposed to each other either as contradictories or as contraries (224b28–9). Thus coming to be and passing away are defined by contradictories, whereas the limits of motion in the strict sense, in quality, in size, or in place, are contraries (225b7–9). Often, of course, such limits are not absolute contraries—the extremes in the genus to which they belong—but they are always at least relative contraries, through their
differing proximity to these extremes (224b30–35). For example, not every change in color is from white to black or from black to white, but all such changes exhibit a basic structure that appears most clearly in motions involving these extremes.

Although Aristotle first mentions continuous motion in *Physics* III.1 (200b16–18), whose topic is motion in the broad sense, his eventual discussion of continuity in V.3 depends on the distinction between change defined by contrar- ies and that defined by contradictories. He first characterizes continuous motion as follows:

The last point (ἔσχατον) of the change is the contrary, and what the changing thing naturally arrives at, changing naturally and continuously, before it changes to what is last, is between. What is moved continuously (συνεχῶς) is what leaves no gap (τὸ μηθὲν … διαλεῖπον), or the least possible, in the πρᾶγμα—not in time (for nothing prevents things leaving a gap, and straightaway after the lowest the highest being sounded) but in the πρᾶγμα in which it is moved (226b26–7, 23–5, 27–31).¹

For a motion to be continuous, then, there must be intermediate states between its beginning and its end. It is both significant, and perplexing for our purposes, as we shall see later, that for coming to be and passing away *tertium non datur*:

Since every change has to do with opposites (ἐν τοῖς ἀντικειμένοις), and the opposites are contraries and those by contradiction, and a contradiction has no middle, it is clear that what is between (τὸ μεταξὺ) is between contraries (ἐν τοῖς ἐναντίοις) (227a7–10).

¹ Lines 226b23–227a10 of this chapter retain Bekker’s numbering, but the passage was reconstructed by Ross for the OCT edition as follows: 226b23, 227a7–10, 226b26–7, 226b23–5, 226b27–35, 227a1–6, 227a10.
The passage from non-being to being or from being to non-being is immediate; there are no intermediate states.

Let us examine this account of continuity more deeply, beginning with the observation that continuity in change, for Aristotle, depends on continuity in something he calls the \( \pi\rho\alpha\gamma\mu\alpha \) of a change. The term \( \pi\rho\alpha\gamma\mu\alpha \) is as flexible as our “thing,” and is likewise often used to denote the reality corresponding to a word or phrase. In this case, however, it refers to the generic attribute—color, for example—with respect to which something changes. In the case of motions, the \( \pi\rho\alpha\gamma\mu\alpha \) involves a continuum of possible attributes belonging to a common genus and bounded by a pair of contraries. Color, whose specific instances constitute a continuous spectrum limited by black and white, is one example. In the passage just quoted Aristotle uses pitch, which varies continuously between low and high. A third example, which turns out to be the most useful, is heat. In each case, however, continuity involves two things. First, the specific qualities between the extremes are arranged in succession (\( \varepsilon\phi\varepsilon\zeta\eta\zeta \)), by a progressively greater difference from one extreme and similarity to the other. Second, any two successive parts of this range share a limit (\( \text{Physics V.3, 227a6, 10–12} \)).

A continuous motion, then, is one that leaves no gap (or the least possible) in the \( \pi\rho\alpha\gamma\mu\alpha \) defined by the limits of that motion. As we have already
observed, however, in the case of a change defined by contradictories there is no such πρᾶγμα. The two possible states are being and non-being; like an electric light with no dimmer switch, there is no intermediate between on and off. Continuity, therefore, as Aristotle defines it in *Physics* V.3, does not apply to coming to be and passing away. This conclusion presents an obvious problem for the description of goal-directed motion in *Physics* II.2. On the one hand, the most obvious and impressive case of change for the sake of an end is the coming to be of living things; on the other, such change, understood as coming to be, cannot be continuous in the strict sense. Later, we shall resolve this difficulty by finding in other texts a broader sense of the term συνεχής. At the moment, however, the strict sense of continuity raises another interesting challenge: the claim, now seemingly implicit in *Physics* II.2, that even so basic a motion as becoming hot, provided that it take places continuously and that the outcome be good, is for the sake of that outcome.

To pursue this suggestion and its implications, we must first consider more carefully the relation between motions and their πρᾶγματα. For as Aristotle’s point about low and high notes makes clear, the right sort of πρᾶγμα is a necessary but insufficient condition of continuity in motion. Only if the successive stages of the motion correspond to those of the πρᾶγμα—if, as Aristotle
puts it, the subject passes from one contrary to the other without skipping anything in between—will the motion be continuous. Metaphysically, this means that for any two successive stages of the change, the εἶδος that defines the earlier defines the later as well, serving as both that to which and that from which. In an actually continuous motion, of course, such stages are only potentially distinct, and the εἶδος that defines them does not actually serve as a limit. An actual pause in time, as Aristotle later points out, would destroy the continuity even without any jump in the πρᾶγµα (V.4, 228b3–7).

Moreover, although Aristotle chooses as his example an instance of change in which discontinuity is possible—think of the flautist who plays one note immediately after another—there are other motions in which discontinuity of this sort seems to be excluded. Such motions are, as we shall see, by far the most interesting, and may well include all causally basic motions in Aristotle’s cosmos. To take an example, we may assume—in the absence of any evidence

---

2 For example, Aristotle’s account of sound commits him to the claim that changes in pitch are causally dependent on changes in the mechanical means by which air is set in motion, such as the fingers of a flautist or the vocal chords of a singer (see DA II.8). Furthermore, Aristotle analyzes all changes of this latter sort in terms of heating, cooling, and locomotion. Finally, it is hard to read Phys. and GC without concluding that changes of these last three sorts are necessarily continuous. Thus, it turns out that a discontinuous change in pitch is causally dependent on continuous changes in temperature and in place. More generally, a careful reading of GC suggests, to me at least, that all bodily changes in Aristotle’s cosmos (where bodily changes are those not depending intrinsically on soul) should be analyzed in terms of locomotion, heating and cooling. This would mean that all causally basic changes are necessarily continuous. This conclusion would be a significant one for what follows, for if a motion only happens to be continuous, it is
that Aristotle had a quantum theory of heat—that an Aristotelian body can become hot only by a continuous motion from its initial relative coldness. As with any continuous motion, therefore, the heated body passes through infinitely many intermediate degrees of heat, each one of which is a potential endpoint. If the motion is necessarily continuous, moreover, then the attainment of every such potential limit is a prerequisite to the attainment of the actual endpoint. The motion is cumulative in such a way that each degree of heat is necessary for reaching the next.

There is a bit more to be said about continuity in Physics V, but we have enough already to consider its relevance to Aristotle’s teleology. As already observed, it follows from Physics II.2 and Physics V.3–4, taken together, that even the simplest continuous motion is for the sake of its outcome—provided, as II.2 adds, that this outcome be good. To press the importance of continuity even further, however, I want to consider before proceeding some evidence that the condition of goodness is—though perfectly correct on Aristotle’s view—superfluous. In Physics II.2, Aristotle begins by claiming that every outcome of a continuous motion is an end. Only after reporting the poet’s “absurd” mistake does he add that this outcome must be good. Now although we have not yet discussed the
doubtful that its earlier stages are hypothetically necessary for what follows. This would, of course, weaken or even refute the claim that the earlier stages are for the sake of the later and of the outcome.
sense in which an animal’s development counts as a continuous change, let us assume for the moment, quite reasonably, that its decline to death is not in any relevant sense continuous with its development to maturity. One follows the other, and that is all. This would mean, however, that the continuity condition already excludes the poet’s mistake. This argument may not be enough to show that goodness is not a distinct necessary condition for goal-directedness. Nevertheless, to catch the poet out with the continuity condition alone would be suggestive—and would make continuity rather more important than it would otherwise seem.³

The claim I shall now try to defend, therefore, is a strong one: on Aristotle’s account, continuity by itself ensures that a motion is for the sake of its outcome, and thus simultaneously ensures that the outcome is good. To begin motivating this claim, I want to consider as a possible counterexample the recently described process of heating. Surely, after all, we have no reason to think that the particular degree of heat at which this process ends is good for its subject. In fact, Generation and Corruption tells us that depending on the nature of the body and the degree of heat, the heated body may actually be destroyed. In this case, we

³ We already know that Aristotle considers the good, by its very nature, an end. Although we have not yet mentioned spontaneity and chance, let me now introduce them by suggesting that if the good is by its nature an end, a good outcome can fail to be that for the sake of which only if it comes about κατὰ συμβεβεκός.
would have a continuous change that results in the destruction of its subject, rather than its good. The heat that brings about such destruction is decidedly bad for the substance destroyed, and previous degrees of heat would thus seem to be neutral at best.

There is, however, a problem with this counterexample as stated. The above account of destructive heating fails to note that on Aristotle’s account of change, heating is the actuality not only of the body heated but also of the agent that heats it (Phys. III.3, 202a13–21). Moreover, although the immediate agent may simply be transmitting heat, such a process always involves a primary agent that is hot by nature. Such an agent may act indiscriminately, and so incidentally destroy what it heats, but the heat it confers is nevertheless the outcome of its natural activity, its ἔργον. On the account of goodness that I summarized above, this is enough to qualify the heat as good.⁴

One is tempted to respond that compared with Aristotle’s clearest examples of goodness in the natural world, namely, living things, calling heat good is at best rather trivial. This appearance of triviality is not damning, however, for it corresponds in Aristotle’s metaphysics to a limitation in the sort of goodness something like heat can have. We have already seen that goodness, for Aristotle, 

---

⁴ Indeed, as soon as we think of the outcome of any change as an achievement of the agent rather than as a πάθος of the patient, there seems to be an intuitive sense in which we can characterize it as good.
is an aspect of actuality, and that actuality corresponds to form and to the activities that flow therefrom. Thus, to return for a moment to the hierarchy of form and function in Meteorology IV.12, it follows that the degree of goodness associated with a given form, as well as the intuitive or non-intuitive character of our attributing goodness to that form at all, will depend on its place in the hierarchy. In short, both the difficulty we may find in attributing goodness to heat, and the possibility of making this attribution plausible, are predicted by Aristotle’s account of goodness.

In addition to heat’s low place in the hierarchy of form, the heat in our example has another handicap from the evaluative point of view. Although such heat is the result of a natural capacity, and for this reason can be thought of as an achievement, this achievement is not directed so as to be “good for” anything. As Aristotle points out in Metaphysics Θ.2, capacities not accompanied by reason act necessarily on any receptive substance with which they come into contact. Thus, as we have seen, the heat that corresponds to an agent’s nature may be irrelevant to or even destructive of the patient. Only by chance—in the absence of some more comprehensive principle of order—might it turn out to be good. Perhaps for this reason, Aristotle recognizes a sense of δύναμις which he actually contrasts with nature: whereas nature (φύσις) is a substance’s capacity to act on
itself, δύναμις is the capacity to act on another or on oneself as other (Metaph. Θ.9, 1049b5–10). The existence of natures thus constitutes the most basic way in which a substance’s capacity to act can be directed toward an effect that is good in a stronger sense: it is good for—in the sense of preserving or bringing to perfection—the agent itself.

In fact, however, despite these limitations, the heat in our example is not purely random. The capacity to heat is, by definition, for the sake of heating, and heating is for the sake of heat. If, therefore, an agent possesses the capacity to heat by nature, then heating fulfills an aspect of its nature. This heating is, in its own way, an ordered activity; it proceeds gradually toward a limit set by the capacities of the agent and the patient and by the immediacy of their contact. This minimal goal-directedness is not diminished by the fact that the activity is in other respects random and imperfect—less perfect even than the activity fire exhibits in feeding and sustaining itself on whatever lies at hand, and far less perfect than the unchanging, self-sufficient identity of thinker and thought which all goal-directed activity—especially activity aiming at self-preservation—tries to imitate.
2. Continuity and unity

In considering what is involved in one body’s heating another, we have developed a plausible argument to the effect that every continuous motion is for the sake of its outcome, an outcome that therefore has at least a minimal goodness. We have by no means exhausted, however, the various dimensions of the connection between continuity and goal-directedness. The next step in our investigation begins when we notice that in the chapter following his discussion of continuity in motion, Aristotle argues that to be continuous and to be numerically the same motion (to be unqualifiedly one) are equivalent. In doing so, he adds two requirements for strict continuity, over and above that of not skipping any portion of the πρᾶγμα. The first is that the motion be specifically the same in all its parts—that is, it must involve only a single πρᾶγμα, defined by a single pair of contraries. The second requirement is that the motion be accomplished without pause, in a single continuous time. Now although he relies on the intuitive plausibility of these criteria to justify their place in his account of motion, the thesis that every continuous motion is for the sake of its outcome suggests that further reflection is in order. For if this thesis is correct, then we should expect each aspect of unqualified unity to be necessary in order to guarantee the teleological character of an apparently continuous motion.
In fact, a more detailed consideration of the various components of unity reveals a point-by-point correspondence between the components of unity and the various ways a motion might fail to be for the sake of its outcome. Consider, for example, a change whose analysis reveals two πράγματα, defined by two pairs of contraries. Such a change would actually consist of two successive or overlapping motions, which might only happen to occur together. In this case, even if the first change were to bring about the starting point for the second, it would do so incidentally, and not for the sake of the second and its outcome. Or, consider a change that involves a single pair of contraries but skips part of what lies between them. In such a change, likewise, the immediate outcome of the earlier stages may be related only incidentally to the later stages. Third, even without such a jump, a gap in time between the earlier and later stages may well indicate that the agent cause of the earlier stages was not aiming at the later stages when causing the earlier, but simply at the outcome of the earlier. A final case of discontinuity, which Aristotle does not mention in Physics V.4, would be an apparently continuous change that is in fact caused by two agents, one picking up at the exact moment when the other leaves off. Even without a gap in time, causal discontinuity of this sort means that the change need not be goal-directed as a whole. If, for example, the second agent carries the process farther than the
first agent could have done, the first agent was obviously not acting for the sake of the outcome produced by the second. Once again, the motion as a whole would fail to be for the sake of its result.

In considering the relation between continuity and goal-directedness in motion, I have slipped in an aspect of motion that Aristotle does not mention in *Physics* V.3–4, but that is essential to his analysis as a whole. This is, of course, the fact that motion is the actuality of its agent cause as well as that of its subject. Although this added perspective has been necessary to uphold the thesis that every continuous motion is for the sake of its outcome, I think it can also be justified independently of this thesis. First, if we return to *Physics* V.1 it is clear that the mover is an essential part of the analysis of any motion. Second, the definition of motion in *Physics* III.1–3 makes it clear that every motion, despite being a single reality existing in its subject, nevertheless pertains causally to both the agent and the patient. Therefore, we may safely assume that the individuation—and therefore the continuity—of motions must involve the agent as well as the patient.

In any case, it is clear enough that continuity from the patient’s point of view is not sufficient to guarantee goal-directedness. This point has already emerged, to some degree, from our discussion of heating as a goal-directed action. There I pointed out that in a cumulative change such as heating, the earlier
stages are necessary for the later. Now if such a change is goal-directed, then this necessity turns out to be nothing other than Aristotle’s well known hypothetical necessity. What first made the claim that heating is goal-directed plausible, however, was not simply its status as a cumulative change, but the fact that it is the actualization of a natural capacity of some agent. Moreover, it is now clear that being cumulative only guarantees one aspect of complete continuity: a cumulative change might include more than one πρᾶγμα, and it might not take place in a continuous time. Both of these possibilities, as we have seen, are significant because they reveal various ways, not all of them teleological, in which a given outcome can occur.

We will grasp more clearly the importance of the agent in goal-directed change if we reflect further on hypothetical necessity, and on the necessity involved in a cumulative change. For Aristotle, as the discussions to which I have already referred make clear (see Phys. II.9, PA II.1), hypothetical necessity is part of an important mode of explanation in natural science. Because each form can only exist in matter appropriate to it, the supposition (ὑπόθεσις) that a given form is to be carries necessary consequences for the structure, materials, and coming to be of the matter whose form it is. In the simple case that we have been considering, a given degree of heat can be produced only in a body that has
already been heated to nearly that degree. What makes this type of necessity ex-
planatory, however, is not simply the dependence of form on matter, but the
supposition that the form is to be. Without this supposition, as Aristotle fre-
quently points out, nothing is explained. When a motion fails to be continuous
(or unqualifiedly one) in one of the ways mentioned above, however, the merely
incidental connection between the earlier and later stages of the motion makes
the ultimate outcome irrelevant to explaining the first part of the motion. Any
supposition of the end would be purely arbitrary, and so the necessity fails to be
explanatory.

In the last analysis, therefore, the continuity or unity of motion required if
an action is to be for the sake of the end depends on the way in which the out-
come of the motion, and therefore also its various stages, come about. More spe-
cifically, a motion will turn out to be continuous in the relevant sense if and only
if (a) it is caused by an agent whose natural effect is the outcome of the whole
motion, and (b) this agent in fact causes the various stages of the motion as stages
in the production of the outcome. In the case of heating, the natural activity of
the agent is to bring about the maximum heat that it can produce, and that the
body on which it is acting can receive. Any intermediate degree of heat that such
an agent produces is, as a matter of fact, produced on the way to this maximum,
because no intermediate degree corresponds fully to the natural capacity of the agent relative to the patient. On the other hand, if the agent’s activity happens to be interrupted before it reaches its natural goal, then, in accord with Aristotle’s analysis of contraries, the degree of heat actually achieved will serve as a surrogate end-point for the change. Despite having failed to reach its natural goal, the process of heating will still have been directed toward that goal.

At the beginning of the chapter, I stated that Aristotle’s notion of καθ’ αὑτό agent causality serves as a key link between continuity and goal-directedness. By considering continuity from the perspective of the agent, we can see how this is the case. Fully understood, the continuity condition ensures that a given change is directed by the source of that change, the agent, toward a predetermined outcome. In Aristotle’s terminology, this correlation between a change’s specific outcome and the specific active potentiality of its cause is a καθ’ αὑτό causal connection. As we have just seen, moreover, the existence of such a connection entails that the intervening change is for the sake of an end. If we add to this picture the account of spontaneity and chance in Physics II.4–6, it follows conversely that absent any καθ’ αὑτό connection—that is, when an agent acts κατά συμβεβεκός—the outcome is not an end, but rather a result of spontaneity or chance.
The alternative between καθ’ αὑτό agency for the sake of an end, on the one hand, and κατὰ συμβεβεκός agency leading to a chance outcome, on the other, is one that Aristotle himself proposes in his most important discussion of natural teleology, in Physics II.8. It serves as a premise in his first argument there that nature acts for an end, although the conditional form in which he states the premise has left it uncertain whether he actually accepts it. The continuity condition of Physics II.2, however gives us an independent reason to take the alternative quite seriously: εἰ οὖν ἢ ἀπὸ συμπτώματος δοκεῖ ἢ ἓνεκά του εἶναι (“so if it seems to be either spontaneous or for the sake of something”: 199a3–4). Later, in section four, I shall return to this argument and also briefly consider the analogy between art and nature that follows it in Physics II.8. It is already significant, however, that our discussion thus far supports a straightforward reading of Aristotle’s preferred argument that nature acts for an end.

Thus for the simple motions considered in Physics V–VI, Aristotle’s continuity condition and the accompanying requirement of καθ’ αὑτό agency guarantee goal-directedness. At the same time, simple motions like heating are far from Aristotle’s favorite cases of goal-directed change: the coming to be of living substances. Moreover, because such complex changes involve really distinct stages, not to mention the coming to be of a new substance, they do not seem to fit the
definition of continuity provided in *Physics* V. To understand Aristotle’s favorite example of goal-directedness, therefore, we need to see how he conceives of the unity and continuity of complex changes. However, before going on it may be helpful to review what we have accomplished already.

Drawing on the account of continuity in *Physics* V.3–4, I have argued that the continuity condition of *Physics* II.2 should be taken quite seriously—indeed, that continuity itself is a sufficient condition for goal-directedness in motion. To summarize the argument briefly, the outcome of every continuous motion is the actuality and defining limit of a natural capacity of the agent responsible for that motion, or at least an approximation of that limit. In such a motion, each successive stage is hypothetically necessary for, and is in fact brought about on the way to, the limit that defines the motion as a whole. In short, the continuity condition ensures that an outcome of change is actually playing, in the entire process that leads up to it, the causal role designated by the phrase τὸ ὀὗ ἑνὲκα.

In addition to explicating the continuity requirement, the account of τὸ ὀὗ ἑνὲκα that I have just given corresponds perfectly to the basic schema that emerged from chapter one. What is required for all coming to be alike, we saw in *Generation and Corruption* II.9, is the form that comes to be, the matter in which it comes to be, and the agent that brings it to be. As Aristotle observes in his
critique of the early naturalists, moreover, agent and form are correlative causal principles: agency, on his account, consists in transmitting to matter a form to which that matter is in potency, but which it cannot give itself (see also *Metaph.* Ζ.7–8). In the terms of the foregoing discussion, the form that comes to be is the defining limit of the agent’s capacity to bring it about, and it is this that qualifies the form as an end and that for the sake of which.

3. Continuous coming to be

The question that we must now confront, however, is this: given that the generation and development of an animal or plant consists of many actually distinct stages, and involves the coming to be of a new substance, in what sense can it be considered a continuous change? If the above discussion of simple motions is any guide, we are looking for a sense of continuity that entails the unity of the process in question under the direction of a καθ’ αὑτό agent. Moreover, despite its general silence on the subject of complex changes, *Physics* V.4 does suggest how we should think about this unity, pointing out a second sense in which we may call a motion one:

Further, we also call one a motion that is complete, whether in genus, whether in species, or whether in οὐσία, just as also in other things the complete and the whole pertain to what is one. But sometimes an incomplete motion may be called one, if only it be continuous (*Phys.* V.4, 228b11–15).
Aristotle’s reference to “the complete and the whole” in these lines is of obvious interest. Before discussing the unity of complex changes directly, however, let us prepare the ground by considering how he uses the term “continuous” to characterize changes that are not continuous motions in the strict sense of Physics V.3–4.

To set the context for our remarks on this topic, it will be helpful to recall what we have already seen about the curious ontological status of the process by which a natural substance comes to be. On the one hand, the generation and development of an animal is clearly not the sort of change Aristotle has in mind when in Physics V–VI he considers the properties of motion. Nor, on the other hand, does it fit his remarks about change between contradictories in V.1–3: rather than an immediate passage from non-being to being, it is an extended process that includes both the coming to be and the development to maturity of a new substance. As we saw earlier, however, following Broadie, the definition of motion developed in Physics III.1–3 does seem to fit: in these chapters, Aristotle is concerned with processes of all sorts, from jumping to house-building (cf. III.1, 201a18–9). Having already considered the simple processes that fit his definition, our task now is to deal with those that are more complex.

Although complex κινήσεις do not have the sort of continuity or unity that Aristotle explicitly discusses in Physics V.3–4, I shall argue that the concept
of continuity articulated in these chapters can nevertheless be extended to some complex processes—and in particular, to the coming to be of living substances—in such a way that continuity remains a sufficient condition for the goal-directedness of any change. More specifically, building on Aristotle’s own use of the term συνεχής to describe complex processes, I show in the present section that such processes can verify the two conditions for continuity in general prescribed in Physics V.3, even if they do not meet the following chapter’s requirements for strict continuity in motion. In section four, this broad notion of continuity will enable us to develop the idea of a change that is not simply complex, but constructive. Despite its lack of simplicity, a constructive change can nevertheless have a unity of completeness or wholeness, the sort mentioned in the lines from Physics V.4 quoted above. This unity, like that of a simple motion, corresponds to the operation of a καθ’ αὑτό agent, and thus ensures that the constructive change is goal-directed.

A good starting point for extending the notion of continuous change is a passage in Generation and Corruption I.2 in which Aristotle contrasts coming to be and passing away with alteration:

There is unqualified coming to be and passing away ... whenever it changes from this to that as a whole.... For in the subject there is something as account (κατὰ τὸν λόγον) and something as matter (κατὰ τὴν ὕλην). Whenever, therefore, the change is in these, it
will be coming to be or passing away; whenever in the affections and what is incidental (κατὰ συμβεβεβεκός), alteration (317a20–26).\(^5\)

In the following chapter, Aristotle goes on to point out that “we do not say that what learns comes to be unqualifiedly, but that it comes to be knowing, whereas what grows comes to be” (319a9–11). Together, these two statements help us to formulate a succinct Aristotelian description of a living substance’s development:

In the process of development, a living thing comes to possess in actuality attributes that are essential to the kind of thing it is (τὸ κατὰ τὸν λόγον), at the same time as its bodily structure (τὸ κατὰ τὴν ὑλὴν) develops to support these attributes.

On this account, development is not simply a series of secondary modifications of a living substance, but a genuine case of coming to be. At the same time, as just mentioned with regard to Physics III.1–3, it is strongly analogous to motion in the narrower sense. As in the case of heating, for example, in a living thing’s development a form comes to be in appropriate matter over a period of time, in such a way that the earlier stages of this process are hypothetically necessary for the later stages and for the final outcome. The difference in the case of unqualified coming to be is that the form acquired is not a secondary or

\(^5\) See also GC I.5: “In unqualified coming to be or passing away a thing does not remain (οὐκ ὑπομένει), whereas in alteration, increase and decrease the thing increased or altered remains the same, but in one the affection, in the other the size does not remain the same” (321a22–5).
incidental attribute of an already existing substance; it is, rather, the very οὐσία of the composite that it constitutes together with its subject. This is also, no doubt, why Aristotle insists in Generation and Corruption I.2 that in unqualified coming to be, the change is not only in the form but also in the matter, which especially in the case of living things comes to be as it is only for the sake of the form.⁶

Not surprisingly, given Aristotle’s recurring cosmological concerns in Generation and Corruption, it is in the same work that we find his most interesting uses of the term συνεχής and its cognates that depart from the strict sense of Physics V.3–4. Aristotle sets the context for this looser sense of continuity early in book one, when he concludes that coming to be is ceaseless (ἄπαυστον) because the passing away of one thing is always the coming to be of another (GC I.3, 318a24–6). Much later, toward the end of the work, this basic observation sets up an important conclusion regarding the cyclical transformation of the elements: “Hence motion in a straight line, imitating that in a circle, is also continuous (συνεχής)” (II.10, 337a8). Shortly after, he expands on the relevant notion of continuity: “in things moved continuously—coming to be or altering or, in general, changing—we see a succession, and this coming to be after this so as not to leave

---

⁶ See also DA II.1, 412b25–27, where Aristotle claims that in the case of living things the proximate matter neither precedes nor outlasts the existence of the composite.
a gap (ὥστε μὴ διαλείπειν)” (II.11, 337a34–b2). The last phrase is generally translated into some such phrase as, “in such a way that there is no cessation.” However, Aristotle is here analyzing the causal structure of continuous change, and he is quite clear that continuity in time is parasitic on continuity in the change itself (II.10, 337a23–6). The relevant point, therefore—as the rest of II.11 amply confirms—is not simply that no time elapses between one change and another, but that each change is made possible by the one that precedes it: the changes are causally, and not merely temporally successive, in such a way that each follows immediately upon the last.

In this chapter, the last of *Generation and Corruption*, Aristotle is concerned with the causal structure of the continuous change he finds in the cosmos; in particular, he wants to establish that continual coming to be is necessary. This means moving from the claim that each step in a continuous series of changes presupposes the previous, to the claim that the later step follows necessarily upon the earlier. Because an agent can always be impeded from acting, moreover, Aristotle does not think that the causal ability of the successive agents is sufficient to establish the necessity he wants (337b3–10; cf. *Metaph.* E.2–3). He therefore finds the grounds of necessity in the fact that the actually existing sequence of agents,

---

7 See for example Forster 1955, 320; Joachim 1984, 2:552.
both in the coming to be of the elements and in that of living things, follows a circular pattern. Starting with any point in the cycle, we can identify a goal-directed agency sufficient for supposing the recurrence—not with strict necessity, but at least ἐπὶ τὸ πολύ—of the same state at a later time. Given this supposition, in turn, everything in between must first come to be. Aristotle concludes that despite the lack of strict necessity from the point of view of agency, the causal structure of the continuous coming to be that we actually observe is sufficiently stable that the structure itself, unlike the individual events that constitute it, is necessary.

The causal structure in which Aristotle locates the necessity of eternal coming to be is, with one significant difference, the same structure that characterized as teleological the simple instance of heating that we considered above. The difference is, moreover, precisely the fact that heating, as we considered it above, does not involve the kind of reflexivity that according to Aristotle characterizes natural activity in the full sense, and that also, despite important differences, grounds the eternal necessity of coming to be in the sublunary sphere. Heating is a directed, constructive change from the agent’s point of view; but taken by itself, it falls far short of the ordered self-maintenance that characterizes the natural world as a whole. However, it is the similarities that now concern us. In discuss-
ing strictly continuous motion, we ended up identifying a basic causal structure that accounts for the teleological character of such motions. In the supposition of an end, grounded in the existence of a competent agent and itself grounding the necessity of what the end presupposes, we have now found the same causal structure in continuous coming to be. This brings us an important step closer to specifying what it means for coming to be to be continuous.

Before raising this question directly, however, it is worth noting that there are several passages in which continuity of the kind we are now considering is explicitly linked with being for the sake of an end. They are parallels—in the same or slightly different terminology—of the text from *Physics* II.2, but unlike this text they emphasize the fact that an end may be reached through an ordered series of steps. Consider the following three:

Further, in as many things as have some end, the first and the successive are done (πράττεται δ’ τὸ πρότερον καὶ τὸ ἐφεξῆς) for the sake of this (*Phys.* II.8, 199a8–9).

For by nature are whatever, being moved continuously (συνεχῶς) from some principle within themselves, arrive at some end: and from each not the same for all, nor just anything (οὐδὲ τὸ τυχόν), but it would always tend to the same, if nothing impeded” (*Phys.* II.8 199b16–7).  

Matter of just such a sort must be present if a house or another such end is to be; and this first, then that, must come to be and be

---

8 This passage, although not obviously relevant from the text itself, has to do in context with the development of living things from their respective seeds.
moved, and indeed in just this way, successively, up to (ἐφεξῆς μέχρι) the end and that for the sake of which each thing comes to be and is. So also in what come to be by nature (PA I.1, 639b26–640a1).

With these passages in mind, we can now ask what, in the context of coming to be, Aristotle means by the term συνεχῆς.

First, it is interesting to note that in the passage just quoted from Parts of Animals and in the first of those from Physics II.8, Aristotle uses the technically weaker term ἐφεξῆς just as he elsewhere uses συνεχῶς. As noted in section one, for two things to be ἐφεξῆς, or successive, is the first of two conditions for continuity, the second being that they have a common limit. In the text from Parts of Animals, moreover, ἐφεξῆς occurs in a description reminiscent of the negative phrase μὴ διαλείπειν, which we saw applied to coming to be in Generation and Corruption II.11. In each text, the underlying cluster of concepts is the same: In describing a complex process, each step in which is hypothetically necessary for a given end, Aristotle uses terms such as μὴ διαλείπειν, ἐφεξῆς, and συνεχῶς to indicate the relation between any step and that which must take place immediately before it. Let us reflect briefly on these terms.

First, recall that in Physics V.3, the terms ἐφεξῆς and μὴ διαλείπειν express Aristotle’s requirement that a continuous change cover the entire πρᾶγμα defined by its limits. In considering this requirement, we discovered that for certain cumulative changes, at least, this continuity is part of the stronger, causal
continuity necessary for the goal-directedness of simple changes. What, then, do these terms express when it comes to the more complex changes that we are now considering? Beginning with ἐφεξῆς, suppose that for a given step B to occur presupposes the occurrence of step A. Now states A and B will be successive, in the causal sense that Aristotle needs, if and only if no intervening step C is involved in B’s following from A. In other words, A must be a sufficient, and not merely a necessary condition for B. If, conversely, A were not sufficient for B, then B could only follow from A by way of C, in which case A and B would not be successive. The phrase μὴ διαλείπειν expresses a closely related point. Literally, it excludes any gap in process, whether temporal or simply causal. For instance, suppose that although B follows immediately upon A in time, it is in fact caused not by A but by C, which is causally unrelated to A. In this case there is a causal gap in the temporally continuous process leading from A to B, and so we do not have a significant case of continuous motion. In the change as a whole, the stages up to and including A would be related only incidentally to the stages following A and thus to the end.

What, then, of συνεχῶς? We are now in a position to say exactly to what extent Aristotle’s account of continuity in motion applies to continuous coming to be; moreover, the answer turns out to be surprisingly straightforward. The
two share exactly what is signified by the adjective “continuous” when taken by itself, apart from either substantive. Continuity in general, as Aristotle defines it in *Physics* V.3, is present when successive things of whatever sort—place, time, motion, and so forth—share a limit. Not only in simple motions, however, but in complex continuous changes as well, causally successive steps do in fact share a limit. To see this, let us return to our first example, in which step A in a complex process is a sufficient condition for B. This can only be the case if the possibility of B is immediately contained in the actuality of A, so that as soon as A occurs, B will follow. In fact, Aristotle explicitly discusses this sort of immediate possibility in *Metaphysics* Θ.5 and 7, one conclusion of which is that strictly speaking, we call something possible only when nothing further is necessary for it to come about, so that if nothing interferes it will, in fact, occur. In short, although it would be incorrect to say that the potentiality for step B is identical to the actuality of A, it is nevertheless contained within it, so that the limit-to-which of A’s coming to be is the limit-from-which of B’s coming to be. This is what it means to say that two steps in a complex change are continuous.

4. Unity of coming to be

It is now clear enough that natural coming to be is continuous in a straightforward sense. By contrast with continuous simple motions, however, it
is not clear what kind of unity we can attribute to an instance of coming to be. Because the equivalence of continuity and unity played an important role in establishing the goal-directedness of continuous motion, an investigation of unity in coming to be seems called for. The need for such an account will be clearer, however, if we first point out what is not included in the kind of continuity I have just described. First, although it was drawn from Aristotle’s discussions of natural coming to be, the above account of continuity in change does not guarantee that the sequence of changes so characterized is directed toward any particular outcome. In the case of simple motions, this difficulty is ruled out by the observation that a single continuous motion involves only one \( \pi\rho\alpha\gamma\mu\alpha \). In the case of coming to be, however, even if every change in the sequence can be described separately as the \( \kappa\alpha\theta\' \alpha\upsilon\tau\omega \) effect of some agent, the sequence as a whole may still be a random concatenation. Moreover, the fact that such a sequence of changes may include instances of coming to be and passing away removes even the continuity that would be present in a progressive series of motions undergone by a single subject. In short, what we have so far is only sufficient continuity to keep the world as a whole moving: there is no guarantee yet that some complex portions of it are going anywhere in particular.
To find this guarantee, however, we need only return to the basic structure of goal-directed change, which emerged first in chapter one and then in sections one and two of the present chapter. As in the case of continuous simple motions, what can ensure the continuity of a complex change as a single, goal-directed process is its relation to a single agent, aiming at a single outcome. This agent need not be the immediate cause of each stage in the process, but it must be responsible for the process as a whole. That this is frequently the case, however, Aristotle takes to be obvious from the simple fact that complex changes leading to familiar outcomes occur repeatedly, in the same or similar circumstances, throughout the world as we know it. As already observed, that which happens always or for the most part does not happen by spontaneity or chance, that is, κατὰ συμβεβεκός. Not to happen κατὰ συμβεβεκός, of course, is to have a καθ’ αὑτό agent cause, and the operation of such a cause provides grounds for supposing that the final outcome will take place. This means, finally, that the process is hypothetically necessary for, and takes place for the sake of, its outcome (see Physics II.8, 198b34–199a8). The continuity we are looking for, therefore, is that which characterizes a process of this sort.

We are now on familiar territory, and there is no need to explain in detail or to provide extensive references to the various arguments by which Aristotle
establishes the teleological structure of the complex but ordered changes that characterize his cosmos. Most or all of these arguments are found in Physics II.8; most or all are found elsewhere as well, especially in Parts of Animals I.1 and I.5. Moreover, the first argument from Physics II.8 is the argument we have just mentioned concerning what takes place always or for the most part. A brief discussion of the second, however, the well-known analogy between nature and art, will help us develop the concepts of continuity and unity in change that we need to complete our commentary on Physics II.2. It will also help make clear what sort of agent we need if a complex change is to have a καθ’ αὑτό cause.

In the art analogy (199a8–32), Aristotle points out that many of the changes we observe in nature have the characteristic structure of things done for the sake of an end—a structure that we have already discerned, he thinks, by observing and experiencing human activity. When we set out to make something, for example, the various steps in which we engage are cumulative, and cumulative in just such a way that they result in this particular end. Their cumulative character is analogous to that of a unified simple motion, although more complex changes—we might call them “constructive” rather than “cumulative”—typically involve multiple steps that differ from each other in kind. Aristotle’s point in the art analogy is that if we know what to look for, we find in nature instances of
constructive change as diverse as ants working cooperatively to store food for
the winter, and plants growing to maturity and producing fruit. To make the
analogy as forceful as possible, he comments that if a house came to be by nature
rather than by art, it would nevertheless have to be constructed in more or less
the same way as it is now. Taking the art analogy together with the cumulative
character of goal-directed simple motions, we can now identify the cumulative or
constructive nature of goal-directed change as one of its general features.

However, the analogy between nature and art is not meant to stand alone
as an argument that natural processes are goal-directed, nor can it do so. In prin-
ciple, a constructive change of the sort Aristotle has in mind might take place
spontaneously, in which case it would not take place for the sake of an end. Nev-
ertheless, the analogy serves to focus our attention on the fact that constructive
change is typical of, though it does not necessarily imply, goal-directed action.
Such change, moreover, is the norm rather than the exception in nature. Thus the
central premise in Aristotle’s first argument—the dichotomy between goal-
directed change and spontaneity or chance—also provides the context within
which the art analogy gains its force.

Aside from the art analogy and its place in Aristotle’s arguments, the no-
tion of a constructive change and its difference from a merely cumulative change
such as heating is worth dwelling on. We have already seen the grounds of this
difference in the fact that a motion like heating has a stable substrate, constituted
by material and formal principles that remain the same throughout. The change
is cumulative in that each degree of heat attained disposes the subject to receive a
greater degree of heat. As long as the change remains a motion in the strict sense,
the basic existence of the substance involved is not affected. Coming to be, as we
have seen, is quite different. Here the change involves the matter and the form
themselves, so that a new substance comes to be. If we think back to the hierar-
chy of matter and form in Meteorology IV.12, moreover, we can easily see why the
notion of constructive change seems so appropriate for coming to be. As we saw
in chapter one, coming to be involves producing in appropriate matter a higher,
and therefore a more determinate, and so in turn a more clearly discernible form.
In this way, each level of actuality builds upon the previous.

The notion of construction turns out to be even more appropriate if we
consider that in Aristotle’s cosmos, the actuality or form of every natural sub-
stance is something complex. As we saw in chapter one, even the simplest bodily
substances, the elements, each have two defining qualities, and these qualities
are determinations of a generic determinable, corporeality, which although it
cannot exist separately nevertheless underlies and provides a foundation for the
formal qualities that define the four basic sorts of body. In the homogeneous bodies composed of the elements, we find a much greater variety of active and passive bodily qualities. These in turn make possible a wide array of heterogeneous bodies, which display an even greater complexity insofar as they can be formed from more than one homogeneous body and are characterized not only by their directly perceptible qualities, but also by their shape and by the interrelations of their parts. All this, finally, makes possible the complex forms of living things, a complexity that Aristotle points out at some length in *Parts of Animals* I.2–4, when he argues that living substances cannot be defined by dichotomy.

Although the point cannot be discussed thoroughly here, Aristotle also makes it clear in *De anima* and especially in *Generation of Animals* that within the complex form of a living thing, some capacities are for the sake of others. The best example of this is that the perceptual faculty depends for its existence on the nutritive faculty, and the nutritive faculty in a given animal must therefore be of the right sort to support its perceptual life. Any serious study of living things reveals an enormous complexity of similar relations, a complexity barely hinted at in *De anima* but investigated at great length in the biological works. The production of a living form in matter, therefore, is necessarily a complex process in which the most essential aspects of the nutritive faculty come to be first, in the
parts that serve as their matter. These are followed by other parts and other faculties, and so on until the animal reaches maturity and is capable of producing another like itself. Here, obviously, the notion of a constructive process is fully exemplified.

Before returning to the question of unity in complex processes of coming to be, there is a final point to be made about the constructive character of such processes. When we considered simple motions like heating, I pointed out that such motions, though teleological at a basic level, are relatively weak examples of goal-directed activity. This was mainly because of the disconnect that can exist in such cases between the actuality or good appropriate to the agent cause of the motion and that appropriate to its subject. In the worst case, the agent’s self-fulfillment can result in the patient’s destruction. Obviously, however, this cannot be the case in coming to be. Here, whether the agency is internal or external, the agent of coming to be necessarily—by definition, in fact—expresses its nature constructively. This colloquial use of the English “constructive” brings out the fact that the outcome of coming to be is, necessarily, the actuality and the good both of the agent and of that upon which it acts.

We are now in a position to complete our discussion of unity in coming to be. Earlier, in order to pursue the notion of continuity, I postponed discussion of
a text from *Physics* V.4 according to which a motion can be one in virtue of being complete. It is now time to see how continuity and the notion of a constructive change can help us approach the notion of completeness, and thus arrive at the unity of coming to be. Let us first review the text:

Further, we also call one a motion that is complete, whether in genus, whether in species, or whether in ὀὐσία, just as also in other things the complete and the whole pertain to what is one. But sometimes an incomplete motion may be called one, if only it be continuous (*Phys.* V.4, 228b11–15).

In writing these lines, Aristotle may or may not have been thinking particularly of motion in the strict sense. In any case, the point clearly applies to coming to be as well. A change is complete, on Aristotle’s account of completeness, when it reaches the limit to which it is directed (cf. *Metaph.* Δ.16, 1021b24–5). In the case of coming to be, this is obviously the form of the resulting composite, and what takes place between the starting point of the complex change and this final limit is obviously a single process of coming to be. As we have already observed, moreover, and will consider further in just a moment, these two limits are the boundaries of a constructive process governed by a single agent. The steps that fall between the limits are therefore continuous with each other in a way in which they are not continuous with any changes that fall outside them. Thus for

---

⁹ There are actually two agents, but the goal of the first is to produce the active power of the second; see below.
coming to be as well as for motion, unity and continuity in this final sense coin-
cide. This sense of continuity is, finally, the sense in which continuity guarantees
that a process of change, whether or not it is a continuous motion in the first
sense of *Physics* V.3–4, takes place for the sake of its outcome.

As in the case of simple motions, therefore, the key step in establishing
that an instance of coming to be is continuous is to show that such coming to be
is brought about by an agent whose activity is aimed at producing the mature
form. Of course, this is not always the case: the subterranean production of vari-
ous homogeneous bodies is, no doubt, largely a matter of chance. Aristotle’s fa-
vorite case of natural coming to be, however, is the reproduction of living things,
and this can easily be distinguished from random or partly random processes by
applying the criteria developed in Aristotle’s discussion of chance (Phys. II.4–6).
Several times in the course of this discussion, Aristotle writes that chance events
are those that, although they might have come about for the sake of an end, in
fact have only an accidental cause. Such events unfold, perhaps, just as they
would by nature or art, but there is no καθ’ αὑτό cause of this development. As
with the continuous and pseudo-continuous motions discussed in section two, it
is the presence or absence of a καθ’ αὑτό agent that determines whether coming
to be takes place for an end.
Aristotle holds that apart from the few cases of spontaneous generation, the coming to be of living things is always the result of a *καθ’ αὑτό* cause—or rather of two such causes, the role of the first being to produce the second. To summarize the account found in *Generation of Animals* II.1–4, the agent cause of an animal’s coming to be as a distinct substance is the male parent. This parent, either through its own activity directly or through the motions it imparts to an instrument, the semen, produces from the material supplied by the female a conceptus (κύημα). This conceptus is a relatively undifferentiated mass separated by a membrane from the more fluid parts of the καταμήνια, and thus structurally distinct from the mother. As soon as it exists, the active role of the male in the reproductive process is complete: the conceptus already has a primitive nutritive soul, and if it receives adequate nutrition, it can develop to maturity by its own agency.

Once the conceptus has come to be, the first internal organ to develop is, Aristotle argues, that which is to direct the rest of the process. In sanguineous animals, this is the heart, which from the first moment of its existence is the source from which both nutriment and the heat required to concoct it come to the various parts of the developing animal. He is clear that the temporal priority of the heart in coming to be must be understood teleologically: for the process of
reproduction to succeed, the parts and faculties most essential to the offspring’s existence and development must come to be as quickly as possible. The male parent, therefore, produces a distinct substance which has a primitive structural integrity through its surrounding membrane, and within which the master organ, the source of nutrition and later of perception, is on the point of coming to be.

We must, therefore, distinguish two stages in the coming to be of a living thing. In the first, the parent or parents produce a new substance distinct from themselves. The immediate result of their action is the instantaneous, unqualified coming to be of the offspring, its passage from non-being to being. However, this coming to be is incomplete in that the form at which it aims has not yet been achieved. The complete process of coming to be continues under internal direction by the nutritive faculty, which exists primarily in the heart or in an analogous organ. Despite this division into two stages, however, coming to be is a single, continuous, and complete process because the first of the two agents involved has as its καθ’ αὐτό effect the nutritive soul, that is, the faculty by which the coming to be is completed.
5. Conclusion

The conclusion of this chapter is, briefly put, that Aristotle considers continuity a sufficient condition of goal-directedness for any sort of change. The robust sort of continuity that can play this role, however, is that of a single process, directed by a καθ’ αὑτό agent toward an end determined by the agent’s nature. The matter in which this end is produced is hypothetically necessary for the end, as is each of the stages or steps by which the agent brings about the end. In short, our discussion of continuity has ended up reconstructing the causal nexus within which, in Generation and Corruption II.9, Aristotle locates τὸ ὑ.erb. This nexus consists of the form produced, the matter in which it is produced, and the agent that produces it. In teleological terms, that for the sake of which is the form, that which acts for the sake of the form is the agent, and that which is hypothetically necessary for the sake of the form is the matter.

As we saw in chapter two, however, Aristotle is not content with a teleology that coheres with the rest of his causal analysis of nature. Rather, his teleology is grounded firmly in his first philosophy. At this level of analysis, its precise foundation is the claim in Metaphysics Θ.8 that actuality is prior in λόγος and in οὐσία to potentiality. With this metaphysical thesis in hand, Aristotle can claim not only that his teleology is empirically adequate, but that it is necessarily true.
Matter as such involves a potentiality for form, and agency is the potentiality to produce form in matter. Form, by contrast, is actuality. Both matter and agency are thus for the sake of form, whereas form, as such, is not for the sake of anything. We have also seen, however, that insofar as it involves a set of capacities for activity, form does involve potentiality, and this potentiality is for the sake of the corresponding activity.

To move past the exegetical, or perhaps better, the reconstructive work of these first chapters, to the level of philosophical analysis, is a task for another occasion. Nevertheless, it is worth at least stating the central point on which such reflection must focus. For Aristotle, unlike for many of his predecessors and contemporaries, and unlike the approach most congenial to heirs of the scientific revolution, the basic explanatory regress in natural science is not toward matter and agency, but toward form and activity. At the level of ultimate explanations, he argues—the level of metaphysics—it is actuality, not active or passive potentiality, that is basic.
CHAPTER FOUR

GOAL-DIRECTEDNESS AND MECHANICAL SYSTEMS

The foregoing chapters, although by no means a complete discussion of Aristotle’s teleology, nonetheless allow us to ask a further question. If this teleology is more or less what I have argued, is it now of purely historical, or also of philosophical interest? In this and the following chapter I argue that the core tenets of Aristotelian teleology, which I shall review in a moment, are worthy of serious consideration. I am well aware, however that in view of the currently dominant conception of nature, so thoroughly mechanistic and Darwinian in inspiration, this suggestion will meet with a good deal of skepticism. With respect to Darwinism, recent discussions of functional explanation have led to a fairly broad consensus among philosophers of biology that functional notions and claims in biology should be explicated in Darwinian, adaptationist terms.\(^1\) Because Aristotle did not have a theory of evolution—and because function is the

---

\(^1\) On this consensus see Neander 1991, 168; Amundson and Lauder 1994. The classic statements of the “etiological” or “selected effect” account of function are Wright 1998 and Millikan 1984. Some important criticisms are found in Cummins 1975/1984; Boorse 1976; Bigelow and Pargetter 1987; Bedau 1992; Amundson and Lauder 1994.
only teleological concept much discussed in recent decades—it is often assumed that his teleology must be significantly different from any that might currently be accepted.\textsuperscript{2} With regard to mechanics, moreover, final causes have been excluded from the physical and chemical sciences since long before Darwin, by a mechanistic view of nature that emerged from the revival of ancient atomism and the development of modern, mathematical mechanics.

In this chapter and the following, I shall focus on the relation between Aristotle’s teleology and modern mechanics, and in particular on the teleological concept of goal-directedness. One reason for these choices is simply that one has to start somewhere. There are, in addition, several advantages to mechanics and goal-directedness. With respect to mechanics, an important factor is that I consider the “etiological” or “selected effects” account of function, which links the concept of function in biology to natural selection, simply mistaken. Because this is not the place to continue the recent debates about functions, and because simply assuming a non-etiological account of function would beg the question of Aristotle’s relevance, I prefer to avoid the question of evolution. Moreover, I have been arguing that Aristotle’s teleology is not even tied to his biology; rather, he provides a unified, teleological account of all natural substances. If this is so, then

\textsuperscript{2} But see Depew 1997.
any attempt to connect the teleological claims found in his biology with contemporay biological issues would presuppose that these claims can be detached from the more general teleology in which they are embedded. This may or may not be the case, but it certainly requires further discussion.

My reason for focusing on goal-directedness also stems from my disagreement with the etiological account of function, and the inadvisability of either ignoring this account or attacking it here. The best available account of function, in my view, is the “general causal contribution” account defended, for example, by Boorse and Adams. According to this view, however, the concept of function is subordinate to that of goal-directedness. Because no one has argued that the concept of goal-directedness is dependent on that of function, goal-directedness seems a good place to start. Another problem with the recent discussion of function is that since about 1970, many writers have considered this topic while making little or no reference to the problem of teleology taken more generally. Whatever the merits of this strategy in itself, for present purposes it is clearly a disadvantage. Goal-directedness, however, is always considered a teleological phenomenon, and is generally discussed in relation to the overall question of the place of teleology, if any, in contemporary science. A final reason for
preferring goal-directedness to function is that in discussing Aristotle, we have spent considerably more time on the former.

The account of goal-directedness we shall consider is known variously as the “cybernetic,” the “system-property” or simply the “systems” view. Originally developed in the nineteen-forties and fifties, its best known exposition is chapter twelve of Nagel’s *The Structure of Science*; more recently, it has been defended by Christopher Boorse. Because the systems view, as I shall call it, is in large part an attempt to distinguish mechanical systems and processes that are goal-directed from those that are not, it is particularly suited to a discussion of Aristotle’s teleology and modern mechanics. Moreover, although my purpose is not to defend it against objections, I do think that it captures, at least for a range of important cases, the reasons we are inclined to think of certain systems as goal-directed. Another point in its favor is that on the other prominent view of goal-directedness, the “intentional” analysis, attribution of goals does not have an objective basis in the apparently goal-directed system or behavior, and so has at best a

---

3 See section 3 for references and a slightly longer history.

4 Boorse (2002) convincingly answers a number of objections to his “general goal-contribution” account of function; some of these objections and responses are directly relevant to his account of goal-directedness as well. The objections I shall formulate below are quite different.
certain heuristic value.\textsuperscript{5} Clearly, there is no point in comparing Aristotle’s teleology, as I interpret it, with non-realistic accounts of teleological language.\textsuperscript{6}

In examining the systems account of goal-directedness, I shall argue that its strengths in fact depend on the covert use of Aristotelian concepts. Once these concepts are uncovered, moreover, the class of mechanical systems that must be recognized as goal-directed expands considerably. In this chapter and the next, moreover, I shall argue that an adequate account of “directively organized systems,” to use the phrase coined by Nagel, cannot be separated from a teleological understanding of mechanical phenomena in general. In exploring this last, more controversial thesis, I hope to show how a relatively modest teleological approach to classical mechanics may lead to new insights in the theory of scientific explanation.

The main task of this chapter, therefore, will begin in section two with an introduction of the system-property view as articulated by Nagel and Boorse. In section three, while acknowledging the strengths of this view, I develop several Aristotelian challenges to its adequacy as a general account of goal-directedness in nature. The issues raised in this section focus on the relation between

\footnotesize{\textsuperscript{5} The classic statement of the intentional approach is Woodfield 1973. For recent defenses see Schaffner 1993, chap. 8; Nissen 1997.}

\footnotesize{\textsuperscript{6} But see below and n. 7, concerning interpretations that make Aristotle’s teleology less “causal.”}
teleological and mechanical considerations in the analysis of mechanical systems, and prepare the way for further discussion of teleology and mechanics in chapter five. Before any of this, however, it will be helpful first to review and to comment briefly on the main tenets of Aristotle’s teleology, as this has emerged in chapters one through three.

1. Aristotle’s teleology reviewed

Our discussion of Aristotle’s teleology in the foregoing chapters was heavily exegetical, and in this sort of argumentation it is sometimes difficult to see the forest for the trees. It is thus important to step back for a moment, and take a look at the thicket through which we have passed.

In chapter three, I argued at length that for Aristotle, the outcome of every continuous change is an end, or that for the sake of which the change takes place. The significance of this striking claim came into focus when we saw that a continuous change, on the strong notion of continuity that the claim requires, is one that has a καθ’ αὑτό agent cause. In Aristotle’s view, that is, finality and agency are correlative: the καθ’ αὑτό effect produced by any agent is that for the sake of which it acts. In chapter two, moreover, it became clear that Aristotle’s teleology, including his understanding of agency and change, can be reformulated and clarified in metaphysical terms. Thus the outcome of a unified or continuous
change is that for the sake of which it takes place because it is the actuality and limit in terms of which the change must be understood. Likewise, an agent’s proper or καθ’ αὑτό effect is the fulfillment and limit of its own agency; it is in terms of this effect, therefore, that the agent as such must be understood.

As we also observed in chapter two, Aristotle holds that in general, every potentiality and every incomplete actuality must be understood in relation to the corresponding complete actuality: only complete actuality, that is, has the definite, positive character required to make the natural world intelligible. It is precisely this relation between potentiality and actuality, I have argued, that Aristotle has in mind when he uses the explanatory phrase “for the sake of,” or refers to something as an end. Thus incomplete actuality, or change, is for the sake of the form in which it results. Likewise, the capacity to act is sometimes for the sake of a complete activity, and sometimes for the sake of a change and of the outcome of that change. The capacity to be acted on, too, is for the sake of the corresponding change and of its outcome. Matter, finally, is for the sake of form. In each of these relationships, complete actuality emerges as a kind of ontological anchor in the natural world, and “that for the sake of which” emerges as a basic causal and explanatory category.
Before turning to the systems view of goal-directedness, I want to raise a vexed question that we cannot discuss here as it requires: in what sense, exactly, is that for the sake of which a cause? In the foregoing paragraphs I explicated Aristotle’s “for the sake of” relation several times by stating that a potentiality or incomplete actuality must be “understood in terms of” the corresponding complete actuality. This phrase, however, implies of itself only a conceptual dependence of “that which is for the sake of something” on “that for the sake of which it is.” And in fact, some commentators have argued that in current philosophical vocabulary, Aristotle’s οὗ ἑνεκα should be thought of as an “explanatory factor,” of epistemic, but not of ontological significance, rather than as a cause properly speaking. Now it should be clear that this interpretation does not sit well with the metaphysical understanding of Aristotle’s teleology for which I have argued, and especially with Aristotle’s claim that actuality is prior to potentiality not only in λόγος, but also in οὐσία. Yet if we claim an ontological role for Aristotelian final causes, we are immediately faced with teleology’s basic riddle. Potentiality, in many cases at least, precedes actuality in time: the capacity to act or to be acted upon can exist without the corresponding actuality. It goes without saying,

---

as well, that every motion precedes its own outcome. How, then, can what exists depend on what does not yet exist, and may never exist at all?

This is an old and difficult question. It arises, however, only given an affirmative answer to the previous question: Is nature, as our best science reveals it to us, genuinely teleological? It is this question that we are now considering, and I think the other may be put off for another occasion. In chapter five, however, I will suggest that the difficulty is not with teleology alone. It is, rather, a problem that arises for any realistic, causal understanding of natural phenomena. That this seems to be Aristotle’s view as well is interesting enough to warrant a brief discussion, which can be found in the appendix to this chapter.

2. The systems view of goal-directedness

The systems view of goal-directedness has an interesting history. Especially in its early formulations, it was strongly influenced by contemporary work in cybernetics, which had produced apparently goal-directed artifacts such as self-guided missiles. It was also affected by the operationalist or behavioristic understanding of science. The direct influence of cybernetics is found primarily in the original formulation by Rosenblueth, Wiener and Bigelow in 1943; and the desire for a behavioristic or operational account of goal-directedness is most explicit in Sommerhoff’s 1950 discussion. The best-known articulation of the
systems view, however, and the one on which I shall mostly rely, is found in chapter twelve of Nagel’s *The Structure of Science*. In 1976, Nagel’s discussion of function in relation to goal-directedness inspired Christopher Boorse to present a modified version of the systems view, drawing on Sommerhoff, in order to ground an alternative to Wright’s etiological concept of function. Boorse’s version of the systems view is notable in that he explicitly drops Sommerhoff’s demand for a behavioristic or “black box” analysis. An attractive feature of Boorse’s approach is the flexibility with which, through a series of careful distinctions, his concepts of goal-directedness and function can be applied in various cases. In the following pages, I shall draw on his discussions to complement Nagel’s.\(^8\)

For present purposes, however, the systems view can be considered apart from the twists and turns of its history, as an approach to teleology based on four assumptions about the physical world and apparently goal-directed behavior. First, mechanical systems in general are not goal-directed. Second, the activity of certain biological and artificial systems is nonetheless appropriately described as taking place for the sake of, or in order to attain, a specific end. Third, goal-directed behavior can be identified and analyzed simply on the basis of the

---

\(^8\) See Rosenblueth, Wiener, and Bigelow 1943; Sommerhoff 1950; Nagel 1961, 1977; Boorse 1976, 2002. Note that Nagel’s 1961 is a development of his 1953. Other discussions include Beckner 1959, 1969; Adams 1979; and Bigelow and Pargetter 1987. Each of these discussions differs from Nagel’s in one or more respects; except as noted in the text, however, their concepts of goal-directedness do not differ at the level of analysis required for our purposes.
system’s behavior, or its behavior and internal structure. Thus, although some
instances of goal-directedness may in fact be results of natural selection, evolu-
tionary considerations are irrelevant to the question of what constitutes goal-
directed behavior at present. Finally, instances of goal-directedness in purely
mechanical systems have a common structure with goal-directedness that in-
volves awareness of, and intentional pursuit of, a goal. Given these assumptions,
the aim of a systems analysis of goal-directedness is to discern the common
causal structure of systems we identify as goal-directed.

To introduce the analysis itself, I shall first summarize Nagel’s discussion
in *The Structure of Science*, then add a helpful formulation from Boorse, and fi-
nally consider several of the more important objections that their analysis has
faced in the past ten years. Because we shall consider some of Nagel’s examples
in section three, I will keep this initial summary brief and therefore relatively ab-
stract. Rather than following his analysis step by step as he develops it, it will be
helpful to focus on three major points.

First, in order to provide an intuitive starting point for his analysis, Nagel
claims that teleological statements involve “expressions signifying a means-ends
Although he does not explicitly return to this claim, we shall see that it calls for further reflection.

Second, after suggesting that basic teleological phenomenon—and thus the condition of applying any teleological concept—is goal-directedness, he proposes the following account of goal-directed systems. Let $S$ be a system in state $C$, in virtue of which—in an appropriate environment $E$—$S$ either has or tends toward a state or process $G$; and let $C$ therefore be called a $G$-state. Given this much, $S$ is goal-directed, or “directively organized,” with respect to $G$, to the extent that disturbances in $C$ or in $E$ that tend to keep $S$ from $G$ result in compensatory changes in $S$, whereby $S$ is returned to some $G$-state. From this definition, Nagel extracts two essential, quantifiable features of goal-directed systems: persistence and plasticity. A system’s degree of persistence corresponds to the range of internal disturbances for which it can so compensate as to sustain the goal $G$; its degree of plasticity corresponds to the range of circumstances, or variations in the environment, in which it can sustain $G$.

Third, following Sommerhoff, Nagel requires that the compensatory changes by which $S$ returns toward $G$ not result simply from the general laws governing the various parts of the system. They must, rather, result from the

---

9 Nagel 1984, 323.
organization of S as a whole. Because of its original formulation, this requirement is generally known as “the independence requirement.” It is meant to exclude a broad class of systems, such as pendulums, that would otherwise fit the systems view but that are not, at least intuitively, goal-directed.  

Boorse, who considers these requirements the core of the systems view, also provides a helpful informal summary. “To say that an action or process A is directed to the goal G,” he writes, “is to say not only that A is what is required

---

10 The formulation of the independence requirement given in the text is mine, and requires a brief history. As first introduced by Sommerhoff, what it required was the “independence (‘orthogonality’) of system variables. Each must be nomologically able to vary independently of the others, a test which rules out some equilibrium-seeking systems such as a pendulum or a marble rolling around in a bowl” (Boorse 2002, 69). However, Woodfield (1976, 69) later raised an objection to this requirement that Nagel considered “the only serious difficulty” in the systems view (1977, 211): if a system is in fact goal-directed, then by definition the state variables in terms of which it is described are not functionally independent (see discussion in Nissen 1997, 17–18). Nagel responded to this difficulty by clarifying the requirement thus: “apart from those situations in which determinate relations hold between the variables because of their role in goal-directed processes, the known (or assumed) ‘laws of nature’ impose no restrictions on the simultaneous values of the variables” (1977, 212). To this reformulation Nissen has raised several additional objections (1997, 18–19), one of which is worth considering here: Without prior knowledge of which systems are goal-directed, we cannot determine which correlations among variables result from their role in goal-directed processes, and which result from the “laws of nature” alone, and thus we cannot know whether a given correlation violates the independence requirement. It is precisely the independence requirement, however, that is supposed to distinguish goal-directed from non-goal-directed systems. Thus Nagel’s criterion for making this distinction is circular.

It seems to me, however, that Nissen makes too much of this problem. The purpose of the revised independence requirement is, I think, to distinguish effects of the general physical laws governing a system’s parts from the behavior that we attribute to the system as a whole, in virtue of its structure, and thus to distinguish law-like behavior in general from behavior that is also goal-directed. If this is correct, however, then another objection arises—which I will develop below—based on the fact that the laws of nature alone do not account for the behavior of any system at all. If, on the other hand, we interpret the revised independence requirement more strictly, as does Nissen, then his objection appears to be fatal. See the discussion below, in section three.
for G, but also that within some range of environmental variation A would have been modified in whatever way was required for G." He provides the following example: “Capturing a bird may be the goal of a cat’s behavior in so far as this behavior not only is appropriate for capturing a bird but would also have been appropriately modified if the bird had behaved differently.”¹¹ This simple formulation nicely expresses the intuitive appeal of the systems view.

As mentioned above, discussions of teleology since about 1970 have focused on the concept of function rather than on that of goal-directedness. Boorse himself is primarily interested in defending an account of function as goal-contribution, for which a defensible concept of goal-directedness is of course prerequisite. The same is true of Frederick Adams, who has defended an account of function similar to Boorse’s.¹² However, the systems view of goal-directedness remains a standard topic for those who consider teleology in general, rather than focusing on the concept of function.¹³ Thus it is not surprising that over the past fifty-five years, a number of objections have been raised. Although it is not my purpose here to defend the view, I suspect that Boorse’s version can handle a

¹³ See for example the criticisms by Bedau 1992; Schaffner 1993, 365–8; Nissen 1997, chaps. 1–2.
good number of these objections. In any case, there are surely a large number of systems and activities that we intuitively consider goal-directed for reasons that are at least roughly similar to those that the systems view proposes. Common examples include thermostats, self-guiding missiles, automatic regulation of body temperature and of water-content in the blood, and animals’ pursuit of food. If the systems view gets examples such as these even approximately right, then it serves my purposes.

In addition to objections focusing on one or another aspect of the systems view, however, there are also global objections to any realistic—that is, nonintentional—analysis of natural goal-directedness. Because these objections would also apply to an Aristotelian view, I would like to consider two of them here. The first is Schaffner’s claim that the neither the systems view nor any other realistic account of natural teleology provides any reason for using the concept of a goal outside the context of human intentions. The second is Nissen’s argument that

---

14 The criticisms leveled at the systems view are discussed at great length by Nissen (1997, chaps. 1–2). I suspect that Boorse’s account can handle most of these objections, generally along lines of response drawn from Boorse himself (1976, 2002). It is curious that although Nissen cites Boorse’s attacks on Wright approvingly (142, 155), and addresses his arguments against a purely intentionalistic analysis of function (213–16), he does not consider Boorse’s response to earlier objections that he considers in chapters 1–2. For example, he criticizes behavioristic versions of the systems analysis because they cannot handle the pursuit of non-existence goal objects and the compatibility of a given behavior with multiple goals (4–5, 8, 26, 31). However, Boorse explicitly abandons the requirement of an operational definition, in the face of just these objections (1976, 79). I suspect his response will also handle Nissen’s objection based on non-persistent goal-directed behavior (22–3, 39–40). Boorse’s recent response to objections to his account of function also contains some material relevant to goal-directedness (2002, esp. 83–90).
because teleological explanations, but not causal explanations, are intensional, it is impossible to give a causal account of any teleological phenomenon.\textsuperscript{15} These objections turn out to be closely related, so I will summarize them both before commenting.

Schaffner’s objection is based on two observations.\textsuperscript{16} First, the systems view is intended to provide a causal or “naturalistic” account of teleological phenomena in nature. Thus, it ought to explicate teleological concepts without using intentional concepts—in terms, that is, that can be used of ordinary, non-teleological causal structures.\textsuperscript{17} Second, all proponents of the systems view in fact treat “goal” as a primitive concept. If they were to give an explicit account, they would have to say something like, “a goal is an $x$ that is achieved under a variety of conditions”—and there is nothing to substitute for $x$ that does not already include the notion of a goal.\textsuperscript{18} A neutral term, such as “outcome,” will not do: as

\textsuperscript{15} Note the distinction between the \textit{intentional} view of teleology, to which Nissen subscribes, and the \textit{intensionality} of teleological concepts involved in this particular argument.

\textsuperscript{16} Schaffner 1993, 367–8. His discussion is quite succinct, and I have taken the liberty of filling it out a bit.

\textsuperscript{17} Early proponents of the systems view actually differed over the translatability of teleological into nonteleological explanations. Nagel (1961), for example, argued for translatability, whereas his student Beckner (1959) took the opposite view. If we leave aside the debates over explanation and translatability, however, focusing on the relation between teleology and causation, Schaffner’s point appears to be valid.

\textsuperscript{18} Another way of putting the same point is that on the systems view, it is unclear what adjectives such as “persistent” and “plastic” are supposed to be modifying. After all, persistence and plasticity are supposed to be constitutive features of goal-directedness. But what is it that the
Jonathan Bennett remarks, “every animal is tremendously plastic in respect of becoming dead.” But death, generally, is not the animal’s goal. In short, the systems view involves an apparently intentional concept, “goal,” of which it provides no “neutral” analysis.

In the course of his lengthy discussion, Nissen uses different terms to make a similar point. Like Nagel, he associates teleological concepts with means-end relations. He goes on, however, to drive a wedge between causality and teleology by emphasizing the difference between means-end relations and causal relations in general. More particularly, in criticizing the systems analysis of function proposed by Adams, he draws attention to the latter’s admission that functions are intensional. Thus if an optical scanner with character recognition software has the function of detecting the letter “Q,” it does not have the system does with persistence and plasticity? What takes place persistently and plastically? It is hard to avoid answering these question in terms of an already specified goal toward which the system tends, and toward which, we now wish to say, it tends in a relatively robust manner. For Schaffner, this amounts to begging the question; and at least those proponents of the systems view who consider it a reductive account of goal-directedness would have to agree.


20 Boorse (2002, 77) suggests that Schaffner finds the concept of a goal problematic because he thinks it is implicitly evaluative; he thus responds to the objection by denying that this is the case on the systems view of goals. As Bennett’s example shows, however, the problem of evaluation is deeper than Boorse seems to think. From an Aristotelian point of view, the assumption that nature is value-free hampers both Schaffner and Boorse.

21 Nissen 1997, 39–43.

function of detecting the seventeenth letter of the English alphabet.\textsuperscript{23} Causation, on the other hand, is not generally intensional: if, that is, $x$ causes $y$ and $y$ is identical to $z$, then $x$ causes $z$ as well. For example, “if hitting an iceberg caused the sinking of the Titanic, then hitting an iceberg caused the sinking of the world’s most luxurious ship.”\textsuperscript{24} Nissen concludes that “the intensionality of teleological concepts and extensionality of causation suggests that a purely causal analysis will not work.”\textsuperscript{25}

Thus both Schaffner and Nissen raise an important question to which, it seems, proponents of the systems view have no answer: Why, exactly, do we think of certain outcomes under certain descriptions as goals, while withholding this title from other outcomes, or even from the same outcome under another description? Because the systems view appears to beg this question, they conclude that there is no naturalistic or causal account of what it is to be a goal.

Before turning to an Aristotelian evaluation of the systems view and of these objections, it is worth noting that Boorse, in responding to several objections to his account of function, also provides a preliminary response to Nissen’s

\begin{itemize}
  \item \textsuperscript{23} Nissen’s discussion of this example shows at least that Adams’ example, at least, is poorly chosen, since the alphabet itself is a product of human convention, and its letters cannot be identified as such without reference to human purposes.
  \item \textsuperscript{24} Nissen 1997, 42–3.
  \item \textsuperscript{25} Ibid., 67.
\end{itemize}
He distinguishes between attributing a function to a given type of object, and attributing the same function to a particular token of that type. In a particular context, for example, a copy of Bonitz’s Index Aristotelicus may have the function of keeping other books from falling off a shelf. We would not, however, say that the function of the Index Aristotelicus in general is to hold books in place.

A similar distinction, however, applies to Nissen’s example of causation. It is certainly true that hitting an iceberg sank the world’s most luxurious ship; but we would not say, I think, that hitting icebergs sinks the world’s most luxurious ships. Just as Bonitz’s utility as a bookend has nothing to do with its content as a book, so the sinking of the Titanic by an iceberg has nothing to do with its being the world’s most luxurious ship. Causality, like function, is intensional.

The intensionality of causality, which Aristotle notes in Physics II.3 (196b25–8), also suggests a nonintentional justification for using terms like “goal” in our analysis of natural phenomena. Rather than pursuing this idea now, however, let us return to the systems view itself, this time from an Aristotelian perspective.

---

26 Boorse 1976, 83–90.
3. Aristotelian queries for the systems view

Especially through the concepts of plasticity and persistence, the systems view plausibly describes the goal-directedness of many systems and activities that we are already inclined to consider goal-directed. In doing so, it also tries to provide a clear and accurate distinction between truly goal-directed systems and other natural phenomena. Given the broad application of Aristotle’s teleology as I have presented it, it makes sense to ask how successfully the systems view restricts teleological concepts to the sorts of systems for which it was intended. In exploring this question, I shall argue that in fact, the plausibility of the systems view depends on its covert use of Aristotelian concepts that apply much more broadly. In fact, as I shall argue in chapter five, it may well turn out that mechanical phenomena in general have a teleological dimension. Otherwise put, Schaffner’s failure to find, in the systems view, a nonteleological definition of the term “goal” may be due not to the fact that the concept of a goal is irreducibly intentional, but rather to the fact that causality in nature is irreducibly teleological.

Before exploring this possibility, however, I want to prepare the ground by developing three Aristotelian comments on the systems view, each of which undermines the attempt to restrict teleological concepts to the “directively
organized” systems for which it is intended. The first comment, closely related to Schaffner’s objection, has to do with the crucial distinction between goals or ends, and mere effects. The second examines the alleged distinction between goal-directed behavior and the operation of the general laws governing a system’s parts. The third concerns the behavior of these parts considered in themselves, apart from their role in the larger system.

My first comment, then, is that upon closer examination the systems view’s criteria for goal-directedness turn out to depend on Aristotle’s requirement that an end be the καθ’ αὑτό effect of its cause.27 We have already seen that

27 In Nagel’s discussion, the first attempt to distinguish ends from mere effects comes in the initial “translation” of function-attributing statements that he proposes before turning to the analysis of goal-directedness. According to this proposal, the statement, “The function of $A$ in a system $S$ with organization $C$ is to enable $S$ in environment $E$ to engage in process $P$” is equivalent to, “Every system $[S]$ with organization $C$ and in environment $E$ engages in process $P$; if $S$ without organization $C$ and in environment $E$ does not have $A$, then $S$ does not engage in $P$; hence, $S$ with organization $C$ must have $A$” (Nagel 1961, 324). It is interesting to note that in light of Nagel’s intuitive starting point—that teleology has to do with “means-ends nexuses”—his task here is to give a nonteleological equivalent of the claim that for $S$ in $E$, $A$ is a means to end $P$. Although the translation he suggests is plausible, it involves two Aristotelian moves that render its claim to be nonteleological rather doubtful. First, his translation of “$P$ is an end”—that is, “Every system with organization $C$ and in environment $E$ engages in process $P$”—is clearly equivalent to Aristotle’s requirement that natural outcomes take place always or for the most part. Second, and correspondingly, his translation of “$A$ is a means” is, in Aristotelian terms, that $A$ is hypothetically necessary for $P$.

To the extent that Nagel’s translation is a reasonable one—which I do not want to insist on here—these Aristotelian features are quite significant. Aristotle’s teleology is, supposedly, the
in order to draw a clear line between teleological and nonteleological processes, 
an analysis of goal-directedness must incorporate a nonarbitrary distinction be-
tween mere effects, and proper goals or ends. The systems view tries to accom-
plish this through the requirements of plasticity and persistence, supplemented 
by the independence requirement. We may begin with this last requirement, 
which according to Nagel is the “formal criterion for distinguishing processes 
that are goal-directed from those which are commonly held not to be such.”28

The independence requirement, as I interpret it, requires that to count as a 
goal, an effect cannot simply result from the laws of nature that govern a sys-
tem’s parts. It must, rather, proceed from the structure of the system in such a 
way that the system’s behavior is not merely “law-like,” but also “directed” in a 
way that ordinary mechanical outcomes are not.29 Thus understood, I think it 
does successfully distinguish between systems that can be thought of as

---

28 Nagel 1977, 211.

29 For a stricter interpretation of the independence requirement as formulated in the lit-
erature, and for what appears to be a fatal objection to the requirement so interpreted, see above, 
n. 10.
goal-directed and those that cannot—although in my second main comment on
the systems view, I will suggest that it draws the boundaries of the former much
more broadly than proponents of the view think. What is required, again, is that
a system tend toward a given effect not merely insofar as its parts obey more
general laws, but rather of itself and as a whole, in virtue of its own organization.
In more Aristotelian terms, such a system has, in its organization, a principle
governing its behavior over and above those that govern the behavior of its
parts. To count as a goal for this system, therefore, an effect cannot simply result
from general laws that yield this particular result because the circumstances
happen to be right. It must, rather, have a καθ’ αὑτό cause in the organization of
the system taken precisely as a unified whole.

Turning now to the criteria of persistence and plasticity, it seems that their
value, as well, lies in the fact that they establish (or strengthen: Nagel, at least,
allows each to come in degrees) a καθ’ αὑτό connection between systems and
their states or effects. On the basis of this connection, we are willing to count
such effects as ends. Consider first the case of persistence. A system that can
compensate for internal disturbances, and thus tend toward a particular state or
behavior G by maintaining itself in one of several functionally equivalent sub-

192

sidiary states, does not just happen to arrive at or engage in G. In other words, G
is not the result of a prior state in which the system merely happens to be. Rather, at any given moment, G is the result of one of several alternative prior states that meet two conditions: (a) the system is structured to assume one or another of them, whenever possible, through the action of various internal cues; and (b) each of them has the same outcome, namely G, as an effect. As a matter of empirical fact, therefore, such a system is accurately described in virtue of its own structure as a G-maintaining system. Conversely, G is a καθ’ αὑτό effect of the system’s structure.

Turning now to the criterion of plasticity, we find that the case is similar. The plasticity of a system with respect to an outcome G ensures that this outcome does not result from circumstances in which the system merely happens to find itself. Rather, a goal-directed system tends toward G, by various means, in a whole range of circumstances. Plasticity, like persistence, strengthens the καθ’ αὑτό connection between the system’s organization and the effect we characterize as a goal or end. Thus, in fixing on these two criteria, Nagel has explicated in mechanical terms two ways in which a system as such—in virtue of its internal order or form, we might say—may tend toward a particular effect. This effect, by contrast with those that result from a chance environment or a random internal state of the system, he agrees with Aristotle in designating as an end.
From an Aristotelian point of view, the plausibility of the systems analysis lies in the fact that persistence, plasticity and the independence requirement all serve to identify, with respect to a system of a particular sort, a behavior or state that is typical of that system as a whole. The independence requirement, especially, requires that the system actually have a typical behavior: if its structure does not give rise to any predictable mode of activity, over and above the law-like activity of its parts taken separately, we cannot identify any particular state or activity of the system as an end.

In referring to the plausibility of the systems analysis, rather than to its success, I want to make clear that the Aristotelian viewpoint I am beginning to develop does not depend on the ability of the systems view to meet all the particular objections that have been raised against it. My point is simply that insofar as this view captures our ordinary thinking about goals, it does so by pointing to features of particular systems that guarantee a clear distinction between the behavior of the system as a whole, and the law-like behavior that is typical of the system’s components, independently of their belonging to this particular whole. In doing so, it directs our attention to that effect or outcome in terms of which the
system as a whole can be understood. This understanding, finally, is clearly teleological.\textsuperscript{30}

My second Aristotelian comment on the systems view also has to do with the independence requirement, particularly as formulated by Nagel.\textsuperscript{31} It seems to me that in distinguishing goal-directed processes from the action of general laws of nature, Nagel overlooks an important fact with which, in other contexts, he was quite familiar: that general laws alone do not account for any natural phenomena at all. The behavior of concrete objects and systems governed by general laws is always determined, more particularly, by local conditions. This means

\textsuperscript{30} One might object that such a system can be understood simply in terms of its structure and the laws that govern its parts, without reference to some outcome typical of the system as a whole. While admitting the possibility of a weak reduction of mechanical systems in many cases, I would argue that the objection either fails to exclude a reference to the behavior of the system as a whole, or involves a concept of structure that is useless from an explanatory point of view. A mechanical system is a physical object, and we always characterize physical objects in terms of their mechanical properties. In the alleged explanation, structure itself is already a mechanical concept, unless it is understood purely geometrically—which would be absurd. To describe a system’s actual structure involves describing the constraints placed on the behavior of its parts by their specific arrangement; and to the extent that the behavior of the parts is determined by the structure of the whole, it pertains to the whole rather than to the parts as such. See the following discussion of the relation between general mechanical laws and the particular structures in which they apply.

\textsuperscript{31} On the relation between my formulation of the independence requirement and Nagel’s (1977) reformulation of this requirement in response to Woodfield’s objection, see n. 10. If my contention there that Nissen’s (1997) objection misses the spirit of this reformulation is correct, then—as I am about to argue—the independence requirement works, but includes as goal-directed many systems that proponents of the systems view would like to exclude. If Nissen’s more literal reading is correct, the reformulated independence requirement turns out to be circular and therefore useless. However, Nissen’s conclusion is that Nagel fails to exclude systems that, from an Aristotelian point of view, are in fact goal-directed. This conclusion serves my purposes nearly as well.
that for any set of physical objects we choose to call a system, the behavior of those objects, considered precisely as parts of the system, does not result from the general laws by which the parts are governed. It is not that these laws cease to apply, or that they apply only in a modified form; it is simply that the behavior of a system as such is governed by its structure, and not by the nature of its parts in themselves.\footnote{I owe this observation to Michael Polanyi’s (1968) insightful discussion of the “dual control” exercised by general laws and circumstances in mechanical systems.}

Let us consider, as examples, three systems that Nagel himself mentions, and on which his own view is either explicit or at least fairly clear. The first is a room whose heating system is regulated by a thermostat; the second is a mechanical clock set to chime on the hour; the third is a marble in a hemispherical bowl.\footnote{See respectively Nagel 1961, 330; 1977, 205; 1977, 211.} On the systems view, only the first of these systems counts as goal-directed. In each case, however, as we shall see more particularly in a moment, the specific behavior of the system is due to its structure, rather than to the general laws governing its parts. As we consider each system, finally, it will be helpful to keep in mind that despite their overall similarity, the influence of structure on the system’s behavior is considerably more prominent in the heating system.
and in the clock than it is in the marble and bowl. This is an important point, to which we shall return.

To begin with the first case, that of the thermostat, we have here the simple, homeostatic feedback system that constitutes a parade case of the systems view. If the heat pump becomes less effective, for example, or if the outside temperature changes, the system, through the action of the thermostat, compensates to maintain a constant temperature inside the room. The constancy of temperature is due to the structure of the system as a whole, and this structure is such that the system’s behavior is both plastic and persistent. For Nagel, therefore, the system is obviously goal-directed.

The second example, that of the clock set to chime on the hour, is one that Nagel mentions only in his critique of Mayr’s concept of teleonomy. On his own view, simple mechanical clocks would not, in fact, count as goal-directed systems. First, a mechanical clock is not made to compensate for internal disturbances; rather, it is made so that internal disturbances are unlikely to arise. If they do arise, the clock needs to be taken out for repair. Moreover, only relatively sophisticated mechanical clocks are designed to compensate for fluctuations in the environment, such as changes in temperature, or in the momentum of the clock as a whole. Thus, a simple mechanical clock exhibits neither persistence nor
plasticity. Nevertheless, it seems odd to deny that the clock is goal-directed. For as a matter of fact, it is precisely so organized as to strike on the hour.\textsuperscript{34}

The third system, much less obviously goal-directed, consists of a single marble in a smooth hemispherical bowl. The behavior that Nagel considers is the marble’s returning to the bottom of the bowl, following any disturbance whatsoever that leaves it within the bowl itself. He denies, of course, that this behavior renders the system goal-directed, on the grounds that the “compensatory” force that brings the marble back to its resting place is simply a function, through the general laws of mechanics, of the displacing force. In this way, he claims, the independence requirement provides an “objective” criterion that coincides with, while at the same time rendering unnecessary, our “subjective” intuition that the system under consideration is not directly organized. Given the point I am now making about mechanical systems, however—and leaving aside the problematic nature of intuitions, on which I fully agree with Nagel—it should be clear that the specific behavior of this system is due most immediately not to any general laws of nature, but rather to the shape of the marble, the shape of the bowl,

\textsuperscript{34} The example of the clock brings up a general difficulty with the systems view. According to Nagel, both persistence and plasticity come in degrees. However, there is an intuitive problem with this admission. Goal-directedness, Nagel claims, is a quantifiable property of a system, the measure of which corresponds to the system’s degree of persistence and plasticity. On this view, a system possessing even the least degree of persistence or plasticity counts as directly organized. Our clock, however, despite being just so constructed as to chime on the hour, does not count as directly organized in any degree, and this seems a bit odd.
and the location of the former inside the latter—that is, to the structure of the system as a whole.

In light of my first comment, concerning the necessity of a καθ’ αὑτό connection between a system’s structure and any behavior or outcome that counts as a goal, the preceding discussion of these three examples suggests that from an Aristotelian point of view, each of the systems discussed is goal-directed.\(^{35}\) And although intuitions do differ, the claim that the displaced marble’s behavior must be understood teleologically is unlikely to go over well with most readers. This reaction, however, is perfectly understandable. Recall that in introducing the three systems, I suggested that in the third and simplest, the influence of overall structure on the system’s behavior is less prominent than it is in the other two.

This is because in a simple, loosely organized system, the effects of the general

---

\(^{35}\) In analyzing these systems, I have focused on the καθ’ αὑτό connection between structure and activity that each of them exhibits. A further consideration, indispensable from an Aristotelian point of view, is whether any simple mechanical system counts as a substance. (For a preliminary discussion of what counts as a substance, for Aristotle, see the appendix to chapter one.) Without attempting to address this question as it requires, I am inclined to think the following: (1) Almost certainly, none of the my three examples counts as an Aristotelian substance; however, each is “substance-like” insofar as it is only weakly reducible (see above, n. 28); finally, because the strength of a καθ’ αὑτό connection between structure and behavior varies with the explanatory importance of the structure relative to that of the components, a weakly reducible, merely substance-like system should be thought of as only weakly goal-directed (cf. below, n. 36). (2) Because the explanatory importance of structure varies more or less continuously among systems whose structures differ, Aristotle’s concept of nature as form does not yield a sharp distinction, among mechanical systems, between substances and more or less substance-like systems. (3) The relevant structural property for making a rough distinction between substances and substance-like systems is the integrity or “wholeness” of the system, which should be understood in terms of the relative nondecomposability of its behavior.
laws governing the system’s components are much less restricted by its overall structure than they are in a more tightly constructed system like the heating system or the clock. Thus, the directedness of the system as a whole, if we may speak thus, is minimal. Nevertheless, the influence of structure cannot simply be ignored. Although the marble may follow many different paths after a disturbance, the specific outcome we are now considering—its coming to rest at the bottom—cannot even be expressed without implicit reference to the shape of the bowl. The system is so organized that, beginning from any one of a limited variety of states, its structure directs the marble to a state of rest at the bottom of the bowl. Thus, I would argue, although it reasonable to maintain that the goal-directedness of such a system is minimal, or even trivial for ordinary purposes, it is arbitrary to deny it altogether.

In considering various degrees of organization and goal-directedness, we have come to a topic quite similar to one we have already seen in discussing Meteorology IV.12. There, Aristotle recognizes that the presence of a proper ἔργον in the elements and the homogeneous bodies is not obvious. This is not, however, because there is no such ἔργον, but because the formal determination of such bodies to a particular way of being and activity is relatively minimal. To deny the presence of a distinct form and corresponding function altogether, however,
would be to deny the existence of these substances. In the same way, a mechanical system will be goal-directed precisely to the extent that it constitutes a unified whole over and above its parts. The greater a system’s unity—and hence the more obvious the causal role of its structure, as distinct from the causal roles of its parts—the more natural it will seem to consider the system goal-directed, because in fact, the more goal-directed it will be.36

My third and final comment has to do with the notion of directedness itself, and with the general laws governing the parts of a mechanical system. In considering the relation between general laws and the various structures within which they can operate, I have just suggested that the effect of such structures is to impose on the behavior of a system’s parts a directedness that is not to be had from the general laws governing these parts alone. The degree and manner of directedness depends on the structure itself, but to the extent that an overall structure can be identified, it not only affects the parts but also determines the existence of a larger whole, with specific characteristics and typical behavior of

36 There is, of course, an important question concerning how to individuate systems. Although a great deal could be said on this point, the basic answer is relatively straightforward. In the case of a highly unified system such as an organism, the distinction between the system and its environment—between items that are part of the system and those that are not—is relatively clear. In other cases, the decision to consider a number of objects as parts of the same system may depend largely on our own interests. On the Aristotelian account of goal-directedness that I am now developing, any καθ’ αὑτό effect of any system can be considered a goal in at least a minimal sense. However, only systems demarcated by observation and causal analysis provide a fully objective basis for goal attributions.
its own. Given this much, however, it is easy to see that the very possibility of a mechanical system presupposes the existence of parts that have at least some degree of directedness independently of the structures in which they may happen to be found.

In speaking here of a system’s parts, I am referring to any component of the system that can exist independently of it. In both natural and artificial systems, these may include “generic” materials such as iron, or an amino acid ingested with an animal’s food; they may also include specific parts that come to be within the system, as do many proteins, or at least with a view to its manufacture, as the hand of a clock. In claiming that such parts have a certain directedness independently of the whole system, I mean only that each one has a determinate structure and capacities, which are independent of the larger system but on which this system depends.

To the extent, in other words, that a given part of a system is itself a whole, and can be considered as a unit, its behavior will depend on its own structure; in this way, there is a certain analogy between the system as a whole and its parts, insofar as these also can be considered as wholes. It may be, of course, that at some level of physical analysis our ordinary notion of structure becomes problematic. In such a case, however, we could just as easily say that the behavior of
each part, considered separately, is governed by the sort of unit or whole it is: by
the properties we attribute to it as such. Its behavior is open to further determi-
nation, but this determination depends on that which is already present.

Whatever goal-directedness a system as a whole may exhibit, therefore, is
dependent not only on its structure but also on the prior directedness of its parts,
considered in themselves. Indeed, the possibility of a system of any sort depends
on the existence of potential components of that system whose behavior is rela-
tively determinate, yet open to further determination by their place in a larger
whole. To state this is, of course, simultaneously to stress that the parts are less
directed in themselves than they are as parts. But this, to return to Meteorology
IV.12, is precisely what Aristotle tells us in the context of his own science.

4. Conclusion

To conclude this chapter, it will be helpful first to summarize the forego-
ing discussion, and then to evaluate briefly the concept of goal-directedness that
has emerged from it.

In summarizing the core tenets of Aristotle’s teleology, in section one, I fo-
cused in particular on two central concepts: that of a καθ’ αὑτό connection be-
tween a process or activity and its cause, and the distinction between actuality
and potentiality. In virtue of the former requirement, finality and agency often
turn out to be correlative causal factors—indeed, the presence of a καθ’ αὐτό agent is sufficient for an action to be goal-directed. The goal or end of such an action is an actuality corresponding to the potentiality of the agent to bring it about.

On the systems view of goal-directedness, as we saw in section two, there is a twofold relation between teleological concepts and mechanistic causal analysis. On the one hand, teleological concepts and explanations are restricted to systems that exhibit persistence and plasticity, rather than forming part of our understanding of nature more generally. On the other hand, the systems analysis makes goal-directedness compatible with an ordinary causal analysis of systems that appear to be goal-directed.

In section three, finally, I argued that when the goal-directedness described by the systems view is properly understood, it can no longer be limited to the sorts of system for which proponents of the view intended it. Rather, it applies in various ways and degrees to any system that, in virtue of its structure, has a typical behavior of its own, distinct from the behavior typical of its parts when these are considered separately. Such goal-directedness, moreover, presupposes an analogous, though less complete, directedness in the parts themselves.
On the account of goal-directedness that has emerged in this chapter, therefore, a system is goal-directed if, in virtue of its structure, it possesses a typical mode of behavior that is distinctive of the system as a whole.\textsuperscript{37} By contrast, a system fails to be goal-directed to the extent that its behavior is simply the sum of the random or unconstrained interactions of its parts. In Aristotelian terms, a system is goal-directed if it is responsible \(\kappa \alpha \theta \, \alpha \upsilon \tau \omicron \), in itself and as such, for a specific state or activity, which state or activity thereby counts as a goal. Such a goal is also, in accordance with the conclusions of chapter two, an actuality proper to the system, but this will be easier to see in a moment.

In section two, we considered two objections to the systems view—or, more accurately, to any nonintentional analysis of goal-directedness. In order to highlight the various dimensions of the account that I have just offered, let us see how it stands up to these objections. First, according to Nissen, causal and teleological explanations differ in that function and goal-directedness are intensional, whereas causality is not. Now we have already seen that in fact, causality is just

\textsuperscript{37} The weak reducibility of many systems (see n. 28, above) would seem to limit the importance of such goals for our understanding of the system as a whole. As suggested by our consideration of the three examples from Nagel above, a more thorough typology of goal-directed systems, noting the different ways in which a system’s parts and its overall behavior can be related, would indicate that the relative importance of a system’s goal for our understanding of the system will depend on the system’s structure. In chapter five, however, I shall argue that even simple mechanical phenomena such as gravitational motion have a teleological dimension. If this is the case, then even systems whose proper goal-directedness is relatively limited will depend on the goal-directedness of mechanical phenomena in general.
as clearly as intensional as teleology. For while it may be true that an iceberg once sank the world’s most luxurious ship, this formulation does not reveal the causal structure of the event. It is not the case that an iceberg sank the world’s most luxurious ship, as such. Causality is intentional, therefore, and in this context, the contemporary notion of intensionality is equivalent to Aristotle’s concept of what is καθ’ αὐτό.

Applying this point about causality to our discussion of systems, let us say that a system $S$ is in some way responsible for an outcome $G$. Obviously, this is not enough to say that in general, systems of type $s$ tend toward outcomes of type $g$. However, let us postulate that this is in fact the case; we still need to know why it is so. Unless it is the case that systems of type $s$ as such, in virtue of their common structure, tend toward outcomes of type $g$ as such, we do not have a clear and precise statement of the causal connection. For example, we do not understand the causal connection between thermostat-regulated heating systems

---

38 For an extremely rough, initial approximation of a causal explanation, we might try the following: (a) An iceberg, i.e., a large, massive, floating object, ripped through the outer shell of another large, floating object; (b) the buoyancy of the second object (we may call it the *Titanic*, but keep in mind that this is irrelevant) was due to the air trapped inside its outer shell; (c) the breaching of the *Titanic’s* outer shell allowed some of the water in which it was floating to displace the air that had been trapped inside, (d) thus changing the *Titanic’s* buoyancy, so that it sank. The fact that the *Titanic* was a ship, and even more that it was the world’s most luxurious ship, is entirely irrelevant.
and the maintenance of a constant interior temperature until we can say in virtue of what structural features, precisely, the heating system produces this effect.

When we have isolated a causal connection of this sort, however, we are not merely saying that \( G \) is an effect of \( S \), or of \( S \)’s structure, as the demise of the world’s most luxurious ship was the effect of an iceberg. Rather, we are saying that \( G \) is an effect of a certain sort: an effect in terms of which we can understand \( S \) as a whole, as a system of dynamic parts whose individual and joint activity is determined by the system’s overall structure. \( G \), in other words, is an outcome at which \( S \) aims.\(^{39}\) If we leave aside the contingent connotations of particular English terms, it makes perfect sense to think of such an outcome as a goal, and to say that the system is ordered to this goal. Thus we have not only answered Nissen’s objection, but have also provided Schaffner with grounds for speaking of goals apart from intentions. A goal is that at which something aims, and aiming does not required intentionality.\(^{40}\)

\(^{39}\) “An outcome at which \( S \) aims”: nothing we have said so far excludes the possibility that a given system may aim at more than one outcome.

\(^{40}\) One might, of course, insist that words like “goal” and “aim,” are applied properly only to intentional goals, aims, and so forth. Given the need to distinguish between effects in general and proper effects, however, these terms become quite useful in nonintentional contexts as well. If confusion seems likely, we might use impersonal expressions such as “is aimed at,” “is directed to.” In any case, the language we feel comfortable with does not change the fact that we understand systems in terms of their activity. If this fact raises a metaphysical problem, it is a problem we shall have to deal with. (On the possibility of reducing the behavior of a system to that of its parts, see n. 28, above. In chapter five, I shall argue at some length that the same problem arises in the case of simple, law-like behavior such as gravitational motion. On the problem
At the end of section three, finally, I observed that the existence and structure of a system depend on the prior directedness of the items that serve as its parts. This raises a new question, to which we can now turn: If the directed behavior of mechanical systems should be understood teleologically, what about the behavior of physical objects in general? In the first part of chapter five, we shall address this question by considering the most basic phenomena of classical mechanics, namely, simple gravitational motions. In the second part of the chapter, after proposing a final account of goal-directedness in general, I will sketch a number of interestingly different ways in which such goal-directedness seems to appear in nature.

of backward causation, see the preliminary remarks on Aristotle in the appendix to the present chapter.)
CHAPTER FIVE

TELEOLOGY AND LAWS

In chapter four, I articulated what might be thought of as an extremely permissive version of the systems view of goal-directedness. In particular, I argued that if systems of type $s$, in virtue of their common structure, tend toward outcomes of type $g$ as such, then instances of $g$ are goals for instances of $s$. Goals, thus understood, are not simply effects or outcomes; they are characteristic of the systems to which they pertain, and in fact enter into our understanding of such systems. At the end of the chapter, recalling the relation we had observed between the directedness of a given system and the prior directedness of its parts, I suggested that we now ask whether physical objects in general have a teleological dimension. In other words, in the previous chapter we considered the constraints placed by a system’s structure on the law-governed behavior of its parts, whereas in much of the present chapter we shall consider law-governed behavior as such.
In section one, I begin by reviewing the role of laws, especially causal laws, in recent discussions of scientific explanation. Then, in section two, I argue that the explanatory role of what I shall call basic laws is implicitly teleological, and that their value for scientific explanation thus reveals a teleological dimension in the natural phenomena that they describe.\(^1\) In section three, finally, I begin by turning from the discussion of laws to the important issue of what, exactly, is to be gained by adopting a teleological approach to natural phenomena in general. Then, after suggesting a flexible, general account of what is involved in all types of goal-directedness, I consider eight different respects in which particular forms of goal-directedness might differ from each other. Finally, in section four, I conclude by returning to the topic of motivation, reviewing the basic steps in a teleological approach to nature, and pointing out the problems involved in trying to exclude teleological concepts and language, or to restrict their use to the most obvious and intuitive cases of goal-directedness.

1. Laws and explanation

Whatever the difficulties associated with the concept of a law of nature, the importance of such laws for scientific explanation is undisputed. This is evident if we consider the two most important recent accounts of explanation:

\(^1\) A slightly fuller preview of section two will be provided in section one.
Salmon’s causal/mechanical approach, and Kitcher’s unification approach. According to Kitcher, a scientific explanation is a (deductive) derivation of a certain sort, and a law is a universal premise that occurs in an explanatory derivation.\(^2\) The importance of universal premises in deductions, of course, requires no comment. Moreover, Kitcher holds that explanations of this sort aim to present us with the order of the natural world, including its causal order, so his concept of a law seems to be fairly robust.\(^3\)

In articulating his concept of causal/mechanical explanation, Salmon deliberately avoids relying directly on the concept of law. He makes no claims, however, to have eliminated this concept altogether, and in fact proposes an “ontic” interpretation of explanation by means of antecedent conditions and laws.\(^4\) In his late work, moreover, he recognizes the importance of unification-style explanations as complementary to, rather than conflicting with, causal/mechanical

\(^2\) For Kitcher’s most detailed account of the unification view see his 1989; on laws in particular, see ibid., 477.

\(^3\) Kitcher (1989, 494–500) proposes a Peircean account of scientific progress that identifies the structure of the world with that structure presented by our best theories at the ideal limit of inquiry. In his 1993, he suggests that both robust realism and the Peircean view are strong enough to support his accounts of scientific explanation and progress. If Peircean “internal realism” (Putnam’s phrase) is the weakest epistemological position compatible with Kitcher’s account of explanation, it seems that his explanatory universal premises cannot be understood as statements about mere regularities. For a brief discussion of the distinction between explanatory and nonexplanatory regularities see his 1989, 426–7.

\(^4\) Salmon 1989, ch. 1; 1998, 54, 58.
Moreover it is not hard to see that even explanations of the latter type depend at least indirectly on laws. In stressing the need for causal explanation, Salmon distinguishes causal from noncausal laws. As examples of the latter he mentions the ideal gas law and Kepler's laws of planetary motion, both of which he takes to be merely descriptive. In both cases, moreover, the explanation of the noncausal regularity involves an appeal to causal laws. Kepler's laws are reduced immediately to the laws of Newtonian mechanics; whereas the ideal gas law is explained by postulating an underlying mechanism, and then deriving the noncausal law from the properties of this mechanism, which include the applicability of various causal laws. This second case is particularly interesting, because the explanation Salmon offers is clearly a causal/mechanical explanation, yet it depends on the articulation of laws.

---

6 Ibid., 55, 59–60. Salmon does not discuss explicitly what he means by “causal” and “noncausal” in talking about laws, but his two examples are helpful. Presumably, he thinks that a causal law is a law that directly describes a causal process, whereas a noncausal law states a regularity that must be understood in terms of distinct, underlying causal processes.
7 For this reason, I think that Kitcher has the upper hand when he argues concerning such cases that “what counts as a fundamental mechanism is not dictated by the local details of the individual causal processes that occur in these episodes. Rather, the fundamental mechanisms are disclosed by our most general theoretical accounts of a range of regularities” (1989, 430). For example, to describe the mechanism that underlies the ideal gas law—leaving aside the assumptions involved in the use of statistical methods—we use extremely general laws to explain both the interactions among molecules, and their inertial motion between one interaction and the next.
In short, the best recent discussions of scientific explanation reveal the continuing need for an account of the explanatory value of laws. The main challenge facing any such account is to clarify the puzzling connection between causal laws and the particular occurrences that conform to them. As Giora Hon has recently put it, explanation by means of laws involve “the imposition of a pattern, a scheme or an order on phenomena from without.” In other words, there is a gap to be filled between a particular, concrete phenomenon that we say can be explained by reference to a given law, and the general, often quite abstract form of behavior or interaction specified by the law itself. Any progress that we can make toward filling this gap will be an important step forward in our understanding of scientific explanation, and consequently also of the natural world.

In section two, I provide an initial argument that the appeal to causal laws is, in fact, a teleological mode of explanation. My discussion there focuses on what I will call basic laws, of which my example is the statement that objects in spacetime tend to move along geodesics. The notion of a basic law will be refined in the rest of the present section; basically, however, such a law states that objects in a given class tend to instantiate a certain simple, abstract form of behavior. In

---

8 See Cartwright 1989, 350; Humphreys 1989, 289; cf. Hon 2001, 8–9. I shall use the expression “causal law” for any law that is thought to explain, rather than merely describe, the concrete phenomena that it governs, and I shall assume that explanation is not a purely logical affair. More than this, I think, will not be necessary for the rough speculations that follow.

9 Ibid., 8.
explanations involving basic dispositional laws, I shall argue, the form of activity described by the law functions as an end in a recognizably Aristotelian sense. Correspondingly, moreover, the relation between these abstract forms of behavior, and the actual phenomena by which concrete objects instantiate them, is appropriately understood as a relation between ends and means. Finally, the teleological structure of basic laws appears even more clearly in phenomena that result from the influence of two or more distinct causal factors.

To refine the notion of a basic dispositional law, we can begin with Nancy Cartwright’s argument that causal laws are capacity ascriptions. In addressing the relation between a given law—her example is “Aspirins relieve headaches”—and a singular causal claim falling under it, Cartwright suggests that “the causal law attributes a capacity to the featured characteristic—say, being an aspirin—a capacity to produce the appropriate effect in individual cases.”¹⁰ Now clearly, “Aspirins relieve headaches” is not, or does not express, a law that we would want to call basic. In fact, Cartwright has deliberately chosen a “law” whose accuracy depends on a complex sequence of causes and effects, because she is interested precisely in the relation between capacity ascriptions, which can often be stated quite simply, and laws about actual causings, which take the form “in

such and such particular circumstances, x’s cause y’s,” and which may not even be finitely expressible.11

Although there appears to be some overlap between my analysis in the following section and Cartwright’s, I want to pursue a slightly different strategy by focusing on the notion of a capacity ascription, or law, that is nonanalyzable and therefore counts as basic.12 That there are such ascriptions, or at least that there are basic capacities, is suggested by Salmon’s discussion of causal and non-causal laws, on which I have already commented. His two examples suggest two ways, already well known, in which a law can be analyzed. First, as in the case of Kepler’s laws of planetary motion, the law can be derived from more general laws by specifying a restricted class of circumstances in which the more general laws are taken to apply. Second, as in the case of the ideal gas law, we begin by analyzing not the law itself, but the system to which it applies. This analysis reveals parts to which other laws apply, and these laws may themselves be analyzable, in either of the ways just sketched.

A basic law, then, is in the first place a law that can be analyzed neither by showing that it follows from a more general law or laws, nor by analyzing the

11 Ibid., esp. 355.

12 The term “capacity” is not always appropriate: in the case of gravitational motion, for example, we will generally prefer to speak of a tendency or a disposition. These notions all require further discussion, but in these pages, unfortunately, I shall have to take them as basic.
systems to which it applies. We need to add, however, that such a law counts as basic only if our theories indicate that to search for an analysis would be inappropriate, because the law in question expresses what the theory presents to us as an irreducible fact about the world. Newton’s three laws of motion are basic laws in this sense, as is his law of gravitation. In section two, I shall use as an example the statement that replaces Newton’s law of gravitation under general relativity: that objects in spacetime tend to move along geodesics.13 In any case, the concept of a basic law should make it clear that the analysis found in section two is not intended to provide a general account of laws of nature, but to advance our understanding of what appears to be a relatively simple case.

Another helpful distinction is between the use of laws as background information in causal/mechanical explanations, and the direct appeal to a law as itself explaining a class of phenomena. For example, when we explain a baseball’s flight by appealing to the action of the batter, we presuppose a number of laws, or at least informal, law-like beliefs.14 However, what we appeal to as directly explanatory is not any of these laws, but a particular event: the batter’s hitting the ball. On the other hand, to explain the ball’s roughly parabolic motion

13 Although it seems reasonable to look for basic laws or something like them in other theories, to the extent that these have a deductive structure, the existence of such laws outside classical (that is, Newtonian or relativistic) mechanics does not concern us here. I have introduced the concept only to clarify what sort of law I have in mind in the following discussion.

after leaving the bat, we appeal not to a particular event but directly to a law, a
law that redescribes the parabola in terms of features of the world that we con-
sider more basic. This is not a causal/mechanical explanation, but a species of
unification;\textsuperscript{15} in the following discussion, I shall be concerned only with explana-
tions of this type.

In the following section, therefore, my analysis of the relation between
singular events and the laws that govern them concerns those cases in which we
explain a particular event, in particular circumstances, by appealing to a basic
law. In particular, we shall be considering the explanation of particular motions
of objects in spacetime in terms of the law that such objects tend to move along
geodesics.

2. Laws and teleology

To understand better the explanatory role of basic laws, it will be helpful
to begin with a lawlike statement of dubious explanatory value. So, consider the
behavior of medium-sized bodies close to the surface of the earth. All such ob-
jects, if not obstructed or resisted by other solids or fluids, accelerate toward the
center of the earth at approximately 9.8 meters per second per second. There are,

\textsuperscript{15} For a similar example see Woodward 1989, 360; on the explanatory redescription of
lawful behavior see Kitcher 1989, 427.
of course, better ways of describing this motion, but this is enough to begin with:
it involves a well-defined behavior exhibited, always or for the most part, by a
certain class of objects: that is, medium-sized objects near the surface of the earth.
Moreover, not only in ordinary language but also in our scientific descriptions,
we recognize a connection between this behavior and the objects it characterizes
that is not merely incidental. Rather, the phenomenon of gravity enters into our
basic understanding of familiar physical objects.

In the remainder of this section, I want to explore the possibility of a teleo-
logical understanding of basic mechanical phenomena by refining and reflecting
on the very local and approximate description of gravitation just given. We can
begin this exploration by stating several obvious problems with this description,
and in particular with the alleged \( \kappa \alpha \theta' \ \alpha \upsilon \tau \omega \) connection between ordinary physi-
cal objects and acceleration toward the center of the earth. First, the motion of
bodies near the earth toward its center is frequently impeded by the intervention
of other, fixed bodies. Second, even when this is not the case, a falling object al-
most never actually moves downward at the rate given above. Rather, such mo-
tions generally take place in a fluid medium, and are slowed accordingly. Fi-
ally, and most importantly, the tendency to move toward the center of the earth
in a particular way is, we now know, relatively incidental to bodies that possess
it. Move them just a bit, in cosmic terms, and the tendency is significantly modified; move them a bit more, and it becomes irrelevant.16

Of these three difficulties, the first is not particularly interesting; after all, Aristotle’s qualification “or for the most part” plays an important role in his discussion of nature. It means, basically, “when circumstances permit,” and captures not only an obvious fact about natural regularities, but also Aristotle’s conviction that natural outcomes are not, generally speaking, necessary (Phys. II.9; Metaph. E.6, Θ.5). The second and third difficulties, however, do raise important issues, reflection on which will help clarify what is involved in a teleological interpretation of phenomena like that of a falling body. I will address the last first, at some length, and then the second.

In Physics II.3 (196b25–8), Aristotle stresses the need to find appropriate descriptions of causes and effects, such that the καθ’ αὑτό relation between them is made clear. Thus we do not assign the activity of building to a human being as such, but only to a builder; likewise, we do not assign a particular sculpture to the class of sculptors in general, but to the particular sculptor who made it.

16 “Tendency”: Newton sometimes spoke of gravitational attraction; at other times he insisted that his analysis was purely mathematical and descriptive. Uneasy with the idea of action at a distance, he also speculated about additional mechanisms—ethers, for example—that might provide an acceptable causal explanation. In the theory of general relativity, the concepts of force and attraction no longer appear in descriptions or explanations of gravitational motion. The theory takes it as basic that objects in spacetime tend to move along geodesics, a geodesic being the shortest line between two points in spacetime.
Given this basic requirement, it is easy to see why we are reluctant to say that the goal of an object in spacetime is to move toward the center of the earth. For in fact, it is only in particular circumstances that a body may tend in this direction. On the other hand, there are two other possible descriptions in which this problem is resolved. Either we can say that any object occupying a similar region of spacetime tends to move in a similar way, or, avoiding such particulars, we can simply say that all objects in spacetime tend to move along geodesics. Either way, we are giving a precise description of certain bodies in terms of an activity that corresponds to them as such.

We thus have two sorts of description of gravitational motion—one fully specified, the other abstract—that reflect a καθ' αὑτό connection between the relevant characteristics of a moving object and its motion. Although this is already suggestive from an Aristotelian point of view, we need to reflect further on the structure and the relative explanatory value of these descriptions. First,

17 The following discussion will be framed mainly in terms of general relativity, but Newtonian mechanics would work just as well, if we are careful not to interpret statements about gravitational force as describing a causal mechanism. (It seems to me that “force” is simply a useful term for describing how the relations among bodies affect their motions; however, the theory of relativity allows us to avoid this issue, at least with respect to gravity.) Thus in a Newtonian framework, we can say either that two bodies of masses $m_1$ and $m_2$ at distance $d$ tend to accelerate toward each other at a given rate, or that massive bodies in general tend to accelerate toward each other at a rate that varies in accord with the product of their masses and the inverse square of the distance between them.

18 The topic of explanatory value brings us to a common objection to teleological explanations: When asked why bodies move along geodesics, it is famously unhelpful to reply solemnly
consider again the statement that objects in spacetime tend to move along geodesics. When we apply this statement to a particular object, we describe both its circumstances and the resulting motion in relatively abstract terms. If, however, we refer to the moving body not simply as an object in spacetime, but as occupying a region of spacetime with specific characteristics, then we will also provide the specifics of its motion. In stating the tendency that our theory attributes to a given object, therefore, we can appropriately use various levels of abstraction, as long as we are consistent throughout.

Now in comparing these two sorts of descriptions—the abstract formula, on the one hand, and a description with all the variable quantities specified, on the other—Aristotle’s explicit discussion can only take us so far. His general principles of explanation, however, are quite helpful. In *Physics* II.3, as we have seen, he states that we must correlate general causes to general effects, and particular causes to particular effects: thus we attribute statues in general to sculptors, but this particular statue to the sculptor Polyclitus. This distinction is not quite the one we are now considering, for the relation between general and particular causes has a tendency to move in just such a manner. Phillip Sloan has pointed out to me that Molière’s satire is actually directed at the Galenic doctrine of causal “faculties,” and that Galen uses such faculties as placeholders for unknown efficient causes (see *Natural faculties*, I). In any case, explanations of this sort cannot be blamed on Aristotle. As he writes on a related topic in *Metaphysics* Z.17, “to ask why a thing itself is itself is to ask nothing” (1041a14–15). Nor is his cause for the sake of which intended to explain itself. Rather, it is intended to explain phenomena that are distinct from it, but take place because something is acting for the sake of it.
particular claims at the same level of specificity is not explanatory. We do not, that is, explain a concrete, particular phenomenon simply by stating that it always happens thus. That sculptors in general make statues does not explain—except perhaps in a trivial sense—why Polyclitus made this statue; and that bodies in certain well-specified circumstances tend to move in such and such a manner does not explain, except in the most trivial sense, why a particular body in just these circumstances moves in just this way.

We do, however, think that we can explain a well-specified phenomenon—be it the actual motion of a particular object, or more generally the specific motion of any object in these circumstances—by appealing to the appropriate abstract law. The abstract formulations of physical theory, in this case of general relativity, do explain, because they tell us, as a mere generalization does not, what are the causally relevant factors that account for a set of natural phenomena. Thus to explain the downward motion of a particular body on earth, it is not enough to have observed that it fits a specific pattern, namely, that all bodies near the earth tend to move downward in the same way. A reference to geodesics, however, does explain; specifically, it explains the particular features of

---

19 This is a point of general agreement since the demise of purely logical accounts of scientific explanation: see for example Hon 2001, 7; Kitcher 1989, 426–7.
the motion we observe in terms of other, distinct features of the moving object’s situation.\textsuperscript{20}

To see in what respect such an explanation is teleological, we can begin by contrasting it with another common form of explanation. As a preliminary point, we can recall that on the deductive-nomological model and its various successors, a complete explanation of a particular phenomenon refers both to a general law and to antecedent conditions. In practice, of course, one or the other of these factors may be taken for granted, and we can explain the phenomenon in question by stating the other. In the present case, we can explain an instance of gravitational motion either by referring to the law, taking the local conditions for granted; or by referring to the conditions—that is, the shape of a given region of spacetime—taking the law for granted. To add another classic example, we can explain the acceleration of a billiard ball either by pointing to the ball that hit it, taking the relevant laws for granted, or by citing the laws that govern the interaction of these two balls, taking their contact for granted. In each case, we have two explanatory factors of quite different types.

\textsuperscript{20} The abstract formulation does not, of course, explain its own structure: thus we have explained neither the fact nor the direction of the acceleration, and we have discovered, but not explained, the way in which the form of particular instances of gravitational motion is affected by certain factors. We have, however, explained precisely what is derived from the explanatory formula and the specific masses and distance in a particular case: namely, the specific form of motion observed.
What I want to suggest is that when we explain by appealing to an abstract law, taking the relevant antecedent conditions for granted, we are explaining a concrete phenomenon by referring to the general form of activity proper to the object or objects involved. By “proper,” I mean simply that according to the relevant theory, there is a καθ’ αὑτό connection between this sort of behavior and objects of this type; and this is, of course, just the sort of connection that we can express in general laws.21 Thus when we cite an abstract law as explanatory of gravitational motion, we specify a mode of activity that belongs to objects existing in spacetime, as such. We state, in other words, how all such objects tend to behave. From an Aristotelian point of view, however, such an activity counts as an end. I shall develop this interpretation further in a moment, after dealing with a traditional objection.

The suggestion that the type of motion predicted by general relativity somehow counts as an end brings us to a difficulty that has beset teleology since the first formulations of modern mechanics: namely, the association of goal-directedness in the behavior of inanimate objects with Aristotle’s doctrine of natural place. Aristotle, it is well known, considered the natural ends of heavy

21 Sometimes we end up defining the class of objects in terms of this characteristic behavior. Whether or not this is the case may depend on a number of factors that are irrelevant from the explanatory point of view. The important question is whether the behavior as we describe it is actually characteristic of the objects to which we attribute it, and this question is settled by experimental practice together with any relevant theoretical considerations.
and light bodies to be places: the center and perimeter of the sublunary sphere, respectively. It is thus unsurprising, in hindsight, that the well-motivated rejection of this doctrine of natural place should have made difficulties for Aristotle’s teleological view of local motion. The doctrine of natural place, however, is only part of Aristotle’s account of local motion, and other aspects of his thought reveal that his teleology in general is separable from this doctrine. For example, as mentioned in chapter one, Aristotle thinks of the circular motion of the heavenly bodies much more as a complete activity than as a process: it has no end outside itself, but rather is precisely that for the sake of which the heavenly bodies exist. Thus it seems that from an Aristotelian point of view, the breakdown of the distinction between heavenly and earthly bodies, and the corresponding rejection of the notion of natural place, should not have reflected poorly on teleology. Rather, it should have led to a better understanding of the specific, teleological structure of the natural motion of bodies, along the lines that I am now suggesting with respect to gravitation. Rather than being thought of as altogether nonteleological, the motions associated with gravitation should be understood in the same terms as are complete activities like natural circular motion or sight.

In chapter two, we saw that Aristotle distinguishes complete and incomplete activities in terms of their relation to the ends that they involve. An
incomplete activity is a process toward a determinate end, as house-building is a process that terminates in the existence of a house, or the development of an organism terminates in the structure and capacities of the mature specimen. A complete activity, by contrast, already includes the end; in other words, it is itself an end. Seeing and hearing, for example, are complete activities—as is the circular motion of the heavenly bodies, as we have just recalled. This distinction can be recast more perspicuously, from a metaphysical point of view, and connected with Aristotle’s teleology, in terms of actuality: an incomplete actuality is incomplete by comparison with its end, whereas a complete actuality is itself an end. For various reasons, in Aristotle’s cosmology only circular motion, out of all changes in place, could be a complete activity, and the doctrine of natural place depends on this cosmology. In modern mechanics, by contrast, spatial motions are generally understood as complete activities, in Aristotle’s sense—but this does not entail that such motions have no teleological dimension.22

I now want to develop more clearly the claim that simple gravitational motions do in fact have a teleological structure. I have already suggested that when we explain a concrete phenomenon by adducing a basic law, we are specifying a telic mode of activity: a way of acting or behaving that is proper to the

22 Note that in Newtonian mechanics, acceleration is an incomplete activity in Aristotle’s sense.
objects governed by the law, and in terms of which we understand these objects, the behavior we actually observe in them, and their place in the causal structure of the world. Two further considerations will, I hope, serve to elucidate this analysis and provide reasons for taking it seriously. Both involve reflecting further on the relation between the actual behavior of concretely existing objects, and the various abstract formulae in terms of which this behavior is understood. The first returns us to a question I raised in discussing the systems view, concerning the identification of ends in general and their distinction from effects that are not ends. The other brings us to the second of the three difficulties I raised above, having to do with the influence of distinct causal factors on particular phenomena.

As we have seen, Nagel begins his discussion of teleology by calling our attention to “expressions signifying a means-ends nexus.” I have already suggested, moreover, that the important thing here is to distinguish outcomes that we are willing to consider ends, from those that are mere effects. Another way of putting this same point is that without an end there can be no means; rather, there are simply interacting objects and successive events that happen to lead to a certain outcome. One cannot identify a means except by identifying an end. In the case of directively organized systems, we might say that means and ends are
identified simultaneously, when, upon realizing that various states and circumstances of a given system all yield the same outcome, we identify this outcome as a goal of the system and the various paths to it as means. From an Aristotelian point of view, finally, this sort of “means-ends nexus,” and the manner in which we identify it, are particular instances of καθ’ αὑτό causal relations and their discovery. The important thing, that is, in analyzing a natural process or object, is to identify the outcome (if any) to which a process or object of this sort, precisely as such, is directed; and this outcome we call an end. In identifying an end, however, we also identify as means those steps in the process, if any, at which the process or object arrives, or to which it is directed, insofar as it is on its way to or directed to the end.

As an excellent example of this relation between ends and means, we can consider a system $S$ that exhibits persistence in Nagel’s sense. Once again let $G$ be the putative goal or end, and in this case let $S$ have three $G$-states, $M_1$, $M_2$, and $M_3$. According to Nagel, it is the existence of these alternative states, and the tendency of $S$ to assume one or the other of them at any given time, that qualifies $S$ as directively organized with respect to $G$. But why, exactly, does this fact qualify $S$ as what we might call a $G$-system? The answer is clear: $S$, as such, is not primarily organized so as to maintain itself in $M_1$, $M_2$, or $M_3$. This is obvious from
the fact that only in certain circumstances does it tend toward each of these states. Thus, at most, we might say that in circumstances $A$, $S$ is an $M_1$-system, in $B$ it is an $M_2$-system, and in $C$ an $M_3$-system. In all of these circumstances, however, $S$ is a $G$-system, and so it is much more accurate to say that $S$, as a whole, is directly organized with respect to $G$ than to say that it is directly organized with respect to, say, $M_3$. Conversely, however, $S$ is a $G$-system only because it is, in various circumstances, an $M_1$-system, an $M_2$-system, and an $M_3$-system, and because each of states $M_1$, $M_2$, and $M_3$, in $S$, tends to bring about $G$. Each of these states is therefore, in $S$, a means to $G$.

Returning now to the case of gravitation, I want to suggest, striking as it may seem, that we find here an analogous means-ends structure. In both cases, that is, we understand the behavior of the object or system by considering the various ways in which a particular state or activity, the end, can be realized or instantiated. In the case of a falling body, quite simply, the end is expressed by the general law that objects in spacetime tend to follow; the means is the concrete behavior through which a particular body, in a region of spacetime with a particular shape, actually instantiates the abstract behavior described by the law. Now although this is an unfamiliar way of describing the behavior of a falling body, there are at least many respectable means-ends relations in which the
means is an instantiation of the end. In the rest of this section, I shall both illus-
trate this sort of relation, and argue that we can understand basic causal laws in
terms of it, in two stages. The first involves reflecting on the relation between a
single law and the phenomena it governs; the second extends this reflection to
cases that involve more than one causal influence. In doing so, it addresses the
second, and the last to be discussed, of the three problems with which we began:
the fact that objects governed by a general law often only approximate the activ-
ity that this law attributes to them.

Let us begin, then, with some obvious cases in which the relation between
means and end is one of instantiation.23 First, consider a person who owns a
complete set of Mozart’s symphonies on CD, and on a particular evening wants
to listen to one. Going to the relevant part of her collection, she takes out the
nearest CD and puts it on. For our purposes, this event has the following impor-
tant features: (a) the agent’s end is to listen to Mozart; (b) the means, determined
by circumstances, is to listen to a particular recording of the “Jupiter” symphony;
and (c) the latter is an instantiation or realization of the former. Next, for an ex-
ample of nonconscious activity, imagine an overengineered electric sign

23 In fact, the distinction between what are sometimes called constitutive and causal
means to an end should be familiar from recent discussions of Aristotle’s ethics. One of the most
difficult interpretative questions raised by the Nicomachean Ethics is whether goods other than
contemplation are partly constitutive of the ultimate end, or merely prerequisite for the distinct
end of contemplation.
designed to measure and analyze the changing ambient light, and then to display a message in the color that will be seen most easily. The sign is thus directly organized to emit light that bears a specified relation to the ambient light. This activity, however, is achieved—in this case again, instantiated—by the emission of different colors at different times. Once again, we have a means-end relation consisting on the one hand of a constant end, and on the other of a set of means, the selection of which depends on which of them instantiates that end in the actual circumstances.

In the case of gravitation, although there is no distinct mechanism for selecting the means, the basic causal structure is the same. This time the object in question, whatever its circumstances, tends to instantiate by its behavior a certain abstract form of motion. What exactly this instantiation involves, however, differs depending on the circumstances. The acceleration of a falling rock, the orbit of a satellite, and the rotation of a double star are all ways in which the same general form of behavior can be instantiated. If we still hesitate to replace “ways” with “means” in these cases, it is, I would suggest, only because the latter word connotes a distinct process of selection that we do not find in them. Yet at bottom, a “means” is what stands in the middle, between an end to be attained and that which is to attain it. An abstract law such as that which governs
gravitational motion, moreover, is only explanatory insofar as we are willing to think of specific behaviors precisely as instantiations of that law, and thus as something to be understood in terms of it. By thus granting the law explanatory priority with respect to the specific behavior at hand, we identify the abstract form of behavior that the law describes as that which is primarily to be attained or instantiated. We thus simultaneously identify the specific observed motion as the means in and through which this attainment or instantiation takes place.

The second consideration in favor of a teleological interpretation of basic laws extends this same line of thought a bit further; it also addresses the second of the three difficulties involved in our initial description of a body accelerating toward the earth at 9.8 meters per second per second. That is, it involves cases in which the actual behavior of a moving object is partly determined by other causal influences, rather than by gravitation alone. For example, we might considering the behavior of charged bodies, thus bringing electromagnetism into the picture. Having begun with the homely example of a falling body on earth, however, the easiest adaptation is simply to add the resistance such bodies encounter from the earth’s atmosphere. We need not consider how this resistance is ultimately to be understood: the important thing for present purposes is simply to have another causal influence, which we can then consider from the point of
view of gravitation. Now in terms of our current theory, resistance from a me-
dium deflects a moving body from the geodesic it would otherwise follow. The
body undergoes a single concrete motion, but the causal analysis of this motion
includes both the moving body’s gravitational relation to the earth, and the resis-
tance it encounters from the atmosphere.

In reflecting on this example, the first thing to point out is that we do not
have a different law of gravitation for cases in which a moving body encounters
resistance than that which also describes bodies moving in a vacuum. We can
apply this law in cases that involve resistance because, at the level of calculation,
we can add variables to the relevant equations of general relativity in order to
represent a distinct causal influence: resistance from a medium. Such calcula-
tions, however, presuppose that we already have an abstract description of
gravitational motion in itself: that we can describe the motion of objects in space-
time that are not subject to other causal influences, or from which, conceptually
at least, we have removed such influences.

Basic mechanical laws, therefore, are formulated by abstracting particular
causal influences from the complex situations in which actual bodies are gener-
ally found, so as to describe and measure these influences separately. To explain
the actual motion of a body subject to several distinct influences, consequently,
involves applying laws whose formulation depends in large part on the disen-
tanglement of these influences, and which are therefore applicable in the first
place to those cases in which a single causal factor is translated into a particular,
concrete motion. In other words, actual, concrete phenomena must often be ex-
plained in terms of causal influences—or, more helpfully here, causal disposi-
tions—that may not be fully reflected in the phenomena themselves. The actual
behavior of physical objects must be understood as the result of tendencies to-
ward quantifiable modes of behavior, having the ontological status of goals or
ends, that enter into the structure of actual events without themselves being fully
realized or instantiated.

On this point, another analogy with more complex forms of behavior may
be helpful. Both in human life and in engineering, we often simultaneously pur-
sue distinct values whose simultaneous maximization is impossible. Someone
may wish, for example, to be both an excellent musician and an excellent football
player; but it is impossible, generally speaking, to achieve both, and so a com-
promise must be sought. Or, an engineer designing an airplane may wish both to
maximize the time it can stay aloft without refueling, and to minimize its weight.
Again, the two values are in conflict, and a compromise is necessary.24

24 With regard to this second example, it is worth noting that with contemporary comput-
ing techniques, automated decision processes can be layered so that a program itself selects the
Earlier, we observed a structural similarity between relatively simple and relatively complex phenomena with regard to the instantiation of ends by various means. The two examples just given, together with that of an object falling through the earth’s atmosphere, serve to extend this similarity. As in the case of listening to Mozart, displaying easily readable text, and moving precisely along a geodesic, so here the only significant difference between human decision-making and optimization problems, on the one hand, and the outcome of distinct causal influences on the motion of particular object, on the other, is the simplicity of the latter. In both cases, we have distinct tendencies toward outcomes that cannot all be fully realized at once; and the result is a concrete choice, or solution, or motion, that partly reflects each of them.

In the arguments of this section, I hope to have provided at least an initial motivation for the claim that basic mechanical phenomena, or phenomena governed by basic mechanical laws, have a teleological dimension. With respect to the analysis of scientific explanation, correspondingly, I have argued that a teleological notion of explanation helps us understand the explanation of concrete phenomena in terms of general, abstract laws. In conclusion, it may be helpful to integrate these suggestions with the summary of Aristotle’s teleology found at values to optimize in a given situation, thus removing the optimization process further from human intentionality.
the beginning of chapter four, and with my remarks about motion as complete activity earlier in this section.

Aristotle’s teleology, we have seen, has its metaphysical basis in the distinction between potentiality and actuality. A causal account of concrete natural phenomena will, of course, refer to specific potentialities, which are variously actualized in likewise specific circumstances. To identify and describe a potentiality, however, necessarily involves knowledge of the corresponding actuality, and it is this fact that draws us into a teleological understanding of nature. For natural phenomena, if they are to be understood causally, must be understood in terms of an inner dynamism and directionality inhering in natural objects. Ultimately, therefore, natural phenomena must be understood in terms of certain basic forms of being and activity toward which objects of various kinds naturally tend, and which therefore have the status of ends. It is in terms of such basic forms of being and activity, whatever they turn out to be, that the web of nature must be understood. The primary phenomenon that we have considered in this chapter, finally, is a particular form of motion in spacetime, which corresponds to Aristotle’s concept of complete activity and is therefore itself an end.

As we have seen, the basic forms of being and activity to which we refer in scientific explanations are realized in a variety of circumstances. Both from a
theoretical and from an empirical point of view, moreover, it is clear that the same form of being or mode of activity is not realized or instantiated in exactly the same way in all circumstances. What, precisely, is involved in this instantiation will depend on a variety of factors, the nature of which depends on the nature of the end. All this is just as true in the most basic phenomena of mechanics as it is in human life, in engineering, and of course in biology, an area in which examples could easily be multiplied. Finally, once we recognize that basic forms of being and activity only occur in nature through more concrete forms of instantiation, we are able to see in the basic structure of natural phenomena a distinction between ends and means—or, if the term “means” remains too difficult, between ends and the “ways” in which these ends are instantiated.

Throughout the foregoing discussion, I have for obvious reasons been stressing the similarities among various sorts of directed activity in nature. Nevertheless, there are clearly important differences between a robust, directly organized system such as Nagel describes, for example, and the relatively simple phenomena that we have considered at some length in the present section. In the following section, therefore, after dealing with several further objections to a generally teleological approach to nature, I shall discuss a number of different
ways in which the sort of goal-directedness we have been exploring might appear, and in fact actually seems to appear, in nature.

3. Different sorts of goal-directedness

There is no doubt that the notion of a general teleology flies in the face of most current approaches to nature. What I have been suggesting, however, is not that a more conventional view of phenomena like gravitation must be rejected, but that such phenomena can be understood more deeply than they now are, and that adequate reflection reveals a teleological dimension. As I suggested in chapter four, therefore, in trying to understand natural phenomena we do not face an alternative between mechanism and teleology, unless mechanics is understood \textit{a priori} to exclude any teleological dimension.\textsuperscript{25}

At the same time, we might well ask what exactly is to be gained by moving toward a teleological view of natural phenomena in general. Perhaps, for example, it would be more helpful to restrict teleological language to those objects or systems whose goal-directedness is somehow distinctive. Thus one might

\textsuperscript{25} As I suggested in the appendix to chapter four, our options may well be not teleology and mechanics, but rather on the one hand a realistic, causal interpretation of physical theories, and on the other, any number of views that deny to the natural world the substantiality, so to speak, that common sense and scientific practice grant it: a substantiality that consists, mainly, in a fair measure of causal autonomy or integrity. But the whole question of teleology presupposes that we want a causal account of natural phenomena; and once we board the metaphysical bus, we should be prepared to go where it takes us.
interpret the systems view not as denying the similarities between directly organized systems and less robustly “directed” systems or objects, but rather as restricting teleological language to a class of systems that, because of their peculiar features, ought to be marked off as “particularly” directed. Objections of this sort are helpful because they challenge us to state more clearly why a general teleological approach to nature is worth pursuing.

The most obvious response to such a challenge is that in fact, as I have already argued, it is simply misleading to say that the distinctive feature of a robustly directed system is goal-directedness. What is distinctive about persistent or plastic behavior, or about other processes one might consider for special status, is the mode of directedness, rather than the fact itself. Even further, there is an obvious way in which simpler modes of directedness are actually more “robust,” more constant, than those proper to certain relatively complex systems. Thus although a directly organized system may have greater autonomy from its environment in the way it tends toward specific goals, this autonomy is purchased at the cost of a relatively high chance of breakdown. Any object in spacetime, by contrast, can always be counted on at least to approximate the modes of behavior that general relativity describes. Each in its own way, therefore, these two sorts of directedness are both quite remarkable.
The basic advantage, therefore, of a broader use of teleological concepts and terms, is simply that it allows us to treat similar phenomena similarly. One explanatory context in which this is particularly helpful is the relation between a directly organized system—whether in Nagel’s or in any other plausible sense—and its parts. As we saw in chapter four, the specific directedness of such a system is partly parasitic on the more basic directedness of the parts: the system as such does not provide all the directedness it requires. Its existence, rather, depends on the presence of just the right sort of parts, and the basic directedness of these objects is not dependent on their being parts of the system. Because we thus have two levels and degrees of directedness in a single system, and especially because one of these presupposes and builds on the other, there is something odd about not giving them a uniform treatment.

There are, of course, obvious differences among the various sorts of goal-directedness we have mentioned in the course of this chapter. Before moving on to discuss some of these differences, however, we need to consider another, fairly significant obstacle to recognizing a teleological structure in relatively simple phenomena. This obstacle is simply the foundational character of the laws of physics and chemistry in our understanding of nature as a whole. Many of the laws and theories of physics, for example, are so generally applicable, and
explain so much, that they tend to acquire in our perception an unwarranted aura of necessity. It is relatively easy to see, against the fixed background of physics and chemistry, that biological systems and human artifacts are not necessarily the way they are. Thus we more easily recognize their specific organization and behavior—that they are and act in one way as opposed to another—as something that merits our attention. Directedness stands out against a background of unchanging patterns that nevertheless involve a relative openness to further, contingent determination.

Despite our tendency, however, to take the relatively universal behavior of physical and chemical objects for granted, centuries of scientific experience give us no reason to think that basic natural phenomena, whether in biology or in physics and chemistry, are necessary. Despite the attempts of some of our best minds to avoid the effort, we still have to discover the basic structure of the physical world by observation and experiment, by trial and error. Thus when we do manage to discover certain basic structures and forms of activity, in terms of which the tangled skein of nature can be at least partly understood, what we discover appears to us, at least, not as necessary but simply as given. This givenness is in fact the basic problem of teleology, viewed from another perspective: there are certain structures and forms of activity that are simply given in the world as
we know it, and the tendencies toward these structures and activities must simply be accepted.

In short, it seems to me that there are good reasons to consider goal-directedness a relatively general feature of the natural world. Although this feature appears in various specific forms, it may be helpful to state explicitly what these forms seem to have in common. In chapter four, I stated that a system is goal-directed if, in virtue of its structure, it possesses a typical mode of behavior that is distinctive of the system as a whole. In view of the foregoing discussion of basic laws, I think we can make the broader and in fact bolder generalization: a phenomenon is goal-directed precisely to the extent that it is properly understood in terms of an as yet unrealized state toward which the entity or entities involved naturally tend. With this general statement in hand, in the remainder of this chapter I shall discuss the dimensions of an important qualifier that it includes, namely, the phrase “to the extent that.”

A goal-directed process, thus defined, can be qualified in many ways. It can be relatively robust, as in systems whose organization enables them reliably to attain a specific end in a variety of circumstances and despite internal

26 The phrase “to the extent that” allows for the degrees of goal-directedness that we found while discussing goal-directed systems in chapter four. In general, the behavior of a strongly integrated system cannot be decomposed without loss of content into the behavior of its parts, and such a system must therefore be understood in terms of its own typical behavior. In a weakly integrated system, the opposite is true.
perturbations or failures. It can be relatively weak, as in objects that tend toward a particular end only in highly specific circumstances, or whose tendency toward the end is easily obstructed. It can involve relatively more or relatively less internal control, depending on the nature of the means by which the end is attained, and the way in which these means are determined. All these qualifications, however, presuppose the basic structure of tendency toward an end, as I have just defined it. Without trying to provide an exhaustive list, let us consider more closely some of the possible variations.

Given that an object \( x \) tends toward a state or activity \( G \), and assuming evidence of a real causal connection between \( x \) as such and \( G \) as such, it seems that we can ask at least the following questions concerning \( x \)'s tendency toward \( G \). (a) Does \( x \) tend toward \( G \) absolutely, or only in particular circumstances? (b) Is \( G \) fairly determinate, or does it involve a range of possibilities? (c) Does the tendency toward \( G \) depend on distinct internal activities on \( x \)'s part, or does \( G \) simply happen? (d) Does the tendency depend on the apprehension of \( G \) by \( x \)? (e) If so, does it follow automatically from this apprehension, in appropriate circumstances? (f) Is the tendency toward \( G \) always successful? (g) Does \( x \)'s route to \( G \) depend on external circumstances? (h) Does it depend on the internal state of \( x \)? In the following pages I shall comment briefly on each of these questions.
(a) Does $x$ tend toward $G$ absolutely, or only in particular circumstances?

The more $x$’s tendency toward $G$ depends on circumstances, the more $x$ is directed toward $G$ externally, rather than being intrinsically directed to it. If $x$ tends toward $G$ only in circumstances $C$, however, we must also ask whether the tendency toward $G$ in $C$ can be explained in terms of a more general tendency—say, a tendency toward $H$ in $D$. If so, the latter tendency is more characteristic of $x$—that is, it is more internal or intrinsic—than the former. Thus an animal may tend to gnaw off its leg in particular circumstances, say if it is caught in a trap, but we do not therefore describe it as a natural self-mutilator. Rather, we must analyze this tendency in terms of, perhaps, a natural tendency toward food, which results in the more particular behavior as a means to the freedom of movement required to obtain food. Several other examples, including the tendency of objects in spacetime to move toward the center of the earth, were discussed in section one.

(b) How determinate is $G$? Is it a single, fully defined outcome? A continuous range of outcomes the selection of which depends on circumstances? A disjunction of two or more discrete outcomes? If $G$ is not fully determinate, can we revise our description of the relevant circumstances so that it becomes so? If not, the directedness is obviously weaker than if a fully determinate outcome could be specified. With regard to questions such as these we might consider two
distinct cases: a system whose unity is relatively weak, so that its tendencies can be described only with a certain vagueness; and a system that simply has no fully determinate behavior, but nevertheless has a partially determinate behavior.

In the first case, for example, we may have a large ensemble of relatively independent objects, either all alike, as in a volume of gas, or of various sorts, as in an ecosystem. Such an ensemble will have determinate tendencies of its own only insofar as it has at least a minimal unity—insofar, that is, as its existence is something over and above that of its components. To the extent that this is not the case, any tendencies we observe will be wholly due to some isolable component or to circumstances. The areas of science that study such loosely unified systems involve special problems, but with statistical methods they can be more or less successful, and so we must attribute to systems of this sort at least a minimal directedness.

In the second case, we may have a relatively unified object or system whose behavior is, irreducibly, not fully determinate. This seems to be the case, at least, with the objects of quantum mechanics: from a given state of a particular system, one of several outcomes may occur. Thus, we cannot say that the system as such has any one of these outcomes, strictly speaking, as an end. On the other hand, the system’s behavior is partly determined, and we might say that the
system’s goal, in the state and circumstances specified (or specified as fully as the nature of the system allows), is the disjunction of the possible outcomes. If this seems to be a rather trivial observation, then we are precisely on track: the system we are now studying is, in another way, only minimally determined or directed in its behavior. Thus, we should expect its conformity to our general account of goal-directedness to be minimal as well.

(c) To the extent that \( x' \)’s tendency toward \( G \) is intrinsic—that is, independent of circumstances, we must also ask whether this tendency depends on distinct internal activities on \( x' \)’s part, or whether \( G \) simply happens, without further ado. This is a question that Aristotle clearly has in mind when he observes that the locomotion of inanimate bodies is rather something that happens to them than something they do (Phys. VIII.4, 254b33–255b31). We can make the same distinction: the gravitational motion we considered before belongs properly to objects in spacetime, but they are not “responsible” for it in any real sense. When a phototropic plant grows toward the sun, by contrast, this is not something that just happens: it depends on a series of internal causal processes whereby the structure of the plant determines its own behavior. The difference is even more marked in the activity that follows when a hungry animal finds
something to eat: the obvious goal is attained only by a series of perceptual, neural, and muscular processes, all with a fairly strong degree of internal control.

(d) For the sake of completeness, although to comment on them would lead us too far afield, we should include the following two questions. First, if \( x \)'s tendency toward \( G \) is relatively active, does it also depend on the apprehension of \( G \) (or of some aspect of \( G \)) by \( x \)? Here we have, of course, an important distinction between the last two examples given.

(e) Second, if \( x \)'s tendency toward \( G \) is conscious, does it follow immediately from the appropriate circumstances \( C \), including \( x \)'s apprehension of the relevant aspects of \( G \), or does it follow only from some internal activity on \( x \)'s part that is not causally correlated with \( C \)? If the latter, then the tendency not only involves cognition; it is free.

(f) Does \( x \) in \( C \) always arrive at \( G \), or does the tendency sometimes fail to achieve its end? In the case of failure, can we identify in each particular instance a factor \( y \), causally unrelated to \( x \) and \( C \) as such, that accounts for the failure? If so, the failure results from a chance obstacle. Clearly, both mechanical and organic processes or tendencies can be impeded in this way. Because, however, inanimate objects often have multiple, equally natural but simultaneously unrealizable tendencies, we do not normally think of such impediments, in their cases,
as failure. Nevertheless, if we focus on a particular tendency such as that to move
along a geodesic, we can easily distinguish success from failure—allowing, of
course, that both are rather trivial as successes and failures go.

(g) Does $x$ always tend toward $G$ by the same route, or does attributing to
$x$ a tendency toward $G$ in $C$ require a further specification of either $H$ or $I$ as the
means to $G$, depending on further external circumstances $D$ or $E$? If the latter, the
tendency is plastic. A particularly striking example is that of a certain species of
insect, whose members, in order to preserve their lives, will either move about
imbibing nutrition or, in the absence of sufficient moisture, secrete a substance
that coats their internal parts and enables them to dry up completely, and which
they also use for temporary nutrition as they revive, when moisture returns to
the environment.

(h) Finally, does $x$ always tend to $G$ by same route, or does attributing to $x$
a tendency to $G$ also require a further specification of $H$ or $I$ as means to $G$, de-
pending on whether $x$ is in internal state (not external circumstances) $D$ or $E$? If
the latter, the tendency is persistent. Nagel provides a nice example: the human
body uses several mechanisms to maintain the percentage of water in the blood
at ninety.
The foregoing eight questions express at least some of the ways in which goal-directedness seems to be found in nature, and I have also tried to present them in more or less logical order. The basic form of each question, namely, “How exactly does $x$ tend toward $G$?” highlights their common presupposition that $x$ does, in fact, tend toward $G$. And this tendency, I have already suggested, is the essence of goal-directedness.

4. Conclusion

In exploring the teleological dimensions that a broadly Aristotelian approach reveals in the natural world as a whole, the biggest obstacle is not, I think, in making the arguments themselves. It lies, rather, in overcoming the habitual assumption, so deeply ingrained in most of us after three hundred years of repetition, that teleology constitutes an alternative way of understanding nature, one that is incompatible with modern mechanics, and which therefore is useful, at best, only in dealing with a limited number of special cases. My suggestion is that on the contrary, mechanical occurrences, and thus quite plausibly all properly natural phenomena, have a teleological dimension, the recognition of which complements our current understanding of causality in nature quite nicely.27 To

---

27 It is worth stressing that the metaphysical “problem” of teleology—the explanatory priority of what comes later in time over what comes before—has turned out to be identical to the
conclude this chapter, I want to reiterate briefly but systematically the motives
for adopting an understanding of nature that includes teleology.

Basically, a teleological approach to natural phenomena clarifies the
causal structure of these phenomena by revealing the explanatory primacy of
certain basic structures and modes of activity, in terms of which we must analyze
the concrete objects of our experience and their behavior. In this clarification of
causal structure we can distinguish the following steps.

We identify objects and systems, considered as wholes, in terms of
their regular behavior.

We notice that a thing has regular behavior of its own to the extent that
it has its own structure or form, distinguishing it from a merely
random collection of parts.

We recognize that natural processes and tendencies, insofar as they be-
long to a unified whole and cannot, therefore, simply be reduced to
more basic processes or tendencies, must be understood in terms of
their typical or καθ’ αὑτό outcomes: their ends.

We overcome the temptation, in explaining natural phenomena, to lose
sight of the explanandum by attributing it to materials, agents and
processes that are not in themselves directed toward it, but are at
best directed toward it by a further structure or form to which they
are subordinate.

We recognize the common causal structure behind the various ways
that natural objects and systems can tend toward particular ends.

---

problem of causal laws, which I mentioned toward the beginning of section one: the behavior in
terms of which we understand a goal-directed object is already specified in the corresponding
law.
Moreover, it is doubtful that this additional clarity concerning the causal structure of natural phenomena can be achieved without teleological concepts and language. In trying to understand the world around us, we cannot get around the basic and pervasive fact of directedness in nature. Direction, however, necessarily involves goals, whether the external goals of Aristotle’s κινήσεις, or the self-contained activities he calls ἐνέργειαι. As we have just seen, different sorts of bodies and systems seek and attain their natural goals in particular ways. Whatever has no goal, however, is not itself anything in particular: it is a random collection with no unified existence of its own, and thus no typical modes of activity.

Finally, downplaying teleological language, or restricting it to the most obvious or robust cases of goal-directedness, inevitably leads to problems of consistency. If we try to do away with it altogether, we create a gap between our philosophical understanding and modes of expression, on the one hand, and the actual practice of many of the sciences, on the other. Needless to say, it is not the sciences that need to change: their language fits the descriptive and explanatory needs of their subject matter. Moreover, it is precisely the persistence of teleological concepts and explanations in the sciences that has led contemporary
philosophers, despite the mechanistic view of nature that many of them accept, to take such concepts and explanations seriously.

If, on the other hand, we restrict teleological language to the most robust and the most active sorts of goal-seeking, we make these modes of activity more anomalous, more unique in nature than they really are. We create a sharp terminological distinction, underlying which is a complex and relatively continuous field of variations. Not only do we thus lose sight of the similarities among various sorts of directedness, however; we also fail to notice the many other interesting ways in which they are related. The important differences among them, for example, cannot be properly recognized except against the background of a common basic structure. There are also interesting causal relations, including the layering of directedness that we have already noted in organized systems.

In short, there are a number of good reasons for making explicit the fact that in natural science, we are constantly asking not only what things are made of and where they come from, nor even what sorts of object there are. Inseparable from these questions is another, namely, what are the basic sorts of things that happen? Basic happenings, however, the kind we can use to make sense of others activities and processes, are always directed, and direction can only be understood in terms of ends or goals.
CONCLUSION

Some of my hopes for the project developed in these pages were sketched in the introduction. It would be presumptuous to rephrase them as predictions, and so they can stand as they are. These last pages are little more than a review; but they are written with the additional hope that despite the possible objections and the loose ends—of which I know many, and of which readers will find more—I have sketched a picture of Aristotle’s teleology and its contemporary interest that is coherent, and that seems worth pursuing.

Aristotle’s teleology, we have seen, is an essential part of his account of natural substances and their coming to be. In *Generation and Corruption* II.9, he places the cause “for the sake of which” side by side with matter and agent as a principle of all coming to be alike. In *Meteorology* IV.12, he explains that every natural substance is a composite of matter and form, and that form must be understood in terms of function, or ἔργον. Throughout *Generation and Corruption*, moreover, he treats both the elements and the homogeneous bodies as natural substances, not withstanding the obvious differences between their forms and the forms of living things. In short, Aristotle’s τέλοι come in many forms,
throughout his cosmos, and his teleology is a broad and flexible mode of expla-
nation that reveals the causal structure of living and nonliving substances alike.

The metaphysical roots of this teleology lie in the related concepts of actu-
ality, limit, and goodness. In every natural substance matter is for the sake of
form, or first actuality, and the composite itself is for the sake of its activity, or
second actuality. Form and activity, therefore—above all the latter—are ontologi-
cally basic, and are the keys to the understanding of nature. Each serves as a
limit, with form fixing, and activity filling out, the being of each thing. Each is a
good, enriching its subject—whether matter or natural substance—as the actual-
ity of what can exist more fully than it does in itself alone. Each, in short, is an
end and that for the sake of which.

Actuality is also the key to Aristotle’s analysis of change. Whether a sim-
ple motion in quantity, quality, or place, or the complex coming to be of an ani-
mal or plant, change is an incomplete actuality, and as such can be understood
only in relation to its end. The end, to which the change proceeds, keeps it from
being something entirely indefinite; it anchors becoming to being, and thus
makes it intelligible. The end is a complete actuality, a state of being that comes
to be by degrees. As an actuality, it is also a good: of its subject, of its agent, or of
both. All this, however, applies only to unified, continuous changes, those
cumulative or constructive motions by which an actuality comes to be, step by step, in the presence of a καθ’ αὑτό cause.

In the contemporary arena, concepts akin to Aristotle’s turn up in unlikely places. The systems view of goal-directedness was developed to naturalize—indeed, in many cases to mechanize—the goal-directedness of certain systems, and in so doing to distinguish these systems from those others, the majority, whose behavior is not teleological. All systems, however, have structures, and the behavior of a structured system is different from that of its unstructured parts. It is the extent of this separation, I have argued, rather than specific criteria such as persistence or plasticity, that determines the goal-directedness of a given system. The behavior a system tends toward in virtue of its structure is not a random outcome, but one to which it is directed by that structure: one at which it aims. This outcome, in Aristotelian terms, is an end, and the system it characterizes is goal-directed. It may be more or less goal-directed—perhaps trivially so—but to the extent that its behavior flows from its structure, that behavior has the character of an end.

Basic laws, I have argued, confer a similar causal structure on the objects they govern. The general theory of relativity, for example, tells us that objects in spacetime tend to move along geodesics. Although we cannot observe or deduce
a necessary connection between spatiotemporal objects and motion along geodesics, we do not consider this behavior an accidental feature of the objects that exhibit it. It is, rather, a basic feature of the world as our best theories present it to us, a feature that we take to explain, to account for, perhaps even to cause, many of the concrete phenomena we observe. In Aristotelian terms, therefore, motion along geodesics is a natural activity of objects in space time: it pertains to them, in other words, καθ᾽ αὑτό. As an activity toward which its subjects tend in themselves, καθ᾽ αὑτό, it is an end. Turning to the relation between this end, which is an abstract form of activity, and its concrete instances, it is plausible to understand this relation—as we do many other cases of instantiation—as connecting an end with the various means by which it can be attained. Motion along a geodesic is not, in other words, a form of activity that happens to be instantiated; it is an activity that objects in spacetime tend to instantiate, and their concrete behavior is the means by which they do so.

Finally: This teleological way of speaking about mechanical systems is certainly unfamiliar, and there are deeply rooted assumptions that tell against taking it seriously. Perhaps, however, the difficulties teleology raises are its biggest strength. For it reveals, more starkly perhaps than other ways of speaking, a feature of the world that requires our attention. Natural objects do not act
randomly. Their behavior depends on their structure, and structure is more than the lawlike behavior of separate parts. Nor is it merely a geometrical arrangement of parts: it is the integration of a dynamic system, whose behavior it channels and directs. Structure does, of course, depend on the things structured, and on their lawlike behavior. The purpose of scientific laws, however, is to specify proper activities; and in many cases, such as that of motion along geodesics, there is not even a question of reducing these activities to something more basic.

In short, a teleological approach reveals that in the world science shows us, certain forms of activity are simply given. To associate a specific process or activity with a well-defined class of objects by means of a law, for example, does not explain that motion or activity. Rather, unless it is wholly reducible to simpler phenomena, it is the motion or activity through which we understand the object—by which we give our concept of it a more determinate content, and explain why its concrete behavior in various circumstances. Conceptually, therefore—and ontologically as well, to the extent that our concepts reflect the structure of the world—it is activity, not matter or efficient causes or even structure, that is basic. This point, which Aristotle generalized when he argued that actuality is prior to potentiality, is surely worth our attention.
APPENDIX TO CHAPTER ONE

NATURAL SUBSTANCES IN METAPHYSICS Z–H

There is an important objection to the arguments and conclusion of this chapter that requires separate treatment, due to its implications for my interpretation of Aristotle’s natural philosophy as a whole. It goes as follows: Aristotle’s teleology is indeed central to his natural philosophy and his metaphysics. It is central, however, not because it applies throughout his natural science, but because it applies to those bodies that Aristotle is, in the last analysis, willing to consider substances. But, as we see above all in *Metaphysics Z* and *H*, Aristotle in fact thinks that the only true natural substances are living things. Otherwise put, only the integrated functional capacities of a self-maintaining, form-replicating being count as an Aristotelian ὄσια, in the key sense of substance-as-actuality that he develops in these books.¹

The view that the only Aristotelian natural substances are living things provides a strong reason to interpret the texts we have been considering quite

¹ This objection comes from James Lennox, to whom I also owe the terms in which I have just formulated it.
differently than I have proposed. Where the teleological language is not entirely clear, the interpreter will tend to downplay its significance; where it is clear, the tendency will be to look for a biological reference. If neither of these moves is possible, one can still conclude that Aristotle is using teleological language in some sort of extended sense. Each of these strategies, moreover, makes perfect sense as long as the major premise—that for Aristotle only living things count as natural substances in the strict sense—is in place. For this reason, although an extended discussion of Aristotle’s concept of substance is impossible here, I want at least to address briefly the claim that this premise can be found in *Metaphysics* Z and H.

There is no question that an important task of *Metaphysics* Z–H is to determine which of the items commonly considered substances in fact deserve this title. Thus in Z.2, Aristotle introduces his discussion of perceptible substances as follows:

Substance seems to belong most clearly to bodies, and so we say that animals and plants and their parts are substances, and the natural bodies, such as fire and water and earth and every other such thing, and whatever are parts of these or are [composed] of these, whether parts or wholes, such as the heaven and its parts, stars and moon and sun. Whether these alone are substances, or

---

2 As we shall see, however, Aristotle does not discuss which bodies are substances, and which are only accidental unities.
others as well, or some of these or others as well, or none of these but some others, we must now see (1028b8–15).

Throughout the following chapters, however, there is an important ambiguity—or perhaps better, a certain “flexibility”—in Aristotle’s use of this term οὐσία. On the one hand, the bodies listed in the passage just quoted seem to be proposed as candidates for the title “substance,” without further qualification. This absolute use of οὐσία corresponds to the immediately preceding discussion in chapter one, where Aristotle briefly contrasts beings in the category of substance with other beings. Shortly afterward, on the other hand, he introduces a different use of the word οὐσία, which occurs in phrases of the form “the substance of x.”

Thus in the opening lines of chapter three he writes: “Substance is spoken of, if not in more ways, yet in four at least: for both the what it was to be, the universal, and the genus seem to be the substance of each thing, and fourth is the substrate” (1028b33–36).

The locutions “x is a substance” and “x is the substance of y” are not unrelated. For if x is the substance of y, then it is surely a substance, and conversely, if x is a substance, then it may well be the sort of thing that can stand as the substance of something else. Nevertheless, as we shall see later, the distinction is an important one to keep in mind. For some things that Aristotle certainly considers substances—namely, natural bodies—are not in fact called substances because
they are the substance of something else. And to deny that a natural body is a substance of some further being is not to deny that it is a substance, simply speaking.

With these initial considerations in mind, let us consider two passages from book Z in which Aristotle might seem to imply that the only natural substances are living things. The first is in Z.9, where he concludes a discussion of coming to be as follows:

But from these things we can derive a proper characteristic (ἴδιον) of substance, that it is necessary that another substance pre-exist in actuality, which is the maker, (ἡ ποιεῖ), as an animal if an animal comes to be; but for a quality or a quantity this is not necessary, but it may exist only potentially (1034b16–19).

It might be thought that in this passage Aristotle picks out, as characteristic of natural substances, the reproductive capacity characteristic of living things. In fact, however, this characteristic of the coming to be of substances is not even restricted to natural substances: it applies just as well, as the previous discussion makes clear, to artifacts, whose form pre-exists in the mind of the artisan.

If we turn to the elements and the homogeneous bodies, we find that on Aristotle’s account of their coming to be, the pre-existence requirement holds in their cases as well.\textsuperscript{3} An adequate discussion of this point would require

\textsuperscript{3} As mentioned at the beginning of this chapter, Aristotle’s examples of natural substances throughout Metaphysics Z and H are generally biological. As I argue below, however, this
examining a number of texts from *Generation and Corruption*, and would take us far from the present topic. Luckily, however, there is a passage in I.7 of this work that summarizes the principle involved, referring specifically to the two primary qualities that Aristotle considers active, heat and cold. Here, early in the long series of arguments by which he approaches the reciprocal changes of the elements, we find a discussion of action and being affected. In the course of this discussion he writes: “Thus it is now reasonable [to say] that fire heats and the cold cools, and generally that the agent assimilates the patient to itself: for the agent and the patient are contraries, and coming to be is into the contrary; hence it is necessary that the patient change into the agent.” By this principle, he can explain not only incidental changes in quality, but also the transformation of the elements and the coming to be of the homogeneous bodies.

The second passage we need to consider is somewhat more difficult. In chapter 16 of book Ζ, Aristotle puts the finishing touches on the dialectical back and forth that characterizes most of this book, briefly disputing the views that the parts of animals and the elements, or being and unity, or separate Forms, are substances. He begins the chapter thus:

---

is not because the points he is making are restricted to living substances. Rather, he can focus on the most obvious examples for these points because it is not his concern, in the *Metaphysics*, to determine what the natural bodies or substances are. This has already been accomplished in his works on natural science, and his present task is to consider these same bodies insofar as their study is relevant to first philosophy.
It is clear that even of the things that seem to be substances, most are potentialities: the parts of animals (for none of these exists separated, and when they are separated, even then they all exist as matter), and earth and fire and air: for none of these is one, but like a heap, before they are concocted and something one comes to be from them (1040b5–10).

At first glance, this passage may appear to deny the elements the status of substances. However, this interpretation leaves us with a hard question: if earth, air, fire, and water are not natural substances, then what are they? They are certainly in the category of substance. And although they are different from living things—much more so than one living thing is from another—they are not artifacts, nor are they merely accidental unities, like the musical man.

Another possible interpretation of this passage—and, I think, a more plausible one—is that it involves an implicit use of the “substance of x” locution. Aristotle is not simply looking at the elements, and denying that they are substances in any sense. He is, rather, attacking the view that earth, air, fire, and water are “the substance” of the things composed of them. They are, rather, “potentialities”; considered in themselves, they lack the unity that characterizes the animal or plant, say, that is composed of them, and that we are trying to understand.

One reason to support this second reading is that it allows us to give a uniform interpretation of the two sorts of potentialities that Aristotle mentions:
the elements, and the parts of animals. The items in each class, that is, are only potentially a living body, and therefore cannot be considered the “substances” in terms of which this body must be understood. If, on the other hand, we take Aristotle to be denying that the elements considered in themselves are natural substances, then the two examples need distinct treatment. Moreover, on this reading it is not as clear in what sense the elements are potentialities. Does earth become a substance any time it is fused together by heat? If earth is only a substance by its potential incorporation into a living body, then we are back to our earlier question. What is earth, in itself? Is “what it was to be” for the elements nothing other than to be potentially something else?

A second reason to think that “substance” in the opening lines of Z.16 signifies the substance of something is that this interpretation sits well both with the rest of the chapter, and with the conclusions that Aristotle develops in the following chapter and in book H. Near the beginning of Z.17, he writes that “substance is a certain principle and cause,” and the rest of this chapter and all of the following book are concerned with οὐσία in the sense of the principle and cause of perceptible substances. In chapter 16 itself, after some further remarks about the parts of animals, he goes on to argue that neither unity, nor being, nor
separate Forms, are “the substance of things.”⁴ We might even say, therefore, that chapter 16 is the final preparatory step before Aristotle begins his definitive discussion of substance as a principle and cause: after concluding his consideration of the four original candidates for substance (the substratum, what it was to be, the universal and the genus), he briefly reviews the inadequacy of various apparent substances for the causal role he wants to isolate in the chapters that follow.

The interpretation of this passage, considered in itself and in its immediate context, is surely underdetermined. There are, however, other reasons to reject the view that natural substances in *Metaphysics* Ζ–Η include only living things. In the remainder of this appendix, I will argue that far from applying only to natural substances, Aristotle’s discussion in Ζ.17 through Η.6 does not even distinguish between substances and merely accidental unities. For the distinction between substances and non-substances among perceptible things, rather, Aristotle points us back to his natural philosophy.

At the beginning of Η.1, Aristotle distinguishes between substances that are agreed upon by all, and those proposed by particular thinkers. The generally

---

⁴ “The substance of things”: οὐσίαν εἶναι τῶν πραγμάτων (1040b18–19). This phrase occurs in connection with unity and being; immediately afterward comes the argument concerning separate Forms, in which Aristotle approves the existence of eternal substances, unseen by us, while denying that these are principles of perceptible substances.
agreed-upon substances are, he says, “the natural [substances], such as fire, earth, water, air, and the rest, the simple bodies; then plants and their parts, animals and the parts of animals, and finally the heavens and the parts of the heavens” (1042a8–11). The following discussion, he adds a few lines below, concerns these substances rather than the others. Now given the remarks he has already made about the parts of animals, we should not take this introduction as an endorsement of the status of all these items as natural substances. Nevertheless, the following discussion provides no basis for a distinction. In fact, Aristotle goes on in chapter two to consider a wide range of perceptible things, each of which exhibits the ontological structure to which he wants to call our attention: the distinction between a material element, which is only potentially the item for whose substance we are looking, and the actual differentia or differentiae by which the item in question is constituted as an actually existing unity.

In this chapter and in those that follow, none of the examples by which Aristotle illustrates this basic structure is even arguably a natural substance. In chapter two, in addition to the familiar house and other artifacts, he goes so far as to include dinner and breakfast, which differ from each other by the time of day. Moreover, Aristotle seems genuinely interested in the general pattern he is discussing, and not just in the subset of instances that consists of natural
substances. Thus at the beginning of the chapter he writes, “it remains to say what is the substance as actuality of perceptible things,” and goes on to criticize Democritus for accepting only three sorts of differentiae of the underlying matter of things. After giving a number of examples, he concludes: “It is indeed clear from these things that if substance is the cause of each thing’s being, we must seek in these [i.e. the differentiae] what is the cause of the being for each of these” (1043a2–4).

Aristotle does not, of course, actually identify any of the items he mentions, or their respective actualities, as substances. Thus he continues, immediately after the sentence just quoted, “So although none of these is substance, even coupled [with matter], nevertheless it is the analogous in each” (1043a4–5). There is, then, as we should expect, a distinction between the actuality of a substance and that of a non-substance. However, Aristotle does not go on, either here or later in the book, to explain which perceptible things count as substances. For this distinction, rather, he refers us to his works on nature.

At the end of the first chapter of book Η, Aristotle provides a brief review of the things that can be called substance in the sense of matter. In particular, he contrasts what underlies changes in place, quantity and quality, with the underlying factor in changes with respect to substance, that is, in generation and
corruption. This passage concludes with a statement that is sometimes placed at the beginning of chapter two, but clearly pertains to the preceding material as well: “What, then, the difference is between unqualified and qualified coming to be, has been stated in the writings on nature (τοῖς φυσικοῖς)” (1042b7–8). In commenting on this reference, Ross points to Physics V.1 (225a12–20) and Generation and Corruption I.2 (317a17–31). In the latter passage, rejecting the view that unqualified coming to be takes place by association and dissociation, Aristotle provides a succinct statement of his own view: “For unqualified coming to be and passing away are ... when it changes from this to that as a whole” (317a20–23). This description, in turn clearly depends on the original account of nature as form in Physics II.1.

For a general statement of Aristotle’s view on what counts as a substance in the strict sense, then, Metaphysics H refers us ultimately to the claim, in Physics II.1, that a natural substance is any body that has, in itself and as such, a principle of motion and rest (192b21–3). Moreover, there is a statement in H.3 that serves the same purpose as H.1’s reference to unqualified coming to be. Its tentative expression again reveals that in Metaphysics Z–H, Aristotle is not concerned to distinguish carefully between substances and accidental unities; this distinction,
rather, is a task for another treatise. Immediately after mentioning houses and utensils as examples of bodies whose “substance” cannot exist apart, he adds: “Perhaps neither these, nor anything else not constituted by nature, are substances at all: for someone might posit that nature alone is the substance in corruptible things” (1043b21–3). Here again, Aristotle suggests that perceptible substances are those things that exist by nature—that fit, in other words, the account of Physics II.1.

The final and most important step in my argument, which brings us from the general statements about natural substances found in the Physics and elsewhere, to the elements and homogeneous bodies in particular, is simply to point out that in Generation and Corruption, one of the conclusions that Aristotle is most concerned to establish is that the unqualified coming to be and passing away of the elements and the mixed bodies is not merely apparent. To the passage already quoted from the second chapter of this work, any number of others might be added. Particularly relevant is the lengthy discussion of mixing in I.10–11, and Aristotle’s insistence there that in mixed bodies, the defining features of elements are present only potentially. In addition to such particular features of Generation and Corruption, we have the very title of the work, which he provides in its opening line. In short, if the elements and mixed bodies are not natural substances,
then why does Aristotle insist at such length that their coming to be is unqualified coming to be, and not merely a matter of association or dissociation, or of alteration?

My main concern in this appendix has been to show that there is no reason to think, on the basis of *Metaphysics* Z–H, that for Aristotle only living bodies count as natural substances. For this reason, the preceding paragraph provides only a gesture, though I think a powerful one, toward *Generation and Corruption* as a work in which Aristotle sets out, among other things, to argue that the elements and the mixed bodies are natural substances, and not merely accidental unities. I want to conclude with a statement from another work altogether, *De caelo*, which nicely summarizes the thesis of the present chapter from the point of view we have taken in the appendix. At the end of book three, concluding his argument against the *Timaeus*’s geometric account of the elements, Aristotle writes:

That, therefore, the elements do not differ by their shapes, is clear from what has been said…. The decisive (κυριώταται) differences of bodies are those in their affections, functions and capacities (αἱ τε κατὰ τὰ πάθη καὶ τὰ ἔργα καὶ τὰς δυνάμεις)—for we say that each of the things [that exist] by nature has functions and affections and capacities (307b20–24).
APPENDIX TO CHAPTER FOUR

TELEOLOGY AND BACKWARD CAUSATION

In section one of this chapter, I mentioned the problem of “backward causation” that is associated with teleology. If, that is, we find that we need genuinely teleological explanations to make sense of the natural world, this seems to imply that what exists already can somehow depend on what does not yet exist, and may never exist at all. This is a serious metaphysical conundrum, which constrictions of time and space prevent me from discussing at any length in this dissertation. In the following appendix, however—without suggesting that I have worked out a solution to the problem of backward causation—I want to reflect briefly, at least, on its relation to Aristotle’s thought, and on some of the conceptual resources he may have for resolving it.

It is important to keep in mind that the metaphysical problem of backward causation is, from an epistemic point of view, secondary to questions about scientific explanation and the structure of the natural world. In other words, the problem only arises once we have reason to believe that
goal-directedness is in fact present in nature. We should not be surprised, moreover, if reflection on nature, and on our explanatory practices, raise metaphysical difficulties that we cannot immediately find the resources to answer. That teleology is a difficult topic does not count, without considerable further argument, as evidence that there are no goals in nature. If such goals can be discerned, rather, they simply provide further evidence that nature does not always respect our philosophical conclusions and intuitions.

As we saw in chapter three, Aristotle often treats finality and agency as correlative causal factors. This is already a step toward the demystification of final causality. After all, if an end is nothing more than what an agent tends to bring about, it is at least no more mysterious than agency. Another important consideration, to which I now want to turn, has to do precisely with the distinction between potentiality and actuality in terms of which I have stated the problem of teleology in section one of this chapter. In fact, it turns out that not only does Aristotle’s concept of potentiality speak to this conundrum of teleology; it is also an essential part of his defense, against thinkers such as Parmenides and Plato, of the reality of change and the autonomy of nature. For this reason, as we shall see, it may well turn out that the difficulty toward which the riddle of
teleology points us is not a choice between teleology and mechanism, but rather a choice between realism and antirealism in our approach to nature.

What I want to suggest, then, is that given the distinction between potentiality and actuality, the riddle of teleology that we are now considering does not arise; or if it does arise, it is much less acute that it would be without this distinction. The reason for this is that Aristotle’s potentiality, and all the more so his incomplete actuality, pertains rather to being than to nonbeing. This understanding of potentiality as a division or way of being is reflected throughout the *Metaphysics*, where in several passages the well-known division of being by the categories is followed immediately by a further division, ranging across all the categories, between being in potentiality and being in actuality. Let us consider it more closely, and examine its implications.

Potentiality and incomplete actuality pertain to being; they pertain to it in each of the categories; and so also they pertain to it in the more particular forms into which these categories are divided. To be potentially red is a way of being, as is to be potentially a man. From this it follows—and here is the crucial point—that being in potentiality is not related to being in actuality as to something wholly outside itself. Rather, whatever is potentially X is, in a manner, already X; and, conversely, to be actually X is presignified and precontained in the
potentiality for $X$. Moreover, Aristotle clearly holds that this relation between potentiality and actuality is primitive and irreducible. Metaphysically speaking, that is, “can” is an unanalyzable term: there is no more basic or more perspicuous way to express the relation between what at one time is not $X$, and at another time is $X$, all the while remaining itself.

To appreciate the significance of Aristotle’s concept of potentiality, it may help to recall briefly the difficulties that faced earlier Greek philosophers as they tried to reconcile the radical metaphysical distinction between being and nonbeing with the continual change they encountered in the sensible world. One common reaction to this dilemma was, of course, to put aside serious metaphysical reflection altogether. Those thinkers, however, who persisted in the endeavor often ended up denying the reality of change, and thus denying or at least severely undercutting the reality of nature. Without the distinction between actuality and potentiality, it seems, one could be a Parmenidean, a Platonist, or any manner of exotic metaphysician; but one could not be a realist about nature. One could not, that is, view the sensible world as a system of real objects, really acting and interacting in accord with the real possibilities and limitations of their own being.

Most immediately, the concept of potentiality seems to lie at the heart of Aristotle’s answer to the problem of change as formulated by Plato, an answer
that emerges in *Physics* I.7–9 and *Metaphysics* Ζ.7–9, H.6 and Θ.1–9. On the Platonic view—that is, according to the theory of separate Forms—whatever is not actual, in and of itself, has of itself no positive content at all. It thus has determinate, actual being only insofar as it receives this being from one or more Forms, which are external principles, of an entirely different order. Only in this way, Plato thinks, can that which is not, or is not $X$, come to be that which it is not.

Aristotle’s response to Plato consists in distinguishing relative from absolute nonbeing, on the one hand, and actuality from potentiality, on the other. An adequate account of changing, sensible objects, he argues, requires first of all a distinction between that which exists in an absolute sense, and can thus undergo change, and the particular respect in which it does not exist, or more perspicuously, the particular attribute it does not have. This distinction, however, enables us to understand change only if in some cases, what does not actually have a given attribute nevertheless possesses that same attribute in potentiality. Unknown as it may be in itself, this being in potentiality has determinate content that can be grasped by reference to the corresponding actuality. In short, Aristotle avoids Plato’s radical solution to the problem of change by pointing out that the problem has been overstated. Only absolute nonbeing truly and completely is not, and to it alone is being always completely foreign.
By distinguishing absolute from relative nonbeing, and actuality from potentiality, Aristotle resolves the Platonic tension between the nonbeing at the heart of sensible objects, and the being imposed on it by participation in the Forms. Relative nonbeing, or perhaps better, nonactuality, may also involve potential being. Potential being, however, is a division of being; it is, more precisely, a mode of being insofar as it involves, not the possibility of receiving something basically foreign to it, but rather an intrinsic dynamism and directionality in relation to the actual. With this concept of potentiality in place, Aristotle thinks, we need not conclude that the apparent being of nature in its own right is illusory: nature acquires an autonomy that for Plato, let alone Parmenides, it can never have.

With this solution to the problem of change and the corresponding problem of the ontological status of nature, it seems to me, Aristotle also provides a teleology that is truly causal, yet does not involve the mysterious action of nonexistent final causes dragging nature along by invisible cords. The causal role of ends, rather, which admittedly do not yet exist, can and must be explicated in terms of what does exist, but potentially rather than actually. To some, of course, to solve the riddle of teleology by appealing to the notion of potentiality may seem to beg the question. After all, there are prominent metaphysicians in the
past half-century who have found the notion of potentiality, as well, unacceptably mysterious. Yet this very fact indicates that the question we face is not one of teleology versus mechanism. It is, rather, as the foregoing references to Plato and Parmenides also suggest, a question of realism versus antirealism about nature.
REFERENCES


