The Four Causes

Der Fakultät für Sozialwissenschaften und Philosophie

der Universität Leipzig

eingereichte

Habilitationsschrift

zur Erlangung des akademischen Grades

„doctor philosophiae habilitatus"

(„Dr. phil. habil.")

vorgelegt von

Dr. Boris Hennig

geboren am 5. März 1974 in Trier

Leipzig, den _____

Contents

Introduction

<u>Causes and Becauses</u>

Aristotle says that in order to really understand a thing, we need to understand its αἰτία, and he distinguishes between four kinds of αἰτία. This term, αἰτία, is usually translated as "cause." However, not all of Aristotle's four αἰτίαι are causes in the modern sense of this word. Perhaps none of them are. There are no English words that are a direct translation, but if one uses "cause," some explanation should be added as to what it is supposed to mean in this context. A common way of doing so is to give an example like the following.

> Take an artifact, such as a silver cup. The material cause of the cup is the silver it
> is made of. Its formal cause is the shape into which the silver was brought when
> the cup was made. The efficient cause of the cup is the person who made it (or,
> perhaps, the capacity of making it). Its final cause is the purpose for which it was
> made, which is presumably the purpose that its maker had in mind.[1]

This way of explaining Aristotle's four causes is misleading in several respects (Sprague 1968). To begin with, it explains all of the causes by using a single example, which Aristotle never does, and this single example is in most cases an artifact, such as a silver cup, a statue, or a house. Although Aristotle refers to artifacts in many of his examples, they are not the main targets of his distinction of causes. His four causes are primarily causes of natural things, and natural things differ from artifacts in precisely the respects that are highlighted in the cup example. When natural things come into being, they are not created with a purpose in mind, and not by shaping a

---

[1] See the similar account given by Heidegger in Die Frage nach der Technik, Gesamtausgabe 7, p. 10-11.

given portion of independently identifiable matter. Rather, they grow by themselves, by taking in and exchanging matter, so that both their matter and their form undergo considerable change during their development. This makes it difficult to apply the distinctions, as drawn in the cup example, to natural things. Given that many natural things were never made out of a given portion of matter, it is more difficult to distinguish their form from their matter than it is in the case of a silver cup. Further, Aristotle writes that the formal and the final cause of a living being are the same (e.g. Physics II 7, 198a24-25). Again, it would be odd to say this of the shape and the purpose of a cup.

Vlastos suggests that we understand Aristotle's distinction of four causes better when we translate "X is the αἰτία of Y" as "Y happened, or happens, or is the case, because of X" (1969, p. 293-4). Accordingly, Hocutt refers to them as the four "becauses" and claims that for Aristotle, causes are explanations (1974, p. 388). This leads us far away from the cup example. For instance, if Aristotle's causes were explanations, silver could not be one of them. Silver is not a kind of explanation.

Some think that what is wrong with Hocutt's proposal is that Aristotle's causes are real phenomena in the world, whereas explanations and their parts are only bits of language (Mure 1975; Furley 1996, p. 60). However, if this were the only problem, it could be easily avoided. Any explanation must state facts and refer to things, and we can easily switch back and forth between explanation and their parts on the one hand, and the facts and things they refer to on the other (Johnson 2005, p. 41). Still, the silver does not cause the cup, any more than "silver" explains "cup." Likewise, many reasons against identifying the material cause of the cup with the explanation, "because it is made of silver," are also good reasons against identifying it with the fact to which this explanation refers.

The intuition that Aristotle's causes provide explanations is often expressed by saying that they correspond to answers to Why-questions (van Fraassen 1980a, p. 24; Irwin 1988, p. 94; Hübner 2001, p. 378). Indeed, Aristotle seems to introduce them in this way:

For since our undertaking is for the sake of understanding, and we do not believe that we know each thing before we can grasp the "why" [τὸ διὰ τί] of it (this is to grasp its first cause), it is clear that this must be done by us with regard to coming to be and destruction and all natural change, whence, knowing the principles of these, we may try to reduce each one of the things sought to these [principles]. (Physics II 3, 194b17-23, tr. Coughlin, modified)

Aristotle seems to say that to know the cause of something is to grasp the "why" of it. Now if Aristotle's causes were answers to Why-questions, it would be obscure why there should be exactly four kinds of them. It is of course easy to come up with a set of questions that match the cup example:

Why is this cup shiny? — Because it is made out of silver.

Why does the cup not fit in the drawer? — Because of its shape.

Why does this cup exist? — Because someone made it.

Why is this cup on the table? — Because this makes the room look nicer.

This list of questions, however, is rather ad hoc. One might easily go on asking Why-questions about the cup, and not all of these questions would clearly correspond to one of the Aristotelian causes. Why is it made out of silver? Why is it so expensive? Why don't we get another one of these? As Falcon notes, "not all why-questions are requests for an explanation that identifies a cause, let alone a cause in the particular sense envisioned by Aristotle" (2008, section 2). Taken as a classification of Why-questions, Aristotle's scheme is thus at best an "over-simplification" (van Fraassen 1980b, p. 131).

On the other hand, the four Why-questions above are not as easily kept apart as one would like, so that there might well be fewer than four general kinds of them. On a more general level, Why-questions seem to divide into two rather than four kinds:

requests for reasons for acting, and requests for causal explanations in the modern sense of "cause." For instance, a satisfactory explanation of why the cup is shiny seems to be that its surface reflects light easily, and this looks like the beginning of an explanation in terms of efficient causes. Further, such causal explanations might still be more fundamental than explanations that cite reasons for acting, so that all kinds of answers to Why-questions might ultimately reduce to one. A good explanation why the cup is on the table might be that someone had a certain desire, which (efficiently) caused that person to put it there. Seen in this way, all causes seem to boil down to efficient causes (cf. Irwin 1980, p. 96; Freeland 1991, p. 50; Furley 1996, p. 62).[2]

Thus the suggestions that causes are answers to Why-questions does not help to preserve the variety of Aristotelian causes. On a very general level, answers to Why-questions seem to divide into fewer than four kinds, and on a less general level, there seem to be more than four kinds of them. This is also the main problem with Hocutt's suggestion that causes are explanations. There are not exactly four kinds of explanation, nor are there exactly four kinds of real phenomena to which explanations refer. Consider, again, the material cause. If causes were explanations, the material cause would have to be a special kind of explanation, which would presumably explain something on the basis of facts about some matter. However, the idea that such explanations constitute their own kind is in at least as bad a position as the idea that any syllogism that refers to an action is a practical syllogism. As Anscombe writes,

> ... one might easily wonder why no one has ever pointed out the mince pie
> syllogism: the peculiarity of this would be that it was about mince pies, and an

_____

[2] There may be further kinds of answers to Why-questions that do not reduce to efficiently causal explanations, such as mathematical explanations. However, even if these should be the same as explanations in terms of formal causes (cf. Physics II 7, 198a17), the argument above still shows that it is difficult to distinguish between explanations in terms of material, efficient, and final causes.

example would be 'All mince pies have suet in them—this is a mince pie—
therefore etc.' (1957, §33).

Anscombe's point is that there are no good reasons for distinguishing kinds of
reasoning only by their subject matter. If practical reasoning is to be taken seriously as
a special kind of reasoning, it must be special in virtue of its logical form. This point
need not apply to all kinds of explanatory reasoning. One might, for instance, define
mathematical reasoning as reasoning about mathematical objects, and thus distinguish
it from other kinds of reasoning in terms of its subject matter. However, it does not
seem as appropriate to distinguish material from formal or efficient explanations by
saying that the former refer to matter, whereas the latter refer to forms or efficient
causes. Here, the subject matter is not sufficiently different. For instance, a material
explanation of why the cup is shiny is not clearly distinct from formal and
(efficiently) causal explanation of the same: It is shiny because of its matter, because
of the form of its surface, and because it reflects light. Therefore, if any explanation
that refers to matter were a material cause, one might easily wonder why no one has
ever imagined a mince pie cause. This would be an explanation that refers to mince
pies. If causes were explanations, it would be difficult to see why there should be four
of them. There seem to be no formal differences between them.

Perhaps we should not suppose, then, that Aristotle's causes are explanations or
answers to Why-questions. And indeed, Aristotle only says that they are answers to
the question διὰ τί, and this question is more general than our question "Why?" It
asks for an account (λόγος, Metaphysics A 3, 983a28), but one may also give an
account of what a thing is, or how it is structured, and this is not to explain why it
exists and why it is structured in this way. The question διὰ τί asks on account of
what, in what capacity, or in virtue of what something is such and such, and asking
"Why?" is only one way of asking this. Other ways of asking διὰ τί are: How did this

happen? Who did this? What's the point? What does it take to be this kind of thing?[3]

For instance, the questions "By virtue of what are these bricks a house?" and "Why are these bricks a house?" are not equivalent. The bricks are a house because someone arranged them in a certain way, so that they provide shelter. These are the efficient and final cause of the house. Now one might as well say that the bricks are a house in virtue of having been used to build one, or in virtue of providing shelter. But the phrase "in virtue of" also has a different sense, in which it does not answer a Why-question. For instance, that in virtue of which the bricks are a house may be taken to be the way in which they are arranged, and to say in what way the bricks are arranged is not to say why they are a house. The question to which it is an answer is not "<u>Why</u> are these bricks arranged so that they constitute a house?" but "<u>How</u> are they arranged so that they constitute a house?"

Also, to ask "Who did this?" is not the same as asking "Why was this done?" The latter question asks for an explanation, the former merely concerns the attribution of an action to an agent. When we say that Polycleitus made a statue, we do not actually say why the statue was made; all we say is <u>who</u> made it. That Polycleitus made the statue also means that one may find out why it was made by asking Polycleitus. Polycleitus is responsible, he should be able to give an answer to the question "Why?" In this sense, he may also be said to provide an explanation. Still, to say who did it is not to say why it was done.

Aristotle's four αἰτίαι are not causes in the modern sense of the word "cause," in the same way in which his question διὰ τί is not our question "Why?" In order to understand what they are, and why there are four of them, we need to understand what the question διὰ τί means, and why there are four ways of answering it. Since διὰ τί does not always mean "why," we cannot understand the four ways of answering it by investigating Why-questions and their answers. When Aristotle says that the silver is a cause of the cup (194b25), he does not say that silver is the cause for the cup's being

---

[3] Charlton translates the διὰ τί in 194b19 as "on account of what," Wicksteed (in the Loeb edition) as: "how and why."

shiny, nor does he say that the fact that silver has been shaped explains the existence of the cup. He simply relates the cup to the silver as one of its causes, and this is not the kind of explanatory relation that would hold between a fact and the explanation why it is so.

Frede points out that Greek philosophers generally distinguish between the αἴτιον of a phenomenon, which is something that is responsible for it, and its αἰτία, which is an account or explanation of why and how an αἴτιον is responsible for this phenomenon. He also notes that Aristotle does not observe this distinction (1987, p. 129-30). Had Aristotle done so, he would probably have consistently used the term αἴτιον rather than αἰτία for his causes (as he does in Physics II 3, 194b24). They are not kinds of explanations or answers to Why-questions. Rather, they correspond to four ways in which one should look at things in order to understand them, and eventually be able to answer Why-questions about them. Just as we may ask Polycleitus in order to find out why and how the statue was made, we may investigate certain aspects of natural phenomena (ask them, as it were) in order to find out why and how they come about.[4]

When Aristotle says that in order to really understand a thing we need to ask four kinds of question about it, he says that we may find answers to Why-questions about this thing by asking four questions that are not Why-questions. These other questions are questions like the following:

> What is this made of?
> What does it take to be this kind of thing?
> What made this happen?
> What is this for?

---

[4] In the sense in which Heidegger says that the question of what nature is must address the movedness of natural things ("bei der Bewegtheit dieses Seienden anfragen," Vom Wesen und Begriff der Φύσις, Gesamtausgabe 9, p. 245).

It is, of course, still not obvious why there should be exactly four kinds of What-questions, any more than it is obvious why there should be four kinds of Why-questions. All we have achieved by turning from Why-questions to What-questions is to direct our attention away from "becauses" and explanations. In doing so, we have also turned back towards the cup example and its flaws. We still do not see why there should be exactly four kinds of Aristotelian causes, and what each of them is. Showing this is the aim of the present book.

Things to Keep in Mind

I take it that the four causes are primarily causes of natural things and processes. All of the following is therefore about such things and processes, and only accidentally about artifacts and intentional actions. The reader should keep this in mind. Counterexamples that involve artifacts, mathematical entities, or intentional agents may not be relevant. For instance, I will eventually claim that the matter of a thing is something that turns into this thing as a result of the thing's natural development, and this is obviously not true for artifacts and their matter. Silver does not naturally turn into a cup, and cups do not naturally come to be out of silver. Also, I argue that the formal cause of a natural thing must be a compound thing, and this does not apply the formal causes of immaterial and mathematical entities. For instance, Frede argues that in Metaphysics ZHΘ, Aristotle claims that the primary substances are pure forms, and that material substances are substances only in a secondary sense (1987, p. 79). In a context that only involves natural things, material substances are as primary as it gets, and therefore, I can leave it open whether pure forms are even more primary.

Another thing that should be kept in mind is that I use the term "cause" only because it is the standard translation of αἰτία (and αἴτιον). As a translation of αἰτία, "cause" behaves in unusual ways. It is not associated with the usual verbs and adjectives, so that in general, there is no causation or causality associated with it.

There is no aitiation or aitiality, as it were.[5] Also, causes in the sense of αἰτίαι do not generally have effects. There are no material effects to result from material causes. This makes it look as though "cause" is really not an adequate translation at all. On the other hand, it makes it easy to distinguish the ordinary sense of "cause" from the one that the word takes on when it translates αἰτία. Whenever there is causation, causality, or an effect involved, the word "cause" can only have its plain English sense. I exploit this in order to keep the Aristotelian and the modern sense apart. By "cause" I generally mean an Aristotelian cause, unless stated otherwise; when I speak of "causation" and "causality" I generally refer to causation and causality in the modern sense of the word "cause," again unless stated otherwise. As for "cause" in its capacity as a translation of αἰτία, I pretend to know nothing about its meaning, so that this meaning may emerge from the ways in which Aristotle actually uses the word that it translates. The labels "material cause," "formal cause," etc. are not used by Aristotle. I also use them without putting any weight on their literal meaning.

Further, I use the word "type" in a somewhat technical sense. In this sense, something is an instance of a type to the extent to which certain standards of typicality apply to it; not to the extent to which it satisfies them. I will not formally introduce this sense. Rather, the word as used by me will gradually acquire it, especially in the course of Chapters 5 and 6. The sense I am giving it is in any case not far from what seems to be its literal sense; for "type" and "typical" are obviously related. That I am using "type" in a special sense is important to keep in mind when reading Chapter 10, where I argue that when something can be shown to contribute to the well-being of a living being, it may be taken to be typical for it.

The reader should further be warned that in a way, this is not a book about Aristotle. Whenever I say something about Aristotle in the following, I have made sure, to the best of my ability, that it is adequate and accurate. I am, of course, liable for possible misreadings and misinterpretations, so that if anything that Aristotle says

---

[5] Aristotle uses the term αἰτίασις once (Poetics 18, 1455b31), but there it means "accusation" (not "causation").

should contradict what I say about him, I would have to defend what I say or take it back. However, my primary aim is not to say correct things about Aristotle or to explain how he came to think what he thought. My primary aim is to say what Aristotle said, not in the sense of reporting what he said but in the sense of repeating and elaborating on it. Repeating a claim comes with more responsibility than reporting it. In general, I endorse the claims that I attribute to Aristotle, so that I am as accountable for their intelligibility and truth as I think Aristotle is. Taking this kind of vicarious responsibility is the best way of finding out whether Aristotle was right or wrong.

Some of the following is not about Aristotle in a more obvious sense. It is, ostensibly, about Hume, Kant, Austin, Anscombe, and others. If this were a book about Aristotle, these authors might play a marginal role in it, by helping modern readers to situate Aristotle's views within their own philosophical context. As it stands, they play a less than marginal role. A discussion of their views is important, and sometimes essential, for understanding the way in which I say and develop what Aristotle said.

The Introductory Chapters Introduced

A previous version of Chapter 1 was accepted for publication a while ago. I have been very conservative about the accepted version (Hennig 2009) and applied many changes only to the book version; as a result, the two versions differ in many details. In particular, I have rewritten section 4 and omitted the better part of the concluding section. Chapter 1 may be read as a summary or outline of the overall argument.

On the face of it, Chapter 2 is not at all concerned with Aristotle and his four causes. It introduces and explains Austin's distinction between two directions of fit. This distinction has been widely misunderstood, and it is important for my purposes to restate it in its original form. I find it convenient to do this by confronting Austin, who gave it its name, with Kant, who would have had good use for it.

The second chapter may be read as a reflection on method, in general as well as in

particular. It gives a rough impression of how a priori knowledge is possible and thus of the status and possibility of genuinely philosophical claims. This concerns the method of the present book, insofar as it is a philosophical work. The distinction between two directions of fit also performs more specific tasks in the following. It makes a brief appearance in Chapters 3, sections 4 and 5, in a gloss on Aristotle's distinction between τι and τοιόνδε, and in relation to Plato's Timaeus and Cratylus. Above all, it is important in Chapter 6, section 5, for explaining how essences differ from properties. I argue, generally, that essences are that in terms of which we may identify an object, without yet describing it as having certain properties. The distinction between two directions of fit explains how this is possible. Incidentally, in section 6, it will also help to address the question whether Aristotle's forms are universal or particular.

The Causes In More Detail

   Beginning with Chapter 3, the project outlined in the first chapter is carried out in more detail, in the form of two chapters devoted to each of the four causes. These chapters vary quite a bit as to their background, topic, style, and method. This is partly due to the fact that I wrote some of them long before I knew that they would get to belong together. Further, Chapter 1 already goes into some of the details concerning Aristotle, and there is no need to repeat things that are already clear enough. Sometimes, it was necessary to add more of the same to the discussion that had already taken place in Chapter 1. In other cases I find it more helpful to engage in an independent discussion and defense of what I take to be Aristotle's position.

   This does not so much concern the two chapters about the material cause (3 and 4). They are based on a published paper (Hennig 2008), which is almost exclusively concerned with things that Aristotle and Plato say. In the published paper, I had treated the Timaeus somewhat carelessly. I have tried to make up for this in the book version, and this has led to considerable changes in the chapters on matter. Both of them differ from the chapters about the other causes in at least two respects. First,

they discuss only one special kind of material cause (namely matter), whereas the other chapters are not restricted to special cases of the respective causes. Second, the chapters on matter refer much less than the others to modern or contemporary debates, such as the problem of material constitution or the concept of matter in modern physics. An excuse for this might be that "in the case of matter, the deepest and most promising insights remain those of Aristotle" (Chappell 1973, p. 680). Still, I could have written about the material cause in the same way as I treat the efficient cause, for instance, or vice versa. I have not, but I do not think that this is a bad thing. It brings in more variety. Chapter 1 shows that the causes form a system, so that many things that may be said about the material cause may also, mutatis mutandis, be said about the efficient cause, and many things that are true of formal causes are also true of final causes. If I had written about all causes in the same way, I would have had to repeat many details.

The two chapters on the formal cause (5 and 6) are somewhat independent and different from one another. The aim of Chapter 5 is to develop the notion of a type and to set it off against competitors, such as extensionally defined sets, intensionally defined classes, biological species, and the like. Chapter 6 defends a reading of Aristotle that is already present in Chapter 1, and that corresponds to the systematic results of Chapter 5. It concludes by invoking the distinction between two directions of fit in order to explain how the primary use of type terms differs from the primary use of predicates. Both chapters have been newly written for this book, though some parts of them are based on a very old unpublished draft.

Causality (that is, efficient causality) is a big topic in modern and contemporary philosophy. Our very idea of what causes are has been shaped by discussions of Hume's conception of causality. Therefore, it is important to directly address Hume and his followers, if only in order to see to what extent their views differ from Aristotle's. Actually, I do not even explicitly compare Humean accounts of causation with Aristotle's notion of an efficient cause. Rather, I discuss Hume's conception in its own right and show in what ways it is mistaken. As it turns out, many of these are

also ways in which Hume's conception differs from Aristotle's. On the face of it, the two Chapters on causality (7 and 8) are rather remote from Aristotle. I trust that what I say about Aristotle's efficient cause in Chapter 1 is clear enough and that it will eventually become clear how it relates to my arguments against Hume and more recent accounts of causation.

In Chapter 7, I argue that efficient causation is not a relation between distinct items. I do not as explicitly argue against the idea of a causal relation between material, formal, or final causes and what they are causes of. Since the efficient cause is the cause that has an effect, it lends itself more easily than the others to the idea of a causal relation between two distinct items (cause and effect). I take it that if it is wrong to speak of a causal relation between distinct items in this case, it is more obviously wrong to speak of causal relations in the case of material, formal, and final causes. The result of Chapter 7 may thus, mutatis mutandis, be extended to all four causes.

The upshot of Chapter 8 is that for a process to be causal is the same as to instantiate a type of natural process. Given my notion of a type, this means that processes are causal insofar as they are subject to standards of typicality. Taken together with claims that I defend elsewhere in this book, this implies that processes are causal if and only if they have a final cause. This latter claim is further elaborated upon and defended in Chapter 9.

In Chapter 9, I draw a distinction between internal and external final causes. This distinction also applies to efficient causes, and it is important for understanding the examples that Aristotle gives of efficient and final causes. I take it that the notion of external final and efficient causes is less basic but more familiar to us than the notion of internal final and efficient causes. In his examples, Aristotle therefore usually refers to external causes. My discussion, in contrast, will most often focus on internal causes. Since external causes can be defined on the basis of the more basic notion, this does not constitute a real difference between my account and Aristotle's.

The final chapter is based on a paper that has been under review for several years

now. In a way, it starts a new agenda. In Posterior Analytics II 11, Aristotle writes that the four causes relate to four kinds of syllogism. How they do so is a difficult question to which I have no general answer. In Chapter 10, I discuss what is probably one of the more difficult cases: reasoning about final causes. Among other things, I suggest canonical forms of practical, teleological, and functional reasoning. In a similar way, one might establish rules for material, formal, and (efficiently) causal reasoning. However, this would easily fill another book, and it is far from clear, for instance, how material and (efficiently) causal reasoning would differ from each other as kinds of reasoning.

Aristotle's Four Causes

It is usually assumed that when Aristotle distinguishes between the four causes a natural scientist should study, he is referring to (1) the matter of a thing, (2) its form, (3) that which initiates its coming to be, and (4) its purpose. On this basis, it must remain obscure what the four have in common such that all and only the four of them should be called causes. In what sense do the form or the purpose of a thing cause this thing? What is the effect of matter or form? And why should there be exactly four causes? For instance, why is not the weight of a thing a fifth kind of cause? If the reason is that in a wider sense of "form," the weight of a thing is one of its forms, why should the material constitution of a thing not also qualify as one of its forms? Moreover, many properties of a thing appear to be reducible to its purely material features. Could we perhaps reduce all of them? It seems that we might either do without material causes or without formal causes. Above all, the final cause does not appear to belong in a list of causes that all natural things must have. In large areas of modern science, we do not any longer refer to purposes, and where we still do so we try not to.

When Aristotle presents his distinction of four causes, he does so without arguing for it. It has therefore been said that we "do not know how Aristotle arrived at the doctrine of the four causes" (Ross 1936, p. 37). In this chapter, I point out that there is in fact a fairly simple rationale for his scheme. This rationale explains Aristotle's division; it need not reflect the way in which he arrived at it. I show that Aristotle's fourfold distinction of causes naturally arises from the combination of two distinctions that apply to all natural processes.[1] First, concerning any natural process, one may distinguish between the thing that undergoes it and the process itself. Neither of these could be studied without in any sense referring to the other. Second, one may

---

[1] Cf. Hantz 1939, who suggests that the four causes emerge "from a consideration of natural productivity" (p. 31).

ask out of what a natural thing comes to be the result of a natural process and one may ask out of what a natural process comes to be what it is. Conversely, one may ask what a natural thing typically comes to be as a result of a natural process, and what the process itself typically comes to be when it occurs. Again, the two questions—out of what something comes to be and what it typically comes to be—can be distinguished but not separated. I maintain that within the cross classification that results from combining these two distinctions, the so-called material cause occupies the slot for that out of which a thing comes to be as a result of its natural development, and the formal cause is what it naturally comes to be. Likewise, the efficient cause is that out of which a natural process comes to occur, and the final cause is what the process naturally comes to be when it occurs.

1.   Natural Processes

We must begin by clarifying what it is for a process to be natural. Aristotle writes that a <u>thing</u> is natural if it has in itself a principle of motion and rest (Physics II 1 192b13-16). He also tells us that in one sense, the typical development of a natural thing may be called its nature (φύσις, 193b12-13). A process should accordingly be natural, first, when the thing that undergoes this process has in itself a principle that governs this process. Second, a natural process should be typical—that is, natural for the thing that undergoes it.

Aristotle also says, misleadingly, that natural things differ from artifacts in that they have an innate impulse to change (ὁρμή, 192b18-19). This seems to imply that a process is natural to the extent that the changing thing is or contains that which initiates the process. But as Aristotle clarifies in Physics VIII 4, that a natural thing possesses an internal principle of motion need not mean that it initiates this motion. The principle may merely govern the way in which it is affected by the impact of other things (255b30-31). Moreover, as Wieland points out, a process may very well be natural for a thing even if the relevant causal chains (in the modern sense of "cause") begin outside this thing (1992, p. 234). By showing that the movements of

an animal or plant have causal antecedents outside this animal or plant, one does not prove that they are not natural for it (cf. Physics VIII 6, 259b1-20). Otherwise, one would have to assume that all natural movements are spontaneously initiated by the things which undergo them, by what Kant would have called "causality of freedom" (Critique of Pure Reason, B 473). This, however, would mean that insofar as a process is natural, it is not governed by the laws of physics. Therefore, we had better not assume that Aristotle's principle of motion and rest is that which initiates motion or stops it.

In Metaphysics Δ 1, Aristotle writes that in one sense of "principle," a principle is "that at whose decision things changed are changed and things altered altered" (1013a10-11, tr. Kirwan). A prince is a principle in this sense: He decides that things are to be done, but need not do them himself. The principle of motion and rest might be a principle in a similar sense. For, in De Motu Animalium 10, Aristotle also writes that just as the government of a city may govern the actions of its citizens without actually bringing them about, the soul of an animal governs the motions of its bodily parts without actually initiating them (703a29-b2). When a government wants something to be done, it issues a law to the effect that doing it is approved or not doing it is sanctioned. The act of issuing the law does not directly initiate the actions that comply with it. It only changes their legal status, as well as the legal status of further actions that promote or prevent them. In a similar way, the soul governs the body by imposing an order on the parts of an animal and assigning functions to them. This does not mean that it initiates the movements by which these functions are discharged. Rather, it determines whether a movement constitutes a function or a malfunction of an organ (cf. Gill 1991, p. 250; Frede 1992, p. 102).

Therefore, when Aristotle speaks of a principle of motion and rest, he has in mind a principle that determines whether a process is natural for a thing and whether the course it takes is typical of a process that is natural for this thing. If a process proceeds in accordance with such a principle, it may be called natural for the changing thing. That a principle of motion and rest is in a natural thing does not mean

that the movements begin within it. Rather, it means that the criteria according to which its movements and its development may be assessed as natural or typical arise from the nature of this thing. The nature of a thing is, in one sense, its form according to its account (τὸ εἶδος τὸ κατὰ τὸν λόγον, Physics II 1, 193a31). The principle of motion and rest arises from this form, and thus from its account. In the case of natural things, we need not look beyond this account in order to see whether what happens with them is natural or not. The principle is not so much in the changing things as it is in the account of their form.

As a consequence, an artifact that initiates its own motions would not thereby undergo a process that is natural for it by Aristotelian standards. This would only be the case if one could give a self-contained account of this artifact that implies a principle according to which its movements qualify as typical or successful. However, artifacts are by definition designed and maintained from outside, and whether they work as they should always depends on what their designers or users want them to do. There is nothing intrinsically wrong with an artifact that does not work. In this sense, artifacts behave according to external principles of motion and rest, even if they move by themselves.[2] In contrast, whether natural things work as they should may be decided by applying standards that are determined by their own nature. That is, the account on the basis of which one can decide whether they are typical and well developed is not primarily an account of other kinds of things. This is why we apply the concept of health to animals and plants but not to artifacts. To be healthy is to be in good condition according to intrinsic standards of typicality.

That the principle of motion and rest follows from the nature of a thing means that it accounts for the way in which the behavior of this thing is distinctive and specific for its type. For instance, what a beaver does when it is building a dam is subject to a certain standard according to which it may count as typical or appropriate. This

---

[2] If there were self-replicating artifacts, their offspring might cease to be artificial to the extent to which their form and development is not any longer governed by external standards of typicality.

standard is tied to the specific nature of beavers. A cat would not do anything typical when building a dam, and it would not be clear what movements would be natural and appropriate for a dam-building cat. Therefore, what a beaver does when it is building a dam is a natural process, whereas what a cat would do when building a dam would not be a natural process. The principle according to which beavers build their dams is not a universal and external law that would also apply to cats but a standard that is specific and inherent to the nature of beavers.

I have said that a process should count as natural, first, when it is governed by a principle of motion and rest inherent in a thing that undergoes this process, and second, when the process is natural <u>for</u> this thing. Since the principle of motion and rest inherent to the changing thing is that according to which the process may be said to be natural or not, both requirements amount to the same. A process is natural for a thing if and only if the account of the thing includes or implies a principle according to which it may qualify as typical or successful. Since this principle determines what processes are typical for a natural thing, and since "nature" may also be taken to refer to the typical development of a thing, the principle of motion and rest inherent to a thing will also determine its nature in this sense.

Natural processes are processes that are governed by standards of typicality and success, which follow from the nature of a natural thing that is involved in them. The following are examples of natural processes. According to Aristotle, when a beaver comes to be, matter that formerly made up a certain amount of menstrual fluid changes into the body of a beaver. When a beaver comes to be mature, a beaver kitten changes into a mature beaver. These are natural processes that beavers undergo. Further, when a beaver builds a dam, a pile of wood changes into a wooden dam. This is also a natural process, since it is natural for beavers to build dams out of wood. However, it is not natural for the wood to turn into a dam. Dams are artifacts, and the principles of dam building apply to a beaver's dam only derivatively, by directly applying to the motions of a dam building beaver. In Physics II 3, Aristotle discusses four questions that a natural scientist should ask about all natural processes and the

natural things involved in them, if she wants to treat them as the natural processes and things they are. The four causes are answers to these questions.

2.   That Out of Which the Thing Comes to Be

The first question that one may ask about all natural processes is out of what the resulting thing comes to be what it is (τὸ ἐξ οὗ γίγνεταί τι, 194b24). The beaver kitten comes to be out of certain matter, the mature beaver comes to be out of a beaver kitten, and the dam comes to be out of wood. However, to determine the required meaning of the phrase "out of ..." is more difficult than it might seem. For instance, a book may be out of paper or out of letters, it might be written out of a certain desire, or it may be out of print. To be sure, some of these uses of "out of" are obviously irrelevant here, but it is not easy to exactly locate the boundary between such irrelevant uses and the relevant ones.

In a first attempt to single out the relevant sense of "out of," one might identify the "cause out of which" with that of which a thing consists. But Aristotle also says that in the relevant sense of "out of," the conclusion of a syllogism may be said to come to be out of its premises (195a18-19). Since it is possible to derive a true conclusion from inconsistent premises, that out of which the conclusion comes to be cannot be that of which it consists. A true statement does not consist of a contradiction in any meaningful sense of "consist." Thus when Aristotle says that the cause out of which remains present in the thing (ἐνυπάρχον, 194b24), he cannot literally mean that in all cases the resulting thing must still consist of that out of which it came to be.

The premises of a syllogism are sufficient for establishing the conclusion, but they need not be necessary. One might therefore think that the cause out of which is something whose continued existence is sufficient for the continued existence of a natural thing. This would also be wrong, however, since the wood out of which a dam comes to be may continue to exist if the dam ceases to exist. Alternatively, one might think that that out of which a thing comes to be is that which makes its continued existence possible, such that it must persist as long as the thing persists. Consider,

however, the matter out of which a beaver comes to be. Since living beings constantly exchange their matter, there will be less and less of this matter left in the growing beaver, and once it has left the organism, the beaver in no way depends on it for its continued existence.[3] If the matter of the beaver were that out of which it came to be, and at the same time that on which it still depends for its existence, we would have to admit that beavers gradually dematerialize during their lives. For there will be less and less of the matter left that is both that out of which they came to be and that on whose presence they still depend. Hence, the original matter of a thing is neither necessary nor sufficient for the continued existence of this thing.

In Posterior Analytics II 11, the place of the material cause is taken by the "necessitating conditions" (94a21-22). This suggests that the material cause of a thing might be something that is necessary for its existence. If this is true, and the matter of a thing is not necessary for the existence of a thing, the matter of a thing is not its material cause. And indeed, Aristotle does not simply identify the "cause out of which" with the matter of a thing. He lists matter (ὕλη) only as one possible instance (195a17).[4] For some natural processes, that out of which the resulting thing comes to be may not be its matter. If this is so, we need not suppose that there is a parcel of matter that persists during all the changes that a living being undergoes in its life. We may thus solve the puzzle of the dematerializing beaver by distinguishing different kinds of changes and focusing on one kind of change at a time. Sometimes a beaver's matter changes its form; sometimes, the beaver exchanges its matter. In the second case, the cause out of which the result comes to be is not its matter.

---

[3] There will be further matter of the same form in the beaver, but not the same amount of this matter.

[4] In (1974, p. 7), Moravcsik therefore refers to the material cause as a "constituent" or "constitutive factor." This is too unspecific, since the formal cause of a thing might also be taken to constitute this thing. And indeed, p. 9 Moravcsik refers to the formal cause as the "structural constituent." Not all constituents can be material causes, then; unless the formal cause is a special kind of material cause.

If the material cause of a thing is not in any case the stuff of which it consists or on whose persistence it depends, we need a more general account of what it is. In Metaphysics H 6, Aristotle indicates the lines along which such an account should proceed. He writes:

> In fact, as has been said, the proximate matter and its form are one and the same thing, one potentially and the other actually. (1045b17-19)[5]

That out of which a thing comes to be is something that potentially is this thing. For instance, a portion of menstrual fluid is potentially a beaver embryo, this embryo is potentially the mature beaver, and a pile of wood and other materials is potentially a dam.

We need to distinguish two kinds of potential here. There are potentials that are, as it were, used up when they are realized. For instance, an unmarried couple has the potential to marry, but (usually) a married couple no longer has this potential. Other potentials are not used up when realized. When a beaver realizes its potential to build a dam, it still has the potential to build a dam. Aristotle emphasizes that the cause out of which a thing comes to be is not used up but remains present in the thing. It would therefore be misleading to speak of the menstrual fluid (rather than the matter it is made of) as that out of which a beaver comes to be. The menstrual fluid does not as such persist in the beaver.

There are two senses in which the cause out of which remains present in the result of a thing's coming to be. First, an actual F is also a potential F. The result of a natural process is thus still the same as the thing out which it came to be. Both are potentially the result; one of them is it merely potentially, the other one is potentially the result

---

[5] See also Chapter 4, section 3.

because it is actually the result (cf. Physics I 9, 192a27-29).[6] The mature beaver is still the same as the beaver kitten insofar as they both are potentially a mature beaver. The only difference is that the beaver kitten changed from merely potentially being something to potentially <u>and</u> actually being that thing. The second sense in which the cause out of which remains present is the following. The matter of a thing is potentially this thing, and the thing that results from realizing this potential is actually this thing. What the matter potentially is and what the thing actually is are exactly the same. What the matter of a beaver potentially is and what a fully developed beaver actually is are the same, namely a beaver. They are both a beaver but in a different modality: one potentially, the other actually.[7] In this sense, the result of a thing's coming to be is still what its matter was.

3.   What the Thing Comes to Be

The matter that makes up a certain portion of menstrual fluid comes to make up a beaver. The beaver comes to be a mature beaver. A pile of wood and other materials comes to be a dam. Taking that as given out of which these things come to be, we may ask <u>what</u> they naturally come to be. That out of which a thing comes to be is something that potentially is that thing. What it comes to be is what this something potentially is. For instance, a young beaver is something that potentially is a mature beaver, and a mature beaver is what this beaver potentially is. There are, of course, innumerable things that a portion of menstrual fluid, a beaver kitten, or a pile of wood might possibly come to be. When we ask what they <u>naturally</u> come to be, we are only interested in a certain range of these. More specifically, we are interested in what a thing would come to be as a result of its natural development, or as a result of a process that is natural for some other thing. The standards of typicality that apply to

---

[6] Cf. Gill (1989, p. 178-80); Frede (1994, p. 192). Bostock reads "merely potential" where Aristotle says "potential" (1995, p. 225), thereby rendering Aristotle's doctrine nonsensical (p. 283-4).

[7] Cf. Sellars (1977, p. 118) and Witt (1989, p. 140). More in Chapter 4, section 3.

the natural development of a thing derive from the nature of this thing. In the case of the beaver dam, the relevant standards derive from the nature of a thing other than the resulting thing.

That into which a thing changes when it develops according to its nature has also been called its formal cause, presumably because the specification of that into which a thing changes involves a specification of the form it takes on. This label, however, is misleading.

There are two reasons why the formal cause cannot be a mere form. First, it is that into which a thing changes, and no thing changes into a mere form (unless it is one). That into which a beaver kitten changes when it undergoes a natural process is a mature beaver, and a mature beaver is certainly not a mere form. Rather, as Aristotle puts it in Metaphysics Λ 3, it is a concrete substance or τόδε τι (1070a11). In 1070a2, he also says that a form (εἶδος) is that into which a thing changes. Since things do not change into forms, he cannot mean a mere form here. Also, when Aristotle says in Metaphysics H 6 (1045b17-19) that matter is potentially what form (μορφή) is actually, form cannot be mere form.[8]

The second reason why the formal cause cannot be a mere form is that the formal cause of a thing is what its definition defines. A definition defines a thing by stating its properties, but this does not mean that it defines these properties. It defines the thing. The correct definition of a beaver does not define a mere form. It defines a formed compound (a beaver). It defines something that has four feet and sharp teeth, and no mere form has four feet or sharp teeth. Therefore, if the formal cause of a thing is what its definition defines, it must be a thing and not a mere form.

In Physics I 7, Aristotle explicitly distinguishes between the formed thing and its form and writes that whereas the underlying thing can be taken to be one (in number)

_____

[8] Again, Bostock diagnoses nonsense, since it follows that the form actually has the features of the formed thing (1995, p. 284; cf. fn. 6 above). But this is precisely what we must assume. (We must not assume that the form of a thing is a separate thing on a par with its particular instances.)

or two (in form), the form that it comes to have must in any case be something further (190b23-29). This seems to imply that the form, as distinguished from the formed thing, must be a <u>mere</u> form. However, in order to comply with Aristotle we only need to assume that the form of a particular thing differs from this thing, and this does not imply that is a mere form. To be sure, Aristotle here also covers cases in which a thing does take on a new quality, and, arguably, qualities are mere forms. But not all forms are qualities. In particular, the formal cause of a thing need not be a quality, and as I show by the end of this section, it may still be distinguished from the particular thing that instantiates it.

Formal causes are not mere forms, but there is a sense of "form" that still applies to them. When we call a living being a form of life, we use "form" in this sense. The formal cause is a form in roughly the sense in which a living being is a form of life. In order to avoid confusion, I will not refer to Aristotelian formal causes as forms but as <u>paradigmatic</u> forms or essences, as Aristotle does in Physics II 3 (194b26: παράδειγμα, τί ἦν εἶναι). Paradigmatic forms and essences are not mere forms. The paradigmatic form of a thing is "what it was to be": the result of its natural development (cf. Johnson 2005, p. 48).[9]

That the paradigmatic form of a natural thing should be that which results from its natural development will strike many as rather odd. For instance, if the paradigmatic form is what a thing comes to be when it develops, the paradigmatic form of a compound thing should also be compound. But does Aristotle not, in Metaphysics Z 4 and Z 10, deny that the paradigmatic form of a thing involves its matter? In Z 10, Aristotle seems to imply that the paradigmatic form of a natural substance cannot be a composite of form and matter. For arguably, the formal cause of a thing is what Aristotle calls an οὐσία in Z 10 (cf. A 3, 983a27), and he says that man, taken in

---

[9] "What it was to be" would not be a correct literal translation of τί ἦν εἶναι, which might be more accurately rendered as "what it is for a thing to be this type of thing." In any case, the "was" does not indicate a time difference, but rather a possible discrepancy between what a thing actually is and what it is supposed to be.

general, is a composite of a sort and not an οὐσία (1035b28-30). What he says might thus indicate that in general, no composite can be a paradigmatic form. However, nothing requires this reading. The reason why man, taken generally, is not an οὐσία may as well be that no universal is an οὐσία, as Aristotle shows in Z 13.[10] Then, the issue would be with the kind of generality, not with the compoundness of man taken in general. What Aristotle means would be: Taken in general, man is a compound other than an οὐσία. Not taken generally, man might be a compound that is an οὐσία.

In Z 4, Aristotle argues that a formula by which a thing is said to be something else cannot be a proper definition (1030a10-11; cf. Posterior Analytics I 4, 73b5-10). For instance, a complex expression such as "pale man" is not a proper definition (and hence does not designate an essence), since to be pale is not in all cases to be a man and hence, in the formula "pale man" something (pale) is said of something else (man). In contrast, to be two-footed is in any case to be a certain kind of animal, and to say that animals of a certain type are two-footed is not to say something else of them; they are not two-footed independently of being this kind of animal. Now it seems that the definition of a substance should not involve a reference to its matter, since the form is said of the matter, presumably as of something different (cf. Z 3, 1029a23-24). If the paradigmatic form of a thing does not involve its matter, it seems that it cannot be a compound. However, as I have pointed out in the previous section, the matter of a thing is something that potentially is this thing, and it remains present in the thing because what it potentially is and what the thing actually is are one and the same. Thus, when we say of matter that it is or makes up a thing, we say of a potential that it is realized, and this is not to say something of something different.

It is true that when we say what a thing is, we do not refer to the particular portion of matter of which it is made. But this does not mean that we refer to a bundle of properties and features instead. For instance, when we say that Socrates is the particular human being consisting of this particular amount of flesh and bones, we say who he is and not merely what he is. We say what he is only when we say that he is a

---

[10] See Chapter 6, section 6.

human being. Again, a human being is not a bundle of features and properties. A human being consists of matter as much as this human being does. The answer to the question raised above is thus that in one sense, a definition should not refer to the matter of the defined thing: It should not involve a reference to this or that particular matter. However, the definition of a thing may very well involve facts about the general material constitution that all things of the respective type share (which Gill 1989, p. 128-33, calls "functional matter"). This material constitution is, in fact, part of their paradigmatic form.

We should recall in this context that natural things are what they are in virtue of an inherent principle of motion and rest. Any account of what a natural thing is must therefore refer to the ways in which such things naturally change. But, as Aristotle argues in Metaphysics Z 11, one cannot treat a thing as something that changes without considering both its material and formal aspects, that is, more generally, without distinguishing between a thing that may come to be something, and what this thing may come to be. Hence, the contrast between the material and the formal cause is essential to all natural things (1036b28-32).[11] If it is essential it must somehow figure in the definition.

I conclude that the formal cause of a natural thing is not its mere form but a

_____

[11] The general point that all natural things are material follows if one reads αἰσθητόν in 1036b28 (cf. Tugendhat 2003, p. 111), for as Aristotle confirms in De Caelo I 9, 278a11, every perceivable thing is necessarily material. If Frede and Patzig are right and one should read αἰσθητικόν (1988, Kommentar p. 210-11), the point is that more specifically, beings capable of sensation must consist of a certain kind of matter. (For a discussion, see Granger 2000.) In any case, that perceivable or perceiving beings necessarily consist of matter does not imply that matter must be part of their definition (Frede and Patzig 1988, Kommentar p. 212). Matter may belong to what a definition defines (i.e., to the essence or paradigmatic form of a thing), but what belongs to the object of a definition need not belong to the definition itself.

paradigmatic formed compound.[12] It is not a property or feature of the thing. This is further confirmed by a passage from De Caelo I 9 (cf. Whiting 1991, p. 634). Aristotle writes that in all things that come to be according to nature, one can distinguish between a form in itself (αὐτὴ καθ᾽ αὐτὴν ἡ μορφὴ, 277b33) and the particular thing whose form it is. He adds that in cases where a type is instantiated only once, as in the case of the world, this might be difficult (278a2-6). In such cases, it is possible to confuse a thing (e.g., the world) with its own paradigmatic form. This shows that the paradigmatic form of a thing is not <u>obviously</u> something other than the thing that has it. There are no circumstances in which a thing would be easily confused with a property. Therefore, if it is easy to confuse the world with its paradigmatic form, this paradigmatic form cannot be a property. What Aristotle contrasts here is not the world and its mere form but this world and a world in general (ὅδε ὁ οὐρανὸς καὶ οὐρανὸς ἁπλῶς, 278a13). The εἶδος and μορφή of this world is thus a generic world.

However, while we must be careful not to confuse the essences or paradigmatic forms of natural things with properties, features, or bundles of them, we must also not confuse them with particular natural things. More specifically, essences must be distinguished from two kinds of particular entity. First, the formal cause of a natural thing is not this particular thing. As Aristotle writes in Metaphysics Z 11, material things do not coincide with their essences (1037b4-5; Albritton 1957, p. 704). They cannot coincide with their paradigmatic form, because the paradigmatic form of a natural thing is one of its causes. This means that we are supposed to be able to explain something about the thing by reference to its paradigmatic form, and as Aristotle points out in Z 17, we cannot explain anything by simply referring to that same particular thing. Therefore, there must be a difference between a particular thing

---

[12] It is sometimes asked whether the object of definition as discussed in Metaphysics H 6 is the form of man or a composite man (cf. Harte 1996). If the essence of man is a paradigmatic compound, it must be both at once.

and its formal cause.[13] Further, the essence of a natural thing is not a particular property instance (sometimes called "trope"). Aristotle writes that although the essence of a house is not generated when a house is built, the essence of this house is (1039b24-25; cf. Albritton 1957, p. 703). One might therefore think that each thing has its own individual essence. Sellars and Albritton (1957) consider the idea that each particular material substance might have its own form, just as each living being has its own life (cf. also Whiting 1986). However, the formal cause is not particular in this sense. It is not a particular form that would only be had by one particular thing.

The paradigmatic form of a particular natural thing differs from this thing not by being something other than a thing (a property or feature), nor by being another particular item, but by being a non-particular thing. The paradigmatic form of a particular beaver, for instance, is what zoology textbooks describe as the life form beaver: not a particular beaver, and not a property or universal, but a paradigmatic beaver. When we say that Betty is a beaver and that beavers build dams, we do not imply that this particular beaver, Betty, builds dams. She might never build a dam and still, the sentence would be true. Betty is an instance of a type of living being that typically build dams. We attribute dam building not to Betty in particular but to her paradigmatic form. This paradigmatic form also functions as a standard with a view to which we may assess to what degree Betty is healthy and typical. Betty differs from her paradigmatic form by potentially being unhealthy and atypical. Paradigmatic forms are actualities (Metaphysics Θ 8, 1050b2-3), and as such, they cannot remain merely potential. As Aristotle says, whereas a particular thing may be said to "strive for" its own perfection, its paradigmatic form "cannot desire itself, because it is not lacking" (Physics I 9, 192a18 and 20-21, tr. Coughlin). Paradigmatic forms cannot fall short of their essences, because they are these essences. For instance, the soul of a man is the same as its essence, for it is the actuality of an organic body (De Anima II

_____

[13] Some things, such as mathematical objects or souls, coincide with their essence. Therefore, not all essences are formal causes. Formal causes are essences of things that do not coincide with their essences.

1, 412b5-6). A particular man, in contrast, is not the same as his essence (Metaphysics H 3, 1043b2-4).

Recall again that a process is natural insofar as it is governed by a principle of motion and rest that is implied by a proper account of the nature of the changing thing. We may now restate this as follows: Whether a process is natural for a thing depends on what the thing is—that is, on its paradigmatic form. The paradigmatic form of a thing determines the typical course of the processes that are natural for it. For instance, the process by which a beaver kitten turns into a mature beaver is natural for the beaver kitten by virtue of the paradigmatic form of beaver kittens. The beaver kitten was to be a mature beaver. Although it may not yet fully exemplify the features associated with this form, since it is not yet mature, a mature beaver is its paradigmatic form, since for a beaver kitten it is natural to grow into a mature beaver.

A natural thing need not actually come to be what it was to be. A beaver kitten may fail to develop into a mature beaver. If this happens, it falls short of its own paradigmatic form. This shows again that the paradigmatic form of a thing cannot be a set of features it has: Nothing can fall short of its own features. It also shows that the paradigmatic form of a natural thing cannot coincide with this thing, for nothing can fall short of itself.

That the paradigmatic form of a natural thing is what it is supposed to turn into as a result of its natural development also means that the paradigmatic form of a thing does not change (Metaphysics Z 8, 1033b5-8). What a developing beaver is supposed to be is clear all along, and it does not undergo any change when the beaver changes. In this respect, the paradigmatic form is also not the same as the particular thing that actually undergoes the process. The paradigmatic form of a thing is what it comes to be in a natural process, but it is not that which changes and comes to be something. The paradigmatic form of a thing is defined by the definition of that thing, and the definition of a thing does not change when the thing changes.

In general, we may refer to a changing thing (1) as a changing compound, (2) as that which has the potential of being the result of the process and remains present in

this result, or (3) as what it was to be (cf. Tugendhat 2003, p. 74). Insofar as a compound is first in one and then in another state, it changes. Insofar as the same thing is potentially the result of a process (the cause out of which), it does not change; since when the process has taken place the thing is still potentially the result of the process. Finally, the paradigmatic form of a changing thing does not change, since its definition does not change. Hence, the matter and form that are involved in a natural process do not themselves change.

Thus what a thing is, its paradigmatic form, differs from the particular thing itself in at least three respects. First, other things of the same type may share it, such that it is what other things are as well. Second, since the paradigmatic form of a naturally changing thing is what it is supposed to change into, things may fall short of their own paradigmatic form. The paradigmatic form sets a standard that the thing might fail to meet, and when this happens the thing is not what its paradigmatic form is. Third, the paradigmatic form of a thing does not change when the thing changes. The thing and its paradigmatic form have different modal properties. However, in one sense the paradigmatic form is still what the thing is: It is not <u>this</u> particular but still <u>a</u> generic instance of its type.

4.   Whence the Process Comes to Occur

Turning from naturally changing things to the natural processes they undergo, we may again ask out of what they come to be, or whence they come to occur (cf. De Generatione Animalium I 18, 724a20-35; Metaphysics Δ 24, 1023a29-30). Aristotle's system of causes would be neat and simple if that out of which a natural process comes to occur were something that has the potential to turn into this kind of process. Just as the "cause out of which" of a natural thing is something that potentially is this thing, the efficient cause of a natural process would be something that potentially is a natural process. Since processes unfold in the course of time, one might identify such a potential process with the beginning of the process that it potentially is. For instance, when a beaver starts biting into the trunk of a tree, one may think of this

action as a potential felling of a tree.

This idea, that the efficient cause of a process might just be its beginning, which potentially turns into the rest of it, is attractive in its own right. As such, it will be spelled out and defended in Chapters 7 and 8. However, whereas Aristotle makes clear enough that the material cause of a natural thing is potentially this natural thing, he does not explicitly say that the efficient cause of a natural process is potentially this process. The closest he gets to this is in Metaphysics Θ 6, where he draws an analogy between material and efficient causes. He writes that οὐσία relates to ὕλη as κίνησις does to δύναμις (1048b8-9).[14] This indicates that δυνάμεις relate to natural processes in roughly the same way in which matter relates to natural things (cf. Frede 1994, p. 183).

Alas, the analogy is much less straightforward than one might wish. Aristotle refers to the efficient cause as that "whence the first principle of alteration and rest" comes (ὅθεν ἡ ἀρχὴ τῆς μεταβολῆς ἡ πρώτη ἢ τῆς ἠρεμήσεως, 194b29-30).[15] We have seen that the principle that governs a natural process arises from an account of the nature of a natural thing that undergoes this process. Accordingly, if the efficient cause of a process is that from which the relevant principles of motion and rest arise, it should be the same as the account of the nature of the natural thing that undergoes it, rather than a potential natural process. Further, Aristotle also refers to the efficient cause as that which produces what is produced, or alters what is altered (τὸ ποιοῦν τοῦ ποιουμένου καὶ τὸ μεταβάλλον τοῦ μεταβαλλομένου, 194b31-32). Here, the efficient cause is something that acts, and its effect is a thing that is created or altered by its action.

Rather than being a potential process, the efficient cause of a natural process is thus a thing, or the nature of a thing, which is the source of a principle that governs

---

[14] Later, he contrasts κίνησις with ἐνέργεια (1048b28) and identifies the latter with οὐσία (Θ 8, 1050b2-3).

[15] Cf. 195a22-23. He also refers to it as ὅθεν ἡ κίνησις (195a8) and ἀρχὴ κινήσεως (195a11).

this process. Although both efficient and material causes make something possible (processes and things, respectively), they do so in quite different ways. An efficient cause actualizes its potential by giving rise to an action, whereas matter actualizes its potential by coming to be what it potentially is. So it seems that the efficient cause of a process is not a potential process. The matter of a natural thing is a potential thing, but as Aristotle describes it, the potential that corresponds to a process is not a potential process. It is, rather, the thing that initiates this process or its capacity of doing so.

Moreover, the efficient causes that Aristotle mentions in his examples are usually things other than the ones that undergo the process they cause. Aristotle calls a counselor an efficient cause of the action he recommends, and the counselor need not be involved in this action as an agent. Likewise, when the father is an efficient cause of his offspring, the doctor of the health of her patient, or the sperm of an animal (Physics II 3, 194b30-31 and 195a21-22), one thing is an efficient cause of something that happens to another thing. Father, doctor, and sperm give rise to processes, some of which result in things; they do not at all turn into these processes or things. Aristotle's efficient causes seem to be strictly external to what they cause.

However, accounts according to which efficient causes are things or natures, whereas their effects are processes, run into well-known difficulties. For instance, if causes and their effects would belong to two different categories (things or their natures vs. processes), there could be no transitive causal chains. Usually, when something A causes something B, it makes sense to ask what caused A, or what the effect of B is. If all causes are things and their natures, but effects are processes, these questions do not make sense. The relation between the nature of a thing and a natural process that is governed by it is what we would nowadays call agency, or agent causation (cf. Chisholm 1976). Agent causation does not admit of transitive causal chains. Where A is an agent and B is an action, it does not make sense to ask who performed the agent, or what further action the action performed.

Further, if causes are things or natures and effects processes, there is no reason

why a cause should act at a particular time rather than sooner or later. The agent is, as it were, one step too remote from the action: It is not the mere presence of the agent that leads to the action but rather something that the agent does. The agent causes her own actions not by just being there but only insofar as she is already acting. This however means that the agent is not actually the cause of her actions. Her <u>actions</u> are causes of further actions or events.[16]

I have said that Aristotle's system would be simpler if the efficient cause were something that potentially is a process; that according to Aristotle, it is rather a thing or its nature; and that an account according to which causes and effects belong to different ontological categories faces certain difficulties. There are thus, it seems, two different accounts of what efficient causes are: one that would work, and Aristotle's account, which does not. According to the first, efficient causes are potential or potentially completed processes; according to the second, efficient causes are agents. In fact, however, the contrast between these two accounts is not as big as it might seem. For we cannot, in this context, separate an agent from what it does, nor can we separate an action from its agent and the agent's capacities. The efficient cause of a natural process may be a natural thing, but it causes this process only insofar as it changes. Further, what an efficient cause causes is never a mere change but always a change in a natural thing.

First, when we call an agent a cause, we cannot separate this agent from what it does (Annas 1982, p. 321). A natural thing may be said to be that out of which a process comes to be only insofar as it acts. A builder is as such only a potential cause of a house; she is an actual cause of a house only if she is actually building one (cf. Physics II 3, 195b3-6 and b16). Therefore, a beaver that does not act would not determine the course of any process.

Second, we cannot separate the action from its agent. When building a dam, the beaver does not just kick off a process that would then run on its own. Rather, it continuously keeps doing things that contribute to building the dam, and it keeps

_____

[16] These arguments reappear in Chapter 8, section 2.

adjusting the movements that, for instance, the wood makes on its own. The beaver thus remains involved in what it causes. It does not simply initiate the movements that take place; it directs them and thereby causes them to be what they are. Further, what the beaver keeps doing is an action of building a dam only insofar as it is governed by the principle of motion and rest inherent to the beaver. The same motions, when performed by an animal of a different type, would not be the same action. They would not be governed by the same principle of motion and rest. Hence, any specification of the steps that are naturally involved in building a beaver's dam must refer to the nature of the beaver. The efficient cause of a natural process cannot be a mere process or the beginning of one, it must be a natural process undergone by some specific natural thing. The nature of the thing determines the effect as much as the occurrence of the process does.

That we cannot clearly separate between things and processes in this context means that it does not really matter much whether the efficient cause of a process is a thing or a process. It can only be a things insofar as it acts, or a process insofar as it involves a certain kind of thing. In the end, the efficient cause is a changing thing that leads to there being a changing thing, or what is the same: a change in a thing that leads to a change in a thing. We may therefore also say what Aristotle does not explicitly say: that the efficient cause of a natural process is something that potentially is that natural process. This is not manifestly Aristotelian. In particular, it does not seem to cover many of Aristotle's example cases, in which one thing causes something to happen to another thing. However, the conception of efficient causation as developed here is not far from what he says; and it may be easily extended so that one can understand what Aristotle says on its basis. It must gain its credibility not by association with a Bekker number but by fitting into an overall plausible account.[17]

Thus in a basic sense, the efficient cause of a process is for this process what the material cause of a thing is for this thing. That out of which the beaver comes to be is that which potentially is a typical and mature beaver, and that out of which building a

_____

[17] This account is given in Chapters 7-9.

dam comes to occur is that which potentially is a successful building of a dam. When the beaver has come to be mature, the same thing that potentially was the mature beaver is both actually and potentially the mature beaver. When building the dam comes to occur, the same process that potentially is the building of a dam is also an actual dam building. This analogy is encouraging, since it indicates that the final cause might be what a process naturally comes to be when it occurs, just as the formal cause is what a thing comes to be when it develops according to its nature. Let me therefore turn to the final cause.

5.   What the Process Comes to Be

   Concerning natural processes, we may again take that out of which they come to be as given and ask <u>what</u> they come to be when they occur. When a beaver is felling a tree and arranging its parts in a certain way, what it is doing may be said to turn out to be, or turn into, the building of a dam. We may also say, on this basis, that the beaver acts in order to build a dam. What a process comes to be may therefore in a first approximation be taken to be any of the following: its end, its direction, the resulting process itself, or that for the sake of which it occurs.

   That into which a process turns should not be taken to be only its final stage. For instance, if a beaver leaves out one of the initial steps that are involved in building a dam, what it is doing might not turn out to be proper dam building (even though a dam might result). A correct account of what a process turns into will therefore involve an outline of the entire pathway by which its final stage is reached. What a process turns into should be taken to consist of several steps and stages. When Aristotle says that it is ridiculous to identify the final cause of a human's life with her death, this may well be what he has in mind (Physics II 2, 194a30-32). For although death does indeed belong to life as its final stage, a complete life will typically involve much more than death. Death is not the final cause of life because death alone

does not constitute the completion of a life.[18]

For similar reasons, that into which a process turns cannot only be its direction. The direction of a process is that into which it proceeds, but it does not determine its entire natural course. In De Anima II 4, for instance, Aristotle refers to a movement that has a direction but no final cause. The movement that fire makes is directed, primarily upward but also in all other directions, and it seems to be typical for fire to move in these ways. However, Aristotle says, it has no limit (πέρας, 416a15-17). Since the Greek equivalent for "to proceed" (περαίνειν) derives from the same root as the word for limit (πέρας), the limit of a process may be taken to be the pattern according to which it proceeds.[19] That the spread of fire does not have a limit does not mean that it will never stop but that it does not follow a specific course that it might complete in a given time. It has an end, to be sure, when there is nothing left to be burnt. However, the point where the spread of fire stops is not determined by the nature of fire; it is determined by what the fire burns and by the circumstances in which the fire burns it.

As Aristotle writes in Metaphysics Δ 17, in one sense of "limit," the limit of a

---

[18] The aim of 194a30-32 is to show that the student of nature must study both the matter and the form of natural things, just as the end of a teleological process should not be thought of in separation from what leads to this end.

Elsewhere, Aristotle writes that wine does not naturally turn into vinegar and that humans do not naturally turn into corpses (Metaphysics H 5, 1044b34-45a6). Obviously, he cannot mean that it is not natural for humans to die. The reason why it is not natural for them to turn into corpses is that humans and their remains have no common nature and that therefore, a change from one to the other cannot be governed by one and the same principle of motion and rest. It cannot be natural for humans to end up being corpses because there is no time at which a human being is a corpse. The human being ceases to exist as soon as the corpse comes to be. The same is true for wine and vinegar.

[19] Cf. Parts of Animals I 1, 641b23-25; Lennox 2001a, p. 146.

thing is its essence (τὸ τί ἦν εἶναι, 1022a9). Therefore, the limit of a process may also be taken to be its essence. Just as the essence of a natural thing is its paradigmatic form, the essence of a natural process will then be a specific course that processes of its kind typically take. For natural processes, this typical course is determined by the principles of motion and rest inherent in the natural things that undergo them. If final causes are for natural processes what formal causes are for natural things, they should be essences of natural processes. If they are, the final cause of a natural process should be its limit.[20]

That final causes are essences of natural processes is, again, not something that Aristotle explicitly says. It follows from the analogy between final and formal causes that he suggests, e.g., in Metaphysics Θ 6 (1048b8-9), and it has consequences that are illuminating and compatible with what Aristotle says elsewhere.

The essence of a process is its limit or typical course. Since natural processes are governed by principles of motion and rest inherent in natural things, the typical course of a natural process is determined by the nature of a natural thing that undergoes this process. If final causes are essences of processes, "for the sake of a final cause" really means as much as "according to a principle of motion and rest."[21] A natural process may thus be said to occur for the sake of the limit according to which it proceeds, so that the final cause of a natural process is primarily the limit or standard pattern according to which it proceeds. This limit is that with a view to which it may be assessed as natural, typical, and complete. It may function as a standard of typicality.

Further, if final causes are essences of processes, it is obvious why every causal process must have a final cause. A cause must always be the cause of something specific. Therefore, as long as a causal process is going on, there must be something specific towards which this process leads. There must be such a thing as the course that one would expect it to take. If the essence of a process is its typical course, all

---

[20] In the Philebus, Socrates establishes a close connection between τέλος and πέρας; cf., for instance, 24A-B.

[21] This will be confirmed in Chapter 9, section 1.

causal processes must therefore have an essence. This is why Aristotle can suppose, in Physics II 8, that everything either happens by chance, or as a by-product, or for the sake of something (199a3-5).

Aristotle sometimes says that the final cause of a process is something good.[22] This is compatible with the assumption that final causes are essences of processes. For in one sense, to describe the essence of a thing is to describe what it takes to be a good instance of its type. Likewise, to describe the essence of a process is to describe what it is to be a good instance of its type.[23] In this sense, the essence of a natural process is something good, and for a process to have an essence is to proceed for the sake of something good. Since the limit of a process is determined by a principle of motion and rest inherent in a thing, a process may also be said to occur for the sake of the good of this thing. The essence of a natural process follows from the essence of a natural thing that is typically involved in it, and in this sense, it proceeds for the sake of the good of the thing.

So far I have argued that the final cause of a natural process may be taken to be its essence, and that the essence of a process is the limit (πέρας) according to which it proceeds. This limit is that with a view to which a particular process counts as a typical or atypical instance of its type. It is the course that processes of its type would typically take. Since the final cause is the typical course that we specify when we say what type of process it is, it is also the essence of the process, in the same sense in which the paradigmatic form is the essence of a natural thing. I now argue further that, just as the paradigmatic form of a natural thing is a generic instance of its type, the essence of a natural process is also a generic instance of its type. The reason for this is that in general, there is no real difference between a process and the course it takes. When we observe a process taking its course, we do not observe anything further beyond the process itself. That a process takes its course is not a further process that the first process undergoes. This is why it sounds awkward in the first

_____

[22] See Chapter 9, section 2.

[23] This point will also be further clarified in Chapter 9.

place to say that a process comes to be or comes to complete a certain course: Its coming to be and completing this course is nothing other than the process itself. Hence, the typical course of a certain type of process is a typical process of this type. As a consequence, the present account of the final cause matches that of the formal cause given above in a further respect. The paradigmatic form of a thing is what its definition defines, and although the object of the definition is not a mere form but a formed compound, it is not a particular instance of the defined type. It is a generic and paradigmatic instance. Likewise, the final cause of a natural process is what this process "was to be": not this, but a process of its type. It is what its correct definition defines, and hence a generic and paradigmatic instance of this type of process. Just as a natural thing may fall short of its paradigmatic form, a natural process may fall short of its final cause. This happens when it is interrupted before it is complete or when it is completed in an atypical way.

This is why Aristotle can say, in De Generatione Animalium I 1, that the formal and the final cause of an item are in several respects the same (715a5-9). One reason for this is that formal and final causes perform the same function concerning things and processes, respectively. The final cause is for a natural process what the paradigmatic form is for the natural thing whose inherent principle of motion and rest governs this process. Both the formal and the final cause are accounts (λόγοι) of the entity in question.[24] Another way in which formal and final causes are closely linked is that natural processes are governed by principles of motion and rest that arise from the nature of things. The principles determine the limit of the process, and they arise from the paradigmatic form of the thing that undergoes this process. Therefore, the final cause of the process follows from the formal cause of the thing. Conversely, since the ways in which natural beings change belong to their nature, the natural processes they undergo belong to their paradigmatic form. Thus, the formal cause of a natural thing follows from the final cause of its development and of the processes in which it is typically involved. The distinction between the paradigmatic form of a

---

[24] Cf. also De Anima II 4 (415b14-15) and Parts of Animals I 1 (639b14-16).

thing and the typical course of its natural development is therefore to a certain extent an abstract one. It may often seem that there is exactly one mature form of a living being and that this is the paradigmatic form. However, natural things develop in specific ways, such that all stages of their typical development belong to what it is to be that type of thing. Therefore, a proper account of the paradigmatic form of a natural thing must include an account of its natural development.

Despite these close relations between final and formal causes, it should be noted that they also differ, formally, in at least one important respect. There might be a thoroughly perfect and presently existing natural thing. It is, for instance, at least possible to imagine a fully typical and complete beaver. However, it is for conceptual reasons impossible to imagine a perfect and still presently ongoing natural process; for instance, a perfect felling of a tree. For to imagine a process as present is to imagine it as ongoing, and as long as the felling of the tree is going on, it is not yet over. But at least for such processes as felling a tree, to be complete would involve being over, so that this process cannot be complete as long as it is going on. If it is over, it is not any longer going on. Moreover, for every stage of felling a tree, at least one other stage will not be going on at present. Therefore, it is impossible to imagine a perfect and yet presently ongoing felling of a tree. In this sense, the final cause of a natural process is external to any concrete instance of its type. The typical course of a type of process necessarily extends beyond any of its particular instances. It includes what should have happened before each present instant and what is supposed to happen after it.

Just as there is a sense in which the material and the formal cause of a natural thing are the same (one potentially, the other actually), the efficient and the final cause of a process are in one respect the same. When all steps occur that are involved in bringing about a process, they constitute this process.[25] Taking steps in order to fell a tree is

---

[25] Aristotle writes that an eclipse is not only caused by the interposition of the earth but the interposition of the earth is also its essence (Posterior Analytics II 2, 90a14-18). Cf. also Chapter 6, section 2.

felling a tree. In this sense, the efficient cause is potentially the same as the process actually is into which it naturally turns. On the other hand, the final cause of a natural process is also what this process is, since it is its essence. It is the typical course that processes of its type take, and as I argue above, this is the same as a generic and paradigmatic instance of this type.

The efficient cause also differs from the final cause, in the same sense in which the paradigmatic form of a natural thing differs from that particular thing. The efficient cause of a natural process is that out of which the process comes to occur, and the final cause is what this process would typically come to be—the course that it takes when it proceeds according to the relevant principle of motion and rest. The final cause is what a process is supposed to be, and what a process is supposed to be is not this particular process but a process of this type. This means that the final cause of a process may be shared by several instances of the same type. One natural process may be what another natural process is as well. Further, the final cause of a natural process does not occur as such when the process occurs (cf. Physics II 7, 198b1-4). Again, we must not confuse the essence of a particular entity with that very entity. As I argue above, the definition of a thing should not include a reference to a particular portion of matter. In the same sense in which the particular matter of a thing does not belong to its essence, the particular time at which a process occurs does not belong to its essence.[26]

6. Conclusion

It should be fairly clear by now how the four Aristotelian causes can be brought into a system and why a natural scientist must refer to all of them in her explanations, as Aristotle claims in Physics II 7 (198a21-24). Since the proper account of a natural thing implies a principle of motion and rest, we can only understand what it is to be

---

[26] There are, of course, events that essentially occur at certain dates, like Ramadan or Christmas. But the date that is mentioned in their definition is a generic date. The case is analogous to the case of functional matter dealt with above.

that type of thing by taking into account in what ways such things typically change. Conversely, since a natural process is a process that is governed by a principle of motion and rest in the changing thing, we could not understand natural processes without also investigating the nature of the things that are involved. Natural things and natural processes belong together. Further, as Plato points out in the Timaeus (e.g. 47E-48A), the student of nature must always ask two kinds of question about the object of her studies: "What is it for, or where does it lead to?" and "What is required or necessary?" (cf. Charlton 1970, p. 115). The first question concerns what a thing or process is supposed to be and do; the second is about that which potentially satisfies such requirements.[27] When investigating persistent things, the natural scientist must investigate both their paradigmatic form and that which may come to realize this form —for instance, their matter. When studying natural processes, the same questions take the form of the question about their final cause or typical course and the required steps that they typically involve. In each case, the two questions are complementary, just as the matter and form of a thing complement each other. The efficient cause is for a natural process what the material cause is for a natural thing. The final cause is for a natural process what the formal cause is for a natural thing. Whereas the material cause is that which potentially is a natural thing, the paradigmatic form is what it naturally comes to be; and whereas an efficient cause is that which potentially is a natural process (or a changing natural thing), the final cause is what this process naturally comes to be.

This explanation why there are four causes involves the assumption that two of them primarily apply to natural processes. It is not obvious, on the basis of what

_____

[27] Aristotle does not distinguish four "explanatory chains," as Moravcsik writes (1991, p. 43), but only two: Material and formal cause belong to one, efficient and final causes to the other (cf. Schofield 1991, p. 39). The two elements of both chains mark the different directions into which one may follow the chain. One of them is the how-direction (where something comes from, what is needed, how it is possible), the other the what-direction (what results, what it is needed for, what is possible).

Aristotle writes, that the efficient and the final cause concern processes and not things. On the face of it, the cause out of which a process comes to be is the nature of the thing that implies the relevant principle of motion and rest, and its complement is the thing that results from this process. For instance, the father is the efficient cause of generation because his nature supplies the principle of typicality that applies to this process, and the final cause of generation is the offspring that is generated. Wieland accordingly assumes that in Physics II 3, Aristotle only discusses the causes of persistent things (1992, p. 111 and 266). Similarly, Moravcsik writes that all four causes primarily apply to substances (1974, p. 5).

However, Aristotle could not discuss the causes of natural substances without at the same time considering the ways in which they change. I have argued that even if the efficient cause is a thing, it can only be a thing insofar as it changes. Likewise, even if the final cause is attributed to a thing, it belongs to this thing only because it belongs to a process that this thing undergoes. Therefore, efficient and final causes are in a basic sense causes of processes and not of things. And indeed, when Aristotle introduces the final cause in Physics II 3 he calls it a τέλος and οὗ ἕνεκα (194b32-33), without saying what it is a τέλος of; it is natural to supply the subject of the previous sentence, so that it turns out to be the τέλος of a process. Also, in Metaphysics A 3, he calls the final cause the opposite of the efficient cause (983a33), and since the efficient cause is the ἀρχὴ τῆς κινήσεως, the final cause should be the τέλος τῆς κινήσεως.

Nonetheless, since Aristotle often refers to things as efficient causes and ascribes final causes to things, the account of the efficient and final cause as causes of processes rather than things is better taken to be a background theory that explains the difference between formal and final causes and the difference between material and efficient causes. It remains in the background because many things we can say about the way in which natural things change may also be said about these things themselves. Therefore, things can be said to be efficient causes or have final causes, even though efficient and final causes primarily belong to processes. This is all the

more appropriate in the case of natural things, whose nature implies a principle of motion and rest according to which they move, change, and develop. Therefore, Aristotle often simply attributes final causes to things, where he could also speak of final causes of processes in which such things are typically involved. It is important to note, however, that this does not at all contradict the assumption that final causes are essences of processes. Whenever a thing has a final cause, there is also a process that may be said to have this or an analogous final cause.[28]

As I have announced in the Introduction, Chapters 3-10 will deal with the four causes in more detail. These Chapters will not just spell out and defend what has been said here. They often go beyond Aristotle, but all are written in what I take to be an Aristotelian spirit. For instance, in Chapter 7 I present independent arguments for the account of efficient causation that I attribute to Aristotle, without deriving these arguments or even the view itself from Aristotle. I argue that in order to understand what (efficient) causality is, we should not primarily ask when a relation is causal but when a process is causal. In Chapter 8, I demonstrate that a process is causal if it is an instance of a type such that it is subject to certain standards of typicality, on the basis of which one may predict its future course. If all causal processes are instances of such types, it follows that they must all have a typical form, of which they may fall short by being hindered or interrupted. This typical form is their final cause.

Before returning to the four causes, however, the following chapter will introduce the distinction between two directions of fit, which will be needed in order to explain the difference between essences and properties in Chapter 6, section 5.

---

[28] This is explained in more detail in Chapter 9, section 3.

Two Epistemic Directions of Fit

Aristotle writes that in order to even begin investigating nature, one must start from the assumption that natural things are subject to change (Physics I 2, 185a12-14). That they may change is thus not something that one could find out about natural things by studying them. For in order to study them, one must assume beforehand that they are subject to change. As I argue in the previous chapter, the four questions to which Aristotle's causes are answers must apply to things that typically undergo natural processes. Natural things are capable of change. This implies that they must be material things with a paradigmatic form, which undergo processes with efficient and final causes. All this follows from something that one must assume beforehand, in order to bring them into view as the natural things they are. In this sense, we may know a priori of all natural things that they have a paradigmatic form that may fail to be realized and that they are involved in processes that may fail to be completed. This is synthetic a priori knowledge. It tells us something about the object of our knowledge, not only about our ways of thinking and talking about it.

Kant claims that synthetic a priori knowledge is possible because the objects of our knowledge conform to our cognition rather than vice versa. When something conforms to something else, the reason is often that the second has acted on the first, or that someone has taken the second as a pattern for making or manipulating the first. That an object conforms to our cognition might therefore be taken to mean that our cognition shapes or manipulates its objects according to the pure forms of intuition and thought. If this were true, however, our cognition would be an instance of what has traditionally been called practical knowledge. Our knowledge would shape its object, and thus be the "cause of what it understands" (Anscombe 1957, §48), or at least the cause of the form of what it understands. I take it that when Kant says that objects conform to our cognition, he does not want to imply that our cognition of them is practical rather than theoretical. I therefore ask how cognition can determine

the form of its object without creating or altering it in any sense. In order to answer this question, I refer to Austin's distinction between two directions of fit (How To Talk, 1979). This distinction is often employed in order to distinguish beliefs from desires, or assertions from expressions of intention (cf. Humberstone 1992; Sobel and Copp 2001). Roughly, a belief is taken to be something that is supposed to fit a state of affairs, whereas a desire is taken to be something to which a state of affairs is supposed to fit. Austin, however, does not at all use the distinction between two directions of fit in this way. In How to Talk, he is only concerned with the differences among several kinds of <u>assertive</u> statements.

I will now first raise the question of how our cognition can determine its object without creating or manipulating it, to which Austin's distinction will be shown to provide an answer. Then I introduce Austin's terminology and work toward an answer to this question. I introduce Austin's distinction in its original form and then take several small steps in order to fit it in the context of Kant's logic and epistemology.

1.   Archetypes and Ectypes

In his letter to Herz of February 21st, 1772, Kant considers the idea that a representation may relate to its object in one of two different ways: such that the object causes the representation or vice versa (Akademie Ausgabe X, p. 130).[1] That he describes the relation between a representation and its object as a causal one seems to raise the obvious question why effects should in any way represent or resemble their causes. Since not every effect of an object represents this object, it seems that no merely causal account of representation is sufficient (Ameriks 1992, p. 332; Longuenesse 1998, p. 18). However, Kant here appeals to what has been called an

---

[1] This is p. 125 in the 1900 edition. Cf. also the Critique of Pure Reason B XVII and 124-5.

exemplary cause.[2] Exemplary causation is a relation not between an efficient cause and its effect, but between a paradigm or model and its copy. That A is an exemplary cause of B means that A is a paradigm of which B is a copy, and it is not difficult to see why a copy should represent what it is a copy of. I adopt Kant's terminology and call the paradigm an <u>archetype</u> and its copy an <u>ectype</u>. In his letter to Herz, Kant thus asks whether representations relate to their objects as archetypes relate to ectypes. It is conceivable that our representations are ectypes of objects and the objects their archetypes. In this case, it would indeed be easy to see why a representation should represent its object: It would be modeled after the object. But this can only apply to merely <u>passive</u> representations. What is much more difficult to explain, says Kant, is how spontaneous cognitive acts, which are not passive, can represent objects. If there are really only two possible ways in which such acts can relate to their object, we have the choice between the following two scenarios.

First, our spontaneous acts could represent their objects by being mere ectypes of them. Then, however, they would represent their objects only <u>insofar as</u> they (the acts) are merely passive and not spontaneous. Everything in them that represents the object would also passively derive from it, and everything that goes beyond what derives from the object would be a possibly distorting, additional, and non-representing element. Now as Kant argues, our sensory experience necessarily involves an activity of the mind; if not judging, then at least what he calls synthesis, which is an act of going through and holding together the manifold of sensory impressions (Critique of Pure Reason, B 103). This implies that one cannot divide our cognitive acts into a passive representing and an active non-representing part. The part of them that is merely passive could not represent anything on its own. It also implies that one cannot check the results of one's cognitive activity against what they represent, since in order

---

[2] Micraelius writes: <u>Exemplaris causa</u> dicitur, qvae in mente artificis est efformata, & ad qvam imago in re ipsa exprimitur. <u>Exemplar</u> est causa imaginis: sed <u>imago</u> est signum repraesentans suum exemplarum. Correspondentia autem imaginis & exemplaris dicitur similitudo (1653, p. 484, s.v. "exemplar").

to bring into view what they represent, one needs to perform these very cognitive acts. If the activity of our mind were a distorting and non-representing element, then all experience would inevitably be distorted by it. There would be no reason why it should not differ arbitrarily in different subjects, which would lead to relativism (cf. B 168). Therefore, if our acts would represent an object only insofar as they are passive, they could not represent an object insofar as they are <u>acts</u>.

Alternatively, our spontaneous acts might be archetypes, of which their objects are ectypes. However, ectypes are copies of archetypes, and they represent their archetypes by conforming to them. Therefore, if the objects were ectypes of our spontaneous cognitive acts, the <u>objects</u> would be representations of these acts and not vice versa. Further, with regard to the object of our theoretical knowledge, the world, there is only one intellect of which one might possibly say that it is archetypical, and this is the intellect of its creator. For the creator, knowing and creating would be the same. The object of her knowledge would conform to it because the world would literally be a result of her cognitive acts (B 145). An intellect that causes its own object would be a divine intellect, and it would not be discursive but intuitive since it would immediately relate to its object (B 135, 139, 145). Its knowledge of the world, however, would be practical knowledge, and the problem is to explain how theoretical knowledge can be spontaneous.

Kant cannot accept any of these scenarios. We cannot solve the problem of how our spontaneous cognitive acts can be cognitions of objects by asserting that instead of being spontaneous acts they are merely passive reactions to objects. Nor can we do this by asserting that instead of being cognitions, they are makings or shapings of objects (cf. Ameriks 1992, p. 338).

Longuenesse claims that according to Kant, our cognitive judgments represent objects not because the judgments cause their objects but because our imagination shapes the object of cognition in an effort to make our judgments possible (1998, for references see her index s.v. "conatus"). Should this be Kant's answer, I shall confess that I do not like it; it amounts to the claim that our judgments represent the

phenomenal world because this phenomenal world was shaped by some unconscious but still goal-directed mechanism in our souls, such as to make this possible. I do not believe that this is the way in which Kant wanted to make his Copernican turn.

In his letter to Herz, Kant concludes that with respect to the objects of theoretical cognition, the human intellect is neither archetypical, since it does not create its own object, nor ectypical, since it is not merely passive.[3] But what else can it be? Is there any way in which something may stand in a representational relation to something else, other than being its archetype or ectype? In other words, how can a concept determine the features of its object, so that the object must conform to it, and still (cognitively) represent this object? With this question in mind I turn to Austin.

2. How To Talk

Austin asks the following question: What is the difference between calling something an F and describing it as an F? In order to tackle this question he introduces a simple model, which he then adjusts in a second step. Within his simple model he introduces the difference that interests me here: the difference between stating that something is an F and casting something as an F.

Austin does not set up his model in order to explain how thought and language relate to reality but only in order to explain how ordinary speakers draw the distinction between calling something an F and describing it as an F.[4] His question is

_____

[3] Pace Longuenesse, who thinks that in his letter to Herz, Kant describes our intellect as an intellectus ectypus (1998, p. 19). Note that in the Critique of Practical Reason, Akademie Ausgabe V 43, Kant also uses the terms "archetype" and "ectype" in order to describe the relation of our practical intellect to its object.

[4] Cousin misses this point when he accuses Austin of appealing to Platonic universals (1955, p. 74-5). As Austin emphasizes in A Plea for Excuses, models such as the one he uses in How To Talk inevitably obscure important differences (1979, p. 202). For a model similar to the one described in How to Talk, see Truth (1979, p. 121-2).

several steps remote from Kant's. After outlining its original shape, I extend Austin's model, thereby gradually changing the context and consequently the meaning of the terminology that acquires its meaning within this context. Each such step will be a step away from Austin toward Kant. In order to fruitfully apply Austin's ideas to Kant, it will be important that each of these steps is taken with sufficient care and circumspection.

Austin proposes to capture the distinction between calling and describing by considering a simple world that contains a number of labeled items, each of which are of exactly one type. The language that he introduces within his model only allows for sentences of the form

The item labeled 'A' is of type T.

These sentences, he goes on to explain, can be used in four different ways, which are best explained by distinguishing different kinds of questions to which they may serve as an answer. In one set of cases, we imagine that someone is confronted with an item, say the one labeled A, and is asked one of the following two questions:

(1) (a) Is A of type T?
    (b) Is type T the type of which A is an instance?

The first question (1a) is a question about item A, which is given; the second (1b) is a question about type T, which is not given. In a second set of cases, someone is provided with a type description and then asked to identify an item that is an instance of this type. This kind of question may also take two different forms:

(2) (a) Is any of the available items an instance of type T?
    (b) Is type T instantiated by any of the available items?

Again, question (2a) is about any one of the available items, which is not given, and (2b) is about the type that is given. In all four cases, the answer may have the form "A is of type T," but in each case this statement has a slightly different meaning. The differences between (a) and (b) are admittedly subtle ones. It is clear that in any case, stating whether A is of type T involves two capacities: to identify a particular item as an instance of a type, and to understand what it takes, in general, to be an instance of this type. Austin's point is that these two capacities are to some extent independent of one another. One of them engages knowledge about a particular item, whereas the other engages knowledge about a general type. For instance, one may be able to correctly describe an item A without knowing of a given type whether A falls under it, or one might be able to state the criteria for instantiating a type T without being able to identify some particular item that satisfies them. Questions (1a) and (2a) assume that knowledge about the type T is already available, and the task is to apply this knowledge to a particular item. In (1b) and (2b), it is assumed that the features of the relevant items are already known, and the task is to decide which type would fit the bill. The distinction between (a) and (b), which Austin describes in terms of the "onus of match," is independent of the question whether the item (1) or the type (2) is

given.[5]

I am here interested in the contrast concerning what is given and what is asked for, i.e., the distinction between (1) and (2).[6] This distinction is the one that Austin draws in terms of the "direction of fit." In cases (1a) and (1b), the task is to come up with a type that fits a given item, such that the direction of fit is type to item (read: type to fit item). In cases (2a) and (2b), the task is to identify an item that fits the type, and the direction of fit is item to type. For my present purposes, I consider only two of the four ways in which the sentence "Item A is of type T" may be used in Austin's model:

(1a) <u>Stating</u>: Given an item A, to answer the question about this item as to what its type is.

(2a) <u>Casting</u>: Given a type T, to answer the question about any of the available items as to whether they fall under it.

In an act of stating, the question that is answered is a question about what is given (the item), and the direction of fit is type to item. Casting provides an answer to a question about what is not given (which is, again, the item), and the direction of fit is

---

[5] It is surprising that of all the distinctions that Austin introduces in How to Talk, the one between onus and direction has been misunderstood by so many readers. He devotes by far the most space to it. Some simply confuse the onus of match with the direction of fit, as, for instance, Price when he writes that the distinction in terms of onus of match may be used in order to explain the difference between beliefs and desires (1979, p. 353). Others have denied that a clear distinction between onus of match and direction of fit can be drawn. Cox, for instance, thinks that Austin describes a situation where someone is given an item and is supposed to "hold fast to" a type that is not given (1955, p. 9). This does not seem to make sense. However, the phrase "hold fast to" is not one that Austin uses. Correct accounts are given by Warnock (1989, p. 48-50) and Chisholm (1969, p. 102-4).

[6] The "onus of match" will reappear in Chapter 6, section 6.

item to type.

It should be clear that this distinction has nothing to do with a distinction between beliefs and desires, or assertions and expressions of intention. In casting, we do not express a desire that A be of type T or an intention to bring this about. When Austin speaks of "producing" an item that fits a description (1979, p. 141), he does not literally mean that such an item is made. He uses the verb "produce" in the sense of "coming up with something."

3.   Sellarsian Sentences

Let me now extend Austin's model in a first step, by taking into account that items may not only be classified as instances of certain types but also furnished with properties. In order to do this, we need to consider a world in which items are of types and have properties, and introduce more complex sentences of the following form:

Item A of type T is P.

These more complex sentences may be broken down into two parts, one of which is exactly of the form Austin describes:

Item A is of type T, — Item A has property P.[7]

I treat property attributions as analogous to type attributions. That is, by uttering a sentence of the form "Item A has property P," we may state that the item has property P or cast an item as having property P. Since there are four different ways of relating types to items and four ways of relating properties to items, sentences of the form "Item A of type T is P" may in theory be understood in sixteen different ways (some of which may not make sense). I consider only the following case:

---

[7] Austin does not break down his statements into further parts (1979, p. 139-40).

Item A, which is cast as a T, is stated to be P.

To cast an item as a T is to single out and present this item as an answer to the question whether it falls under a given type T. To state that it is P is to say what it is like. In this first extension of Austin's model, casting and stating do not any longer refer to self-contained speech acts but to parts of speech acts. I speak of the casting part (or term) and the stating part (or term) of an assertion. As in propositional logic, it seems most appropriate to classify the sentences under consideration by reference to their main logical operation. That is, just as "(p or q) and r" is a conjunctive and not a disjunctive proposition, "A, which is cast as a T, is P" is a statement and not a casting (castment?). Still, it has a casting part.

It is now time to gradually return to Kant. On my way back, I will further modify Austin's model and drop the assumption that all items are labeled. When we cast an item as a T, we do not answer the question whether an item labeled "A" fits type T, but rather the question whether a so far unspecified item fits this type. We do not cast item A as a T; we simply cast a T. Dropping the labels is a decisive step away from Austin's model. In Austin's original model, items are accessible independently of how they are cast, namely by means of their labels, and acts of casting them are acts of bringing these labeled items under a given concept. There is no question how we identify items in the first place. By dropping the label, I turn the casting part of the kind of sentence under consideration into that part of it by which the object is first brought into view. This is in some respects an even more artificial situation than the one originally considered by Austin. In ordinary discourse, we perform acts of pure casting only rarely; in most cases, we refer to items that are already known elements of our universe of discourse. For the time being, however, I am not concerned with the structure of average statements but with the structure of such exceptional ones, where an object is first brought into view by casting it as a T and then stated to have a certain property P.

The sentences under consideration have the form:

   This T (<u>casting</u>) is P (<u>stating</u>).

Since I study this rare specimen throughout the rest of this chapter, I might as well give it a name. I call it a <u>Sellarsian sentence</u> because it has the general form

   This such is so-and-so. (cf. Sellars 1976, §10)

By speaking of Sellarsian sentences, I do not want to give the impression that I buy all or even only the essential claims that Sellars makes about Kant. For instance, Sellars treats the first part of his sentences ("this such") as mentalese representations of Kantian intuitions. I argue that in casting, we do not establish an immediate relation to an object, and since intuitions are immediate relations to objects (B 33), the casting parts of Sellarsian sentences do not correspond to intuitions.

This is not the only disclaimer I should make before going on. I should also emphasize that I do not wish to claim that all assertions (let alone all thoughts) have the form of a Sellarsian sentence. For instance, one may cast an item as having a property instead of casting it as being of a type, and then one may state of it that it instantiates type T (as in "This furry item is a beaver"). Further, when I propose to discuss cognition by considering Sellarsian sentences, I am neither claiming nor denying that all subjects of cognition must be capable of language. I am still arguing within a simplified model in which it is simply assumed that all cognition is or may be expressed by way of uttering sentences of the form "This such is so-and-so."

4.   Affection and Function

According to Kant, judgments are either complex and consist of further judgments or they are categorical. All categorical judgments involve two predicates that perform

different functions. In Reflexion 4634, Kant writes that in a (singular) categorical judgment of the form "S is P" we compare two predicates, one of which stands for a logical subject, the other for a logical predicate (Akademie Ausgabe XVII, p. 616-7). He adds, in Reflexion 4676 (p. 654), that when a concept A stands for a thing so that it is impossible to think of the same thing by its negation, it applies substantively to this thing and is not a (logical) predicate. For instance, whereas one may think of an educated human as someone who is not at all educated, one cannot think of a human as something that is not at all human.[8] Therefore, "educated" is a logical predicate, where "human" is a substance concept. The reason why "human" could not possibly be affirmed and denied of the same thing is that the concept and its negation would not refer to the same thing. Substance concepts serve to identify objects, and different substance concepts serve to identify different objects.

In a categorical judgment of the form "S is P," an object is introduced by means of a substance concept, which is then related to a logical predicate that stands for a property. Since the substance concept introduces the object that is then determined by the predicate, Kant also says that it provides the ground of the act of judging, of which the logical predicate is, in some sense, a consequence (Logik Jäsche, Akademie Ausgabe IX, p. 104). Categorical judgments thus have a structure that is to some extent analogous to the structure of hypothetical judgments ("If A, then B").

To be sure, the parts of a categorical judgment are not further judgments, and the relation between a substance and its attributes is not the same as the one between the antecedent and the consequent in a hypothetical judgment. Kant writes that categorical judgments cannot be transformed into hypothetical judgments because in a hypothetical judgment the condition is not presupposed or stated but merely problematic, whereas in a categorical judgment its truth is at least presupposed (Logik Jäsche, p. 105). However, it should still be possible to transform categorical judgments into judgments of the form "Because this is T, it is P," where the

_____

[8] When one imagines a human that has been turned into a frog, one imagines a human with the shape of a frog, not a human that is not at all a human.

antecedent is not problematic. On the other hand, there also seem to be categorical propositions in which the substance term is problematic, such as Descartes' judgment that even an evil demon could not deceive him. Hence, categorical and hypothetical judgments are sufficiently analogous for making the following move. Just as the hypothetical judgment "Since this is a T, it is P" may be split up into the two judgments "This is a T" and "This is P," the singular categorical judgment "this T is P" may be split up into the two "protojudgments"

    this T — is P.

Kant says that the first of these parts provides a ground, of which the second is a consequence. Protojudgments are not judgments, but in order to see more clearly what is going on, one may as well expand them to actual judgments:

    This is a T — it is P.

The first of these judgments will now be easily recognized as an instance of casting, the second as a case of stating. By casting an object, we lay the ground on the basis of which we may state something of this object. When Kant says that this is what happens in general in a categorical judgment, he seems to assume that what I have called a Sellarsian sentence is the general form of a (singular) categorical judgment. To make such a judgment is to state something of an item that is cast as an instance of a type.

Let me now fill in some epistemological background. Kant distinguishes two capacities involved in cognition: sensibility and understanding (Critique of Pure Reason, B 74). Actualizations of our sensibility are passive affections. By means of affection, an object is given to us. Acts of the understanding, which Kant calls functions, are active: They determine the form of their object (B 93).

Affection relates to intuition, and functions relate to concepts. As Kant uses these

terms, "intuition" and "concept" may both be understood in at least three different ways. First, they may refer to capacities of actively or passively relating to objects. Second, since capacities must be described in terms of their actualization and acts are always actualizations of capacities, Kant also uses these terms for the acts themselves. Finally, they may stand for that which is achieved in such acts; that is, the relatings that are achieved or the relations that are established. In this latter sense, an intuition is an instance of intuiting, and a concept is an instance of conceiving. Intuitions and concepts may thus be taken to be the internal results of the acts of intuiting and conceiving, in the sense in which a dance is the internal result of dancing it. Kant uses the term in this third sense when he says that our sensibility supplies us with intuitions (B 33) and that intuitions rest on affections, whereas concepts rest on functions (B 93). I straighten out Kant's terminology by using "sensibility" and "understanding" for the respective capacities, "affection" and "function" for the acts, and "concept" and "intuition" for their (internal) results.

Since sensibility is a passive capacity and understanding an active one, it seems appropriate to distinguish the respective acts, affection and function, in terms of their direction of fit. Functions are analogues of castings, which determine their objects, and affections are analogous to statings, which register features of their objects. For affections, the direction of fit will then be intuition to object, and for functions, the direction of fit is object to concept. Note that here, the internal result of relating to an object (intuition and concept) is one of the relata. It may be said to relate to the object in the same sense in which a path relates to what it leads to. A path does not only relate to what it leads to but also constitutes or establishes this relation. In the same way, intuitions and concepts relate to objects by being the internal results of relating to objects.

Before going on, I should emphasize that any such sharp division of cognition into affections and functions is an artificial one (an ens rationis). Considered in isolation, neither affections nor functions actually relate to any object. No relation to any object can be established unless concepts and intuitions come together (B 74). This means

that in the absence of all spontaneous synthesis, there can be no intuitions (A 120 n., B 136f.). We must therefore be careful when we match Kant's terminology on Austin's distinction. It is important to keep in mind that the whole point of Kant's distinction between active and passive elements in cognition is that in order for cognition to be possible, both must come together. Mere affections are blind, and mere functions are empty (B 75). This means that there are no cognitive acts that have only one direction of fit. Taken in isolation, functions are not acts of casting, because functions alone do not relate to objects, and to cast an object is a way of relating to it. Likewise, taken in isolation, affections are not acts of stating something of something, since they do not relate to any object, and stating something of an object is a way of relating to it. If functions and affections differ in their direction of fit and both must come together in order to yield cognition, every cognitive act must have both directions of fit at once. As Engstrom puts it, our cognition does not stand in a "single unidirectional direction of fit with its object" (2002, p. 53).

A second difficulty is the following. When applying Austin's distinction to Kant, one might be tempted to say that functions are acts of determining objects by using concepts, whereas affections are acts where an intuition is determined by its object. However, Kant never says that objects determine intuitions. He says, more precisely, that in cognition we relate concepts to objects by determining intuitions (B 74, 128, 304, and passim). Thus, if affections and functions are relatings to objects, they are not simply the converse of one another. Only one of them involves determination; the other one does not.

Since neither affections nor functions alone relate to objects, the only way to distinguish between an active and a passive element in cognition is to begin with a complex act of successfully relating to an object, in which both concepts and intuitions must be involved. Within such a complex act, one may then single out two elements: functions, which are actualizations of the understanding that result in concepts, and affections, which are actualizations of our sensibility resulting in intuitions. These parts may be said to differ in their direction of fit.

Applying Austin's terminology to Kant leads to an answer to the question raised in the beginning of this chapter: How can our cognition determine its object without in any sense creating or manipulating it? When Austin speaks of an item to type direction of fit, he has in mind the following case: A type description is given, and the task is to find (not to make) an object that fits this description. I submit that when Kant says that objects must "conform to our cognition" (B XVI), he basically means the same. He does not say that objects must conform to our cognition because they are made to conform to it. He says that they must conform to our cognition because they are cast as conforming to it. In casting, a type term is given and an item is identified as an instance of this type. It conforms to the type not because it is made or created according to it, but because it is chosen under the condition that it would.

As a test case for applying Austin's terminology to Kant, let us ask the following question: Why does Kant claim, in B 377, that intuitions immediately relate to an object, whereas concepts do so only mediately? He writes that concepts "can never relate immediately to an object but only to some other representation of it" (B 93). This, however, cannot mean that concepts only represent representations of objects. For one thing, if a concept would represent a representation, this representation would be its object, and it would immediately relate to it. For another, Kant says that concepts refer to an object by relating to another representation of the same object, and this seems to imply that concepts are also representations of the object to which they mediately refer. Therefore, concepts are not representations of representations. Kant can only mean that although concepts do represent objects, they do so only through other representations (which do not represent the objects through further representations).

Austin provides the resources for understanding why this must be so. As parts of cognition, functions may be described as relatings to objects with an object to concept direction of fit: The object is supposed to fit a given concept. As I have emphasized, this cannot mean that by a function, our understanding actually shapes its object. What I want to suggest now is that functions can only mediately relate to their object

<u>because</u> they do not shape it.

Let me, for the sake of illustration, replace the distinction between function and affection with a more familiar one: the one between doing and getting something. Like a function, doing is active; like an affection, getting is passive. If there were a distinct and merely passive element in cognition, it would be something like a mere getting. A mere doing, on the other hand, would not be an instance of cognition but an action. Therefore, our spontaneous cognitive acts cannot be described as mere doings. Rather, they are instances of <u>doing something in order to get something</u> (i.e., of fetching something). Successfully doing something in order to get something (or fetching something) is not the same as simply doing or simply getting something. For instance, I may do something in order to get money but not get the money; and I might get money without doing anything in order to get it.

Castings are instances of doing something in order to get something. When we cast an object, we throw a concept at the world, as it were, and if we are lucky, we get something in return. If we do not get anything in return, our casting is a mere doing without a getting, but it will also not be successful. This means that whenever we successfully cast an item, our relation to what we get is in an obvious way not immediate. It is necessarily mediated by a mere getting. The success of casting depends on something that is not a doing but a getting, and in this sense, casting is mediated. The same is true for our spontaneous cognitive acts. As Kant puts it, we think an object in relation to a receptive determination of our mind (B 74). Here, the thinking of the object corresponds to the doing, and the receptive determination to the getting. Both of them, taken together, constitute a sort of fetching (an instance of doing something in order to get something). If functions are analogous to castings, they must be mediated by affection, which is a mere getting. Affection, in contrast, need not be mediated by anything further. On the assumption that functions are acts of

casting, we can thus explain why concepts can only mediately relate to objects.[9] On the other hand, if the objects of our cognition were created by our spontaneous acts, like the world is created by a divine intellect, the objects would directly relate to our spontaneous acts. They would offer no resistance, and we would need no further means for relating to them. There would be nothing in between our acts and their objects; our relation to them would be immediate. Because we do not create or shape the objects of our spontaneous cognitive acts, we do not immediately relate to them.

5. A Priori Knowledge

My account of functions as castings may be used to explain Kant's notion of synthetic a priori knowledge. I show this in two steps. First, I introduce a notion of relative a priori knowledge that is admittedly not Kant's. Then I ask how to modify this conception so that it turns into the Kantian one. As we have seen, Kant describes the substance predicate in a categorical judgment as a condition under which the statement is made. It provides the ground for attributing a property because in a Sellarsian sentence, casting is prior to stating: Before we can state anything, we need to get hold of something to state it of. The first step towards an account of a priori knowledge is to realize that what is known a priori may simply be what is prior. Consider, for instance, the Sellarsian sentence

This ice cube is pink.

When we divide this judgment into its protojudgmental parts and transform these

_____

[9] Sellars (1968, p. 3) and Haag (2007, p. 167 and 313), suggest that intuitions literally are those parts of a judgment by which we refer to a singular subject of predication ("mentalese demonstratives"). As has been noted, my analysis makes clear that the casting part of a Sellarsian sentence ("this such") cannot be an intuition. Whereas intuitions are immediate relations to objects, casting is a mediate relation to an object.

parts into judgments, we get the two judgments

This is an ice cube — it is pink.

The first is not stated in the original judgment but presupposed; before predicating pinkness, we cast something as an ice cube. The direction of fit of the casting part of a Sellarsian sentence is object to type; that is, an object is picked out by virtue of its falling under a type. We pick an object in its capacity of being an ice cube. Now it is important to note that in most cases in which we successfully pick out an object by virtue of its falling under a type, we can know in advance that it will indeed fall under this type. If it did not fall under it, we would probably not have been able to pick it out by virtue of its doing so. That is, whenever we successfully cast something as an ice cube, and do not miscast it as something else, we know a priori that it is an ice cube.

Kripke writes something along these lines. He says that someone may know a priori what Aristotle did if he "fixes the reference of 'Aristotle' as the man who did one of these things" (1980, p. 63). Against this, Donnellan objects that stipulations as to what a name means cannot constitute real knowledge other than in linguistic matters (1977, p. 55). Where an object is cast as satisfying a certain description, real knowledge would have to go beyond the mere fact that it satisfies this description.

So this had better not be my last word on a priori knowledge. Let me nonetheless emphasize already at this point that nothing I have said so far implies that all a priori knowledge is analytic. One might think that in the case under consideration we know the features of the object a priori because we know the meaning of the term "ice cube." But casting an object as an ice cube may involve either more or less than understanding the meaning of this term. For instance, it may involve locating a solid object, by sight or touch, in space and time. We may therefore a priori know certain features of an item that we cast by considering the act of casting it, rather than only considering the meaning of the words that are employed in this act. If the meaning of

the casting term does not fully determine its referent, but only the entire act of casting it does, synthetic a priori knowledge is possible.

Still, this is certainly not the way in which Kant wishes to use the expression "a priori." When we successfully cast something as an ice cube, we know a priori only in a relative sense of "a priori" that it is an ice cube. Kant mentions this relative sense only to contrast it to the notion of absolute a priori knowledge he is interested in (B 2).[10] However, if the distinction between a priori knowledge as I have explained it and a priori knowledge in Kant's sense is the distinction between relative and absolute a priori knowledge, then there must be a way that leads from one to the other. Relative a priori knowledge is a priori knowledge that we have relative to a specific context, as when we cast an item as an ice cube. Absolute a priori knowledge is knowledge that we can have, for the same reasons, in any context where we cast an object.

We may extract the difference between relative and absolute a priori knowledge from §14 of the Deduction. In this section, Kant draws the by now familiar contrast between representations that are caused by their object and representations that determine their object. He claims that in the first case no a priori knowledge is possible. Then he continues:

> In the latter case, representation in itself does not produce its object insofar as <u>existence</u> is concerned, for we are not here speaking of its causality by means of the will. None the less the representation is <u>a priori</u> determinant of its object, if it be the case that only through the representation is it possible to know anything as an object. (B 125, tr. Kemp Smith)

Therefore, in order to see how there can be a priori knowledge in an absolute

---

[10] In B 273, Kant also says that we may know of a thing's existence comparatively a priori when we infer it from the observation of other things. This is a case of relative a priori, since we know that the thing exists before we have experienced it. We may not know it prior to all experience.

sense, we need to ask whether there are specific concepts or capacities that are necessarily involved whenever we cast an object. I have argued that when Kant says that an object conforms to our cognition, he cannot mean that our cognition creates or in any sense manipulates its object. If our cognition would determine the form of its object by manipulating it, it would be practical knowledge. Using the distinction between different directions of fit that I have extracted from Austin, we can see how one may fit an object to a given representation without manipulating it in any way.[11] Representations determine an object when they are used in order to cast this object. Even if an object that is cast is not thereby changed, we may still know some of its features by merely considering the act by which it was cast. Whenever we successfully use a concept in order to pick out an object, we have good a priori reasons for assuming that the concept applies to this object. If it would not apply, we would probably not have been able to pick it out by using the concept, and our act of casting would not have been successful. To be sure, mistakes are possible. In a mistake, one mistakes something for something else. I may cast a tomato as an apple and then discover that the concept by which I cast the object does not apply to it after all. As far as Kant is concerned, it is important to note that this kind of mistake is <u>not</u> possible in the case of concepts that are necessarily involved in all acts of casting objects. According to Kant, all acts of casting objects of experience must involve pure forms of intuition, categories, schemata, and principles. If a proof can be given that such resources are necessarily involved in casting any object whatsoever, the only surprise that can happen when using these resources can be that after all, there was no object whatsoever. This may of course still happen (although strictly speaking, it would not be a mistake). In casting, we depend on there being something to be cast. If there is nothing to be cast, our act of casting is unsuccessful and does not provide us with knowledge of any object. However, if there is nothing that we might cast by

---

[11] We can see how this is possible. This does not yet imply that in fact, we do not inevitably and systematically alter the object of our knowledge by knowing it. Whether we do this or not is obviously not something we can know.

carrying out our most fundamental casting procedures, cognition is altogether impossible. Conversely, if cognition is at all possible, a reflection on the way in which we cast its objects may tell us something about what they are. In B 681-2, Kant considers the possibility that although the pure forms of intuition and understanding apply, there might be no further (that is, material) unity among the objects of experience, such that it would be impossible to apply general empirical concepts. He argues that under such circumstances no understanding would take place, and concludes that for an actual use of understanding to be possible we must suppose that there are things to which at least some general empirical concepts do in fact apply. However, that there might be no item at all to be cast is not a problem for Kant. He would presumably not claim that we know a priori that there are objects; the only thing we know a priori is that if we can have any knowledge of any object, it must be possible to cast one, and that as a consequence of this, the object of our cognition must have certain fundamental features.

We should not expect to get to know these fundamental features by only considering the casting parts of Sellarsian sentences. Outside our simplified model, objects may be cast in all kinds of ways: by locating them in space and time, by pointing or even only looking at them, by taking up anaphoric pronouns, by using proper names, by recalling something from memory, and so on. We cannot expect to find out what is necessarily involved in all acts of casting objects by only considering one kind of simple assertive statements. For the time being, however, I am only pointing in a direction where one should look for a priori knowledge; I need not go there and find it. I have said enough to conclude that a priori knowledge is knowledge that can be derived from a reflection on what it takes to cast an object in general. By using Austin's notion of casting, I am thus providing a more specific version of Allison's general (and, I suppose, largely uncontroversial) thesis: that a priori knowledge rests on the fact that "objects must conform to the conditions under which we can alone represent them to ourselves as objects" (2004, p. 37). The conditions in question are the conditions under which something may in general be cast as an object.

6. Aristotle's Four Causes

The main point of explaining a notion used by a philosopher is to be able to use it, possibly beyond or against this philosopher. I believe that my account of a priori knowledge, rudimentary as it is, opens up at least two ways that may lead beyond Kant. First, if a priori knowledge is knowledge derived from a reflection on what is necessarily involved in casting items, it is possible for it to change. After all, Kant tells us only what is necessarily involved in casting items as objects of experience. I do not know whether any sense can be made of the idea of casting an item other than as an object of experience, but if this could be done, we might learn to cast objects by using concepts other than those that necessarily apply to empirical objects. It might be that science opens up a way of casting items as something other than objects of experience, as Sellars seems to hope (1968, p. 143-50). This does not seem to be conceptually impossible, although one should not infer from this that objects of experience do not really exist (cf. McDowell 1998, p. 469-70).

Second, there may be such a thing as a regional a priori; in something like the sense in which Husserl distinguishes between general and regional ontologies (Ideen, Husserliana vol. 3/1, §9). It might well be that in casting specific kinds of objects (physical objects, living beings, processes, persons, mathematical entities, etc.), different sets of specific but in the respective area still fundamental concepts need to be applied. These concepts might be the subject matter of a special, rather than general, transcendental logic (cf. B 76). Kant's categories may be the concepts by which we cast objects in general, if indeed he is right; but there might also be more specific sets of casting terms by means of which we cast living beings as opposed to mere physical objects, or persons as opposed to mere living beings.

As I have indicated in the beginning of this chapter, this opens up a way of understanding how Aristotle's four causes apply a priori to all natural things, even though they need not apply to all objects of experience. Aristotle does not pick his causes up just as they occur to him. He introduces them by emphasizing that in order to understand natural things and processes we must approach them in a certain way.

We must ask four questions about them in order to see them as the natural things and processes they are, rather than as something else. That is, Aristotle's four causes play an indispensable role in casting natural things and processes as such. If Aristotle is right, natural things must be cast as material things with a paradigmatic form, and natural processes must be cast as processes that proceed toward an end. Further, natural processes must be cast as involving natural things, and natural things as being involved in natural processes. This is not something we find out by observing natural processes; it is something that we need to take for granted in order to bring them into view in the first place.

Explaining a notion found in Kant by reference to Austin, and then applying this notion to Aristotle, is a bold move. It employs conceptual resources that were not available to Aristotle. By making this move, I do not wish to give the impression that Aristotle explicitly thought of the four causes as fundamental casting terms. Still less do I wish to give the impression that Aristotle was, after all, some kind of German Idealist, or a speech act theorist. I am here not interested in establishing general philosophical connections between Aristotle, Austin, and Kant. I am only interested in the distinctions that I extract from Austin, between the two directions of fit, and between casting and stating. I believe that these distinctions, wherever they come from, help explain what Aristotle wants to say. In Physics I 2, Aristotle emphasizes that if one starts from the wrong principles one will not be able to study nature. I find it helpful to put this the following way: Nature can only be studied as such when we cast natural things in a certain way.

In order to see natural things as the natural things they are, we must ask the questions about them to which the four causes are answers. Otherwise, we fail to see their naturalness. Further, there is an obvious sense in which the way in which the questions that we ask about an object are prior to their answers. Questions determine a range of possible answers before any actual answer is given. If nature only appears as such when certain questions are asked, we can know something substantial about nature by reflecting about these questions and the range of possible answers they

admit. The knowledge that we can have in this way is synthetic knowledge a priori.

Something very much like this has been emphasized by Heidegger. In Physics I 2, Aristotle writes that it is clear ἐκ τῆς ἐπαγωγῆς that natural things are subject to change (185a12-14). Heidegger insists that here, ἐπαγωγή cannot mean induction. He writes that rather than inferring from an observation of many cases that natural things change, we know it by looking beyond them and thereby grasping the principles in terms of which each of them can be seen as a natural thing. Failure to address natural things as subject to change leads to what Heidegger calls "φύσις-blindness" (Gesamtausgabe 9, p. 264).

Wieland also says something closely related about Aristotle's four causes. Taking up a suggestion by Brandis, he calls them "concepts of reflection." Brandis writes that such concepts form a system and that they "invariably apply to all determinations of concepts and investigations" (1853, p. 370). Wieland emphasizes that concepts of reflection are "conspicuously void of content," so that they do not refer to metaphysical entities (1992, p. 186-7, 211). They are not concepts with specific content but "points of view that one may take for constructing or finding such concepts more easily" (p. 202-3).

The term "concept of reflection" is Kant's, and the context in which he introduces it is somewhat obscure. In general, reflection is the act of comparing representations to other representations or to their sources (First Introduction to the Critique of Judgment, Akademie-Ausgabe XX, p. 211; cf. Reflexion 5051, XX, p. 211). Now as Longuenesse sees it, all cognition of an object involves a comparison of representations (1998, ch. 6). This, however, does not mean that conversely all comparison of representations leads to the cognition of a further object. Kant's categories correspond to ways in which representations can be related to each other such that they acquire objective content and represent objects. Concepts of reflection, in contrast, correspond to ways of comparing representations that already represent real objects to each other. Such comparison does not yield representations of further objects, which explains why the concepts of reflection are not listed as a species of

representation in B 376-7 (Malter 1982, p. 130). They do not represent their own objects, over and above the objects of the compared representations.

If Aristotle's causes are concepts of reflection in Kant's sense of this term, they correspond to ways in which representations of empirical objects can be compared to each other. When we reflect on them, we take different representations—of the same or of different objects—and compare them in terms of their objects. Concepts of reflection come in pairs, and one of the pairs that Kant discusses is the contrast between a determinable item and its determination. He also refers to this as the relation of matter and form (B 317, 322-4). For instance, we can take a representation of the silver that makes up a cup, and a representation of the cup that consists of this silver, and compare them to each other as determinable (silver) to determination (the cup). According to Kant, this comparison does not represent a further object, such as, presumably, the relation of material constitution. Similarly, we can take a representation of a process as biting chips of wood off the trunk of a tree, and a representation of the same process as felling a tree. These two representations also relate to one another as determinable to determination, and again, their comparison does not yield a representation of a further object, which would presumably be a causal or teleological relation. Although concepts of reflection differ from categories, they may also be employed prior to experience. We can use them in order to reflect on the ways in which we cast objects as instances of certain general types. If Aristotle's causes are concepts of reflection, and if they must be used in order to cast natural things and processes as such, they apply to such things a priori, just as the pure forms of intuition, the categories, the schemata, and the principles do. The only difference is that they do not correspond to further objects.

As I emphasize in Chapter 9, section 4, this does not mean that the causes are not real. When we compare two representations to each other by using the contrast between matter and form, this comparison does not represent a third object. The objects of the two compared representations, however, remain as real as they were before comparing them. Matter and form are real objects; the relation between matter

and form is not a real object.

Τόδε, τι, and τοιόνδε

The material cause of a natural thing is that out of which it comes to be as a result of a natural process. It is something that potentially is this natural thing, and its potential of being this thing persists in the result of the thing's coming to be. Matter (ὕλη) is only one kind of material cause. When material things come to be or change their form, they come to be out of matter. When they exchange their matter, they do not come to be out of matter. This chapter and the following one are about the question of what matter is. After briefly considering various attempts at saying what it is to be the matter of a material thing, I focus on an idea that Aristotle articulates in Metaphysics Z 3: that matter is what remains when we take away all form. I raise the question how it is possible to take away all forms from a thing, and argue that one may do so by taking a complex noun phrase of the form "this such" (τόδε τι) and deleting all parts of it that have any kind of descriptive content. This leaves us with a mere "this" (τόδε), which may be taken to refer to the matter of a thing. I suggest the formula (τόδε + τι) + τοιόνδε in order to bring out the distinction between the matter of a thing, its paradigmatic form, and its properties. I do not mean to suggest that τόδε always refers to the matter of a thing, or that τι and τοιόνδε always refer to types and properties, respectively. Still, the formula (τόδε + τι) + τοιόνδε is useful for drawing certain important distinctions. In section 5, I follow up a suggestion made by Charlton, that the concept of matter as an ultimate subject of predication may be traced back to Plato's Timaeus.

1. What is Matter?

Like other what-questions, the question "What is matter?" can be answered in two different ways. First, one may list instances of matter. Matter is what silver is for a cup, wood for a beaver dam, blood for an embryo, and so on. Second, one might pin down what all such instances have in common—that is, define what it is for them to

be matter, if this is indeed what they are. These two kinds of answers are not always easily kept apart. For instance, when physicists claim that matter ultimately consists in quantum-mechanic particles or fields (or whatever they claim in detail), it is not clear whether they provide a definition of what it is to be matter, or rather describe an instance of matter. On the one hand, they seem to do the same as we do when we say that matter is what silver is for a cup. They say that the things in our actual universe happen to consist of certain elementary particles or fields. They still give instances of matter, only fewer of them and more generally specified ones. On the other hand, a definition of matter is nothing but the most general specification of all instances of matter. When we state the definition of matter, we provide a description under which all instances of matter fall, and thus we describe an arbitrary instance of matter.

In any case, however, when physicists tell us what matter is, they answer an empirical question. They do not define a concept of matter that would apply in all possible circumstances; they only tell us what satisfies the definition of matter in those circumstances that contingently hold in our physical world. Further, to define matter as what physicists find out to be matter would turn the claim that it is indeed matter into a logical truth. But it is not a logical truth. In general, in order to evaluate an empirical claim to the effect that X is matter, "X" and "matter" must be taken to have different meanings. Therefore, we can only evaluate the claim that matter consists in particles or fields when we assume that "matter" and "fundamental particles" or "fields" (or whatever scientists tell us matter is) do not mean the same.

In this and the following chapter, I am not interested in the question of what things in this world actually consist of. The more general systematic question that they addresses is rather what it is to be matter, for whatever anyone may come up with as the actual matter of physical things. I should also make clear in the beginning that I only ask what it is to be the matter of another thing. I am not interested in the concept of absolute or prime matter but only in the concept of relative matter: matter-of-something. Prime matter may probably be defined, in a second step that I do not take, as matter-of-everything. Further, the question I am raising here should not be

confused with a question about the most basic elements of concrete things. An element is something into which a thing may be decomposed or analyzed, which may itself not be decomposed or analyzed into further constituents of a different type (De Caelo III 3, 302a15-18). Matter need not satisfy the latter condition, and hence, the concept of matter is more general.

Now what is it to be the matter of a thing? In Physics I 9, Aristotle seems to suggest that when a thing comes to be or undergoes a change, matter is what underlies this change and persists as the same. However, as I argue in Chapter 1, and as Aristotle says in Physics I 7 (190b8-9), natural things may also exchange their matter, and when they do so their matter is not what persists throughout these changes. In fact, Aristotle identifies the underlying subject of change with matter only in cases where a thing comes to exist (Physics I 9, 192a31-32). This also means that the matter of a thing cannot in general be the matter out of which it originally came to be. For, when a thing exchanges its matter, the matter out of which it came to be ceases to be its matter.

In Metaphysics Z 8, Aristotle writes that what makes Callias and Socrates different is not their form but their different matter (1034a5-8). It is therefore sometimes assumed that matter is that which makes the difference between two instances of the same form. However, in order to make such a difference, the respective parcels of matter should themselves be different instances of the same form, and thus by assumption consist of further matter (cf. Charlton 1972, p. 247). This leads to a regress.

Finally, since matter is the complement of form, one might think that the matter of a material thing must be what is left over when one takes away its form. Aristotle seems to suggest as much in Metaphysics Z 3. In the following, I concentrate on this latter attempt to say, in general, what it is to be matter.

2. The Pale and the Dead Socrates

Aristotle introduces the idea of taking away all attributes of a given thing in

Metaphysics Ζ 3 in order to show that at least in one sense of "substratum," the essence of a substance cannot be taken to be its substratum: the <u>material</u> substratum of a thing is not its essence.

This is the story, as it is usually told, about why Aristotle raises the issue. The most basic question that we can ask about a thing is what it is; less basic are the questions of how it looks and feels, where it is located, how it relates to other things, etc. In Categories 5, Aristotle claims that What-questions are properly answered by reference to substances and that in the primary sense, substances are the particular instances of a species, such as Socrates or a particular horse (2a11–14). In his physical writings, however, Aristotle comes to distinguish between the matter and the form of sensible substances. As a result, there seems to be something more basic out of which Socrates and the particular horse may be said to consist. Socrates once came into being out of some matter, and as everyone knows, he eventually ceased to be and a dead body was left behind. Coming and ceasing to be are changes, and as Aristotle argues in Physics I 7, the proper account of any change involves three principles: that which results from the change, that which changes into that which results, and that which makes the difference between them. Consider, for instance, the change by which Socrates turns pale. The pale Socrates is that which results, the not yet pale Socrates with the potential for being pale is that which changes into the pale one, and the latter's paleness makes the difference. Now what is crucial is that the thing that changed into the pale Socrates must be specifiable as something that may or may not be pale (190a13-21). In the case of the pale Socrates this is easy, for we may refer to him by his name, whether he is pale or not.

In De Generatione et Corruptione I 3, Aristotle distinguishes qualified from unqualified coming to be. In the proper description of a case of qualified coming to be, the "be" is supplemented by a qualifier. When Socrates comes to be pale, this qualifier is "pale." Socrates does not simply come to be, period; he only comes to be pale. The same holds true for ceasing to be. If Socrates ceases to be pale, he does not cease to be, period. He ceases to be, period, only when he dies. However, Aristotle

adds that to a certain extent, the distinction between qualified and unqualified coming and ceasing to be concerns only the way in which we talk. For when Socrates ceases to be pale, we can also describe this change as an unqualified ceasing to be: The pale Socrates ceases to be, period (De Generatione et Corruptione I 4, 319b25–28).

This gives rise to the question of whether one may also describe the death of Socrates as a case of <u>qualified</u> ceasing to be (cf. Physics I 7, 190b1-3). When the pale Socrates unqualifiedly ceases to be, Socrates ceases to be pale. Now Socrates drinks the hemlock, tells Crito that he owes a cock to Asclepius, and ceases to be, period. But perhaps there is something of which we may properly say that it ceased to be something by changing from the living Socrates into his mortal remains. This something would have undergone a qualified ceasing to be when Socrates died: It ceased to be the living Socrates, but did not cease to be, period (cf. Gill 1989, p. 54). And indeed, there seems to be something that satisfies this description, at least when we focus on the immediate instant when Socrates dies. What may or may not be Socrates appears to be the organic matter that is left behind immediately after his death. Now if there is such a parcel of matter that may or may not be Socrates, it seems that one should also say that during his lifetime this matter was Socrates and that Socrates was not a substance but only an accidental compound.[1] When someone points at Socrates and asks what this is, we should say that it is an ensouled parcel of matter: a bunch of organic matter in a certain arrangement such as to be capable of certain movements. If this is true, however, "a human being" is not the most basic answer to the question of what Socrates is. The really basic subject is his matter.

If there is any such tension between the Categories and the Physics, it must have troubled Aristotle in Z 3, for here he argues that matter does not qualify as a substance, and that accordingly, the question of what a thing is may not be answered

---

[1] Cf. Lewis (1982) and Matthews (1990) for the distinction between substances and accidental compounds.

by reference to its matter. This is, at least, the common story.[2] In this chapter I argue that there is no real conflict between the picture of (primary) substances in the Categories and the doctrine of matter and form. In both contexts, Aristotle supposes that sensible substances are composite entities with a formal and a material aspect, and he denies that their material aspect may be considered in isolation. If Aristotle gets his point home, it is indeed misleading to say, pointing at Socrates, that he is a parcel of matter in a certain shape. It is misleading because the phrase τόδε τι, which Aristotle uses in order to refer to substances, does not have the logical structure of an attribution of a shape to a given thing.

3.   On Denuding

In Z 3, Aristotle proceeds in three steps. First, he argues that if we strip a physical thing of everything that is not a substance, we end up with matter, in a certain sense of "matter." Second, he shows that matter in this sense is neither separable nor a "this such." Third, since primary substances should be both separable and a "this such," Aristotle concludes that matter in the specified sense cannot be the primary substance we were looking for. In the course of this argument, he seems to assume that the matter of a thing is what is left over after the subtraction of all its forms and determinations. He writes:

> If matter is not a substance, it is hard to see what else could be; for when all else is taken away, nothing appears to remain. For—while other things are attributes, products, and capacities of bodies—, length, breadth, and depth are quantities and not substances (for quantity is not a substance). Rather, the substance is that primary thing to which these quantities belong. And yet when length, breadth, and depth are taken away, we see nothing remaining unless there be something which is determined by these. So on this view it must appear that matter alone is

---

[2] Cf., for instance, Dancy (1978, p. 383), Gill (1989, p. 16), Loux (1991, p. 50), and Bostock (1995, p. 72-4); cf. also Beere (2006, p. 308).

substance. (Metaphysics Z 3, 1029a10–19, tr. Bostock, modified)

Aristotle points out that when we take everything away from a thing that is not a substance, we must also take away its (determinate) spatial dimensions. That which is left must therefore be a subject of which these spatial dimensions are attributes. According to the view he discusses, this amounts to the claim that what is left over is matter.

The argument quoted above requires that we go on taking attributes away until we have reached the ultimate substratum of qualities.[3] However, there is no way of physically denuding a thing of all its properties since it seems that after each step we are left with something that has further properties. After all, Aristotle writes that the absence of a quality would also be a quality (1029a25-26). When we imagine a thing that lacks a certain quality, we therefore imagine a thing that has a contrary quality. There is strictly no way of even only imagining a thing without qualities.

Some suggest that instead of taking away the properties of a thing, we are only supposed to change them or imagine them changed. By changing the remains of Socrates regarding their shape, color, smell, etc., we would show that none of these qualities are essential to them, and since in every one of these changes from one quality to another, something must persist that may be one or the other, we will also show that there is something underlying the sum of all the changes that need not have any of the qualities in question (Bostock 1995, p. 78). However, this also does not seem to be a way to rid a thing of all its features. For consider how Aristotle describes the change that occurs when one of the elements—fire, air, water, earth—changes into another one (De Generatione et Corruptione II 1–4). All four elements, he tells us, may change into each other. When water changes into air, it ceases to be cold and becomes hot. When air changes into fire, it remains hot but turns dry and ceases to be

---

[3] (Pseudo) Alexander Aphrodisiensis calls this procedure "denuding" (ἀπογύμνωσις) in his Commentary to Aristotle's Metaphysics (CAG 1, p. 464,24). Cf. Dancy (1978, p. 395).

moist. One might thus expect that there is something underlying such a sequence of changes from water to fire that is in itself neither cold, nor hot, nor moist. But this would imply that there is something more basic than the elements, which would imply that fire, air, etc. are really not elements after all. If the elements are indeed elementary, there can be nothing else into which they could be decomposed. The only thing into which fire may be decomposed is fire, and this means that there is nothing in or underlying fire that would not be what fire is: hot and dry. The solution that Aristotle offers is that whenever something changes between moist and dry, it must remain either hot or cold, and whenever it changes between hot and cold, it must remain either moist or dry. Hence, although there is something that remains the same in each change, there is nothing that remains the same throughout all changes—except for a disjunctive property. Whenever we look at a changing element, it will necessarily have exactly two of the qualities that define the elements. At no instant will there be anything that is neither hot nor cold, or neither moist nor dry.

This means that there is no general way of reaching bedrock by freely varying the attributes of a thing. It would not work for the elements. What we cannot strip away from the elements are disjunctive properties: Every element must be either hot or cold and either moist or dry, and cannot be neither hot nor cold or neither moist nor dry. [4] Therefore, even if we change all features of a thing, this need not amount to stripping all properties away. There will be at least some disjunctive properties that resist denuding. This means that changing all properties of the remains of Socrates will not

---

[4] Some say that Aristotle believes in a prime matter that would underlie the change of one element into another, others deny this (Bostock 1995, p. 73). The very fact that he speaks of fire as an element seems to imply that there is nothing more basic than fire underlying it. However, I ignore this issue here since as Charlton points out, "there will be no empirically discoverable difference between a world which contains prime matter and a world ... in which there is nothing more basic than the elements" (1983, p. 197). In any case, all we can say is that when fire turns into air, "the hot" turns from dry to moist; whatever it may be that is hot here.

lead us to anything that might perhaps not have any of these properties. Moreover, it is not clear why a universal variation of attributes should leave us with matter. Given the remains of Socrates, we could also change them by gradually replacing parts, as we may change a thing by replacing the matter out of which it is made (cf. Physics I 7, 190b8-9). It is not at all clear why matter should be whatever survives such a procedure (Dancy 1978, p. 396-7).

4. τόδε τι

We do not arrive at anything that has no properties by taking away its properties, nor by changing all of them, since whatever we do to it, the thing before us will take on some new contrary properties or keep some disjunctive properties. How are we supposed to take all else away, then? In the course of his argument in Z 3, Aristotle articulates what sounds like an official definition of what it is to be matter:

> I call matter that which is in its own right neither said to be such (τί), nor so many (πόσον), nor anything else by which a thing may be determined. For there is something of which each of these is said (κατηγορεῖται), and which itself has a being different from that of each of the categories—for while others are said of substance, the latter is said of matter—and so the last thing will in its own right be neither such, nor so many, nor anything else at all. (1029a20–5)

This definition is interesting because it does not refer to any actual, physical stripping of properties. Matter is what is not <u>said</u> to be τί, nor πόσον, nor anything else by which one may specify what or how something is. This way of putting it suggests that the stripping concerns what is said of a thing, rather than what inheres in a thing. Since τί, πόσον etc. is the initial segment of the list of categories, we are apparently supposed to take a noun phrase that refers to a substance and delete everything from it that falls under any of the categories. Some noun phrases are, of course, non-starters. If we take "Socrates," for instance, and take away what would

fall under one of the categories, we end up with nothing. But take, for instance, "the philosopher who taught Plato." To teach is an action, Plato is a primary substance, and being a philosopher is a quality (or a passion?). What is left is "the ... who ...," since demonstratives and pronouns do not fall under any of the Aristotelian categories.

I take it that this is what Aristotle tells us in Ζ 3. The idea is that we may take a complex noun phrase of the form "this such," which refers to a composite substance, and delete all terms from it that concern the attributes and determinations of the substance. To be sure, Aristotle does not distinguish between noun phrases and their object as clearly as we do. However, since we do distinguish between them, it is better to think of stripping noun phrases rather than stripping what they refer to. When we do this, we will be left with a phrase that merely refers to the matter of the composite substance, which seems to be the ultimate subject of which all the descriptive bits may be said. If this is correct, matter is what corresponds to a mere "this" and "that." This reading has the considerable advantage that to refer to a thing without mentioning its features is certainly not the same as referring to something that does not have any features (Stahl 1981, p. 178).

An important detail to note in the present context is that Aristotle tells us to subtract first the τί, and then the other categories, starting with πόσον. However, whereas πόσον is in fact the expression that Aristotle uses for the category of quantity, τί is not the word he uses for the category of primary substance. Rather, he refers to substances by means of the complex expression τόδε τι ("this such," Categories 5, 3b10).[5] When we take the τι away from this expression, a mere τόδε remains. Matter, as it appears in Ζ 3, is what corresponds to this isolated τόδε. As Anscombe writes:

> Now 'this matter' is τόδε but not τὶ: that is, it is designatable, identifiable, but it
> is not as such of any specific kind or necessarily possessed of this or that property

---

[5] Cf. Smith (1921), Frede and Patzig (1988, Kommentar p. 14-5), Witt (1988, p. 163-4 fn. 15), Gill (1994, p. 68).

or dimensions, ... (1953, p. 86)

Matter is, as it were, the τόδε τι deprived of its τι.[6] Τι, as part of the list of categories and as part the phrase τόδε τι, refers to the type that a thing may be said to instantiate.[7] As Loux writes, it is "a placeholder for substance-predicates" (1991, p. 61), that is, for a reference to a specific paradigmatic form. In contrast, the categories other than the first one specify how a thing looks and feels, how it relates to others, and so on. These are properties or features of a thing, which may or may not inhere in it. For them, Aristotle generally uses the expressions τοιόνδε or ποιόν τι.[8] Τοιόνδε

---

[6] I disregard the grammatical distinction between τί (the interrogative pronoun quid) and τι (the indefinite pronoun aliquid). Both pronouns refer to the same in different modes: in a question or in an answer to this question. "Was" and "wat," in German and Dutch, may also be understood in both ways (as "what" and "something").

[7] Tugendhat writes that τόδε τι does not mean "this somewhat" but "a this" (2003, p. 25 fn. 22). He backs this up with a reference to Categories 5, 3b13–17, where Aristotle first uses τόδε τι and then ποιόν τι, such that in both phrases, the τι seems to indicate particularity rather than being of a certain kind (ποιόν τι = a particular quality, τόδε τι = a particular "this"). In 3b13-17, Aristotle says that secondary substances are ποιόν τι. However, secondary substances are not particular qualities. If at all, they are a highly special kind of determination of a thing. The τι in ποιόν τι does not refer to the particularity of a quality token but indicates that what we have here is an odd kind of quality. It performs the same function as I suppose it does in τόδε τι: A τόδε τι is a special kind of particular individual and the τι is a placeholder for what is special about it, just as secondary substances are special kinds of qualities, and the τι indicates that they are somehow special. The same holds true, e.g., for ἕν τι in Z 12, 1037b27, and τοιόνδε τι in the Sophistical Refutations 178b38: Here, τι does not indicate particularity but that a unity or the bearer of a quality is of a certain type.

[8] See Metaphysics B 6 (1003a8-9) and Z 13 (1039a1-2) for τοιόνδε; Categories 5 (3b15-16) for ποιόν τι.

may function as a placeholder for a quality or an attribute of a substance.[9] When we refer to something as a τόδε τοιόνδε in this sense, we refer to it as an accidental compound, as for instance the bronzen sphere in Metaphysics Z 8 (1033b19-24) or the bed in De Partibus Animalium I 1 (640b26-28). The phrase τόδε τι does not refer to an accidental compound, a thing as having a certain shape or set of features, but to a substance, a thing that instantiates a certain type.

I find it helpful in the present context to clearly distinguish between τι and τοιόνδε. There are, however, also passages in which Aristotle says that "man" refers to a τοιόνδε, rather than a τι or τόδε τι. When he says this, he usually emphasizes that "man" and other things that are said to be <u>common to many</u> are not τόδε τι but τοιόνδε τι (Sophistical Refutations 22, 178b37; Metaphysics B 6, 1003a8-9). This does not mean that "man" simply is a τοιόνδε. "Man" is similar to a quality (τοιόνδε) only when it is said to be common to many, as in the statement "Socrates and Kallias are men." Still, there are uses of this term where it is not said to be common to many, and when used in such ways, "man" is not like a quality.[10] Further, in Categories 5, Aristotle writes that "man" is not a τοιόνδε in the same sense in which a quality such as "white" is one. "Man" does not signify a quality in a substance but rather a qualified substance (ποιὰν γάρ τινα οὐσίαν σημαίνει,

----

[9] Kung (1981, p. 207) and Strobel (2007, p. 34) ignore the distinction between τι and τοιόνδε and therefore assume a direct contrast between τόδε τι and τοιόνδε. In fact, this contrast involves two independent distinctions: the one between τόδε X and X (without τόδε), and the one between τι and τοιόνδε. Had Strobel observed the latter distinction, he might have reached a different conclusion regarding his main topic, self-predicative sentences; for such sentences make more sense concerning types (τι) than concerning properties (τοιόνδε). Man (in general) is an animal, but humanity is not an animal. Likewise, man is human, but humanity is not.

[10] In Z 13, Aristotle argues that none of the things that are predicated as common to many can be an οὐσία; all they can be is τοιόνδε (1038b37-39). See my discussion of Z 13 in Chapter 6, section 6.

3b20-21). As part of the phrase τόδε τι, τι is not said to be common to many, but used in order to single out a particular instance of a certain type.

All in all, one might imagine combining the three linguistic devices, τόδε, τι, and τοιόνδε, in order to refer to a thing that instantiates a type and also has certain properties: (τόδε + τι) + τοιόνδε. If this were proper Greek, it would be the Greek version of a noun phrase corresponding to a Sellarsian sentence:[11] "This such, which is so and so." In this somewhat artificial formula, (τόδε + τι) + τοιόνδε, a "this" (τόδε) is referred to as an instance of a type (τι), and the substance that it thus constitutes is said to have a property (to be τοιόνδε).

It would of course be insane to consistently use words that are as common as τόδε, τι, and τοιόνδε in a technical sense. Τόδε may be used anaphorically in order to refer to all kinds of things, and τι may indicate all kinds of specificity or particularity. We should therefore not suppose that Aristotle always restricts the use of τόδε and τι in the way that he seems to suggest in Ζ 3. Indeed, he sometimes uses τόδε τι when he refers to matter,[12] and on the other hand, he also uses τόδε (without τι) when referring to paradigmatic forms.[13] Nonetheless, the formula (τόδε + τι) + τοιόνδε helps us distinguish between the matter, the paradigmatic form, and the properties of a thing. Moreover, it reveals that the way in which Aristotle speaks of sensible substances in the Categories (τόδε τι) is compatible with the way in which he speaks of matter in Metaphysics Ζ 3 (a τόδε deprived of its τι).

———————————

[11] See Chapter 2, section 3: "This such is so-and-so."

[12] Aristotle refers to matter as τόδε τι in Physics I 7, 190b24–27. Cf. Dancy (1978, p. 402).

[13] See the following passages, taken from Bonitz: De Generatione et Corruptione I 3, 317b9, 21, and 28; Metaphysics Ζ 13, 1038b24; Λ 1, 1069b11; Ν 2, 1089a11 and b32. In Ζ 7, Aristotle also lists the categories as τόδε ἢ ποσὸν ἢ ποιὸν ἢ ποῦ (1032a15). That Aristotle sometimes refers to paradigmatic forms as τόδε or τόδε τι is probably related to what I emphasize in Chapter 1: that the formal cause of a thing is not a property but a generic instance of a type.

5. The Timaeus

Charlton points out that the idea of matter as a bare τόδε ultimately underlying all predication may be traced back to Plato's Timaeus (1983, p. 204).[14] This is the relevant passage:

> Rather, the safest course by far is to propose that we speak about these things in the following way: what we invariably observe becoming different at different times, such as fire, not this (τοῦτο) but in each case what is such to call fire, and water not this but always what is such, and never to speak of anything else as this, as though it has some stability, of all things at which we point and use the expressions, that (τόδε) and this (τοῦτο), and so think we are designating something. For it gets away without abiding the charge of that and this, or any other expression that indicts them of being stable. It is in fact safest not to refer to it by any of these expressions. Rather, what is such—coming around like what it was, again and again—that's the thing to call it in each and every case. So fire, and generally everything that has becoming, it is safest to call what is altogether such. But that in which they each appear to keep coming to be and from which they subsequently perish, that's the only thing to refer to by means of the expressions that and this. A thing that is some such or other, however, —hot or white, say, or any one of the opposites, and all things constituted by these— should be called none of these things. (49D-E, tr. Zeyl, modified)

This passage is interesting in the present context because Timaeus also seems to make a point about the use of the word τόδε (and τοῦτο). Timaeus is here talking about the precosmos, in which all things are constantly changing into different things. In this situation, we cannot reliably identify and track anything in terms of its properties, because all things constantly change all their properties. Therefore,

---

[14] Note also that a significant proportion of the relevant passages listed by Brunschvig (1979, p. 132-5) involve a discussion of Plato.

Timaeus suggests that instead of saying "This is fire," we should say "What is such is fire" (μὴ τοῦτο ἀλλὰ τὸ τοιοῦτον ἑκάστοτε προσαγορεύειν πῦρ, 49D).

It is disputed how to translate this passage. I confine myself to two possible readings, according to which Timaeus either makes what I call the feature placement recommendation, or what I call the predicate logic recommendation. According to the feature placement recommendation, we should not think of fire as a fiery thing, because in a world of constant change, such a fiery thing will eventually cease to be fiery. Rather, we should confine ourselves to statements that register the occurrence of fieriness. Instead of saying "This is fiery," we should say "Fieriness is occurring" (cf. Cherniss 1954, p. 118; Lee 1967, p. 5). According to the predicate logic recommendation, on the other hand, we should always be aware that we can refer to a thing only in terms of its properties (Silverman 1992, p. 112). This means that in a constantly changing world, we should always supplement a demonstrative such as "this" with a term that indicates the property in terms of which the thing referred to is identified. Therefore, we should never simply refer to fire as "this" but always refer to it by means of a complex expression such as "this fiery item" (cf. Gulley 1960; Zeyl 1975, p. 129).

Depending on which interpretation we choose, what Timaeus says relates more or less well to Aristotle's contrast between a bare τόδε and a qualified τόδε τοιόνδε. Aristotle argues that a bare τόδε does not refer to anything specific and separable and that in order to refer to a thing, one needs to supplement the τόδε with a descriptive term. This sounds a lot like the feature placement recommendation, but there is an important difference. In the Timaeus, the added descriptive term seems to be a predicate denoting a property (τοιόνδε), not a type of which the thing is an instance (τι). At any rate, of the two kinds of descriptive term that may be added to a τόδε, τοιόνδε and τι, Timaeus knows only one. If things can only be identified and tracked in terms of their properties, all indexicals, variables, and proper names must be backed up by ascriptions of properties to items. A formal language that conforms to this rule would have no individual constants or proper names among its primitive

symbols, only predicate constants and variables. The only basic way of referring to individuals would be "the x that is F," and proper names or individual constants could only be introduced by defining them, e.g., by "Peter := the x that is F."

Aristotle does not make the predicate logic recommendation since he acknowledges a difference between properties (τοιόνδε) and types (τι). To refer to something as a τόδε τι is not to attribute a property to a τόδε but to identify it as an instance of a type (cf. Rapp 1995b, p. 241). Put in terms introduced in Chapter 2, Aristotle distinguishes between substance concepts, by which one identifies a substance without attributing a property to it, and logical predicates, which may be said of a given substance.

Zeyl writes that according to the "old λόγος," our descriptive terms are used in order to identify things that may change their properties and nonetheless be tracked over time, and that according to the λόγος that Timaeus is going to put in its place, our references can only be used in order to describe things, not in order to identify them. Zeyl thinks that this reform is necessary because the assumption that descriptive terms may be used in order to identify things leads to "absurdities and contradictions" (1975, p. 128-9). However, although this might be what Plato's Timaeus says about the precosmos, it does not seem to be what Plato himself says about the sensible world. If it were, it would be difficult to reconcile with some implications of the overall argument of Plato's Cratylus.

In this dialogue, Cratylus argues that names can only refer to their bearers by describing them (Cratylus, 429C). This is the position that Zeyl attributes to Timaeus, confined to names, but later in the Cratylus it is extended to all kinds of predicates, type terms, and even sentences (λόγοι, 431B). Hermogenes takes the opposite view, that names refer to their bearers only by convention, without in any way describing them (384D). Plato argues against both Cratylus and Hermogenes. He shows that no reference would be possible if all speech acts were acts of calling things by

conventional names or all of them were acts of describing things.[15] For if all references were made in the way in which one calls something by a conventional name, it would be difficult to see how one can make a false statement about anything. The statement that the cat is on the mat would not describe the cat as being on the mat; it would only attempt to call some state of affairs by the conventional name "cat being on the mat." If there is no cat on a mat, this attempt fails, and no statement has been made about anything, a fortiori not a false one. (This is not exactly Plato's argument, but compare 385B-E.) On the other hand, Plato shows that not all references can be made by describing a referent. In a lot of cases, it is possible to think of a name as a description, as when one takes "Hermogenes" to stand for "born of Hermes," but this cannot go on ad infinitum. At some stage the parts of our speech must refer to objects without describing them (435A-D). This means that in order to refer to anything we must use language both in order to call items by conventional names and to describe them. Now if Plato proves that in order for our language to refer it must do so both by convention and description, he also shows that in a world as Timaeus imagines it our language must fail to refer. According to Timaeus, we can only refer to things by describing them, and according to the Cratylus, this implies that we cannot refer to anything. And indeed that we could not refer to anything in a world of constant change is something that Plato claims in the Theaetetus (182D). Aristotle's distinction between τι and τοιόνδε might be motivated by such arguments. When we refer to something as τόδε τοιόνδε, we describe a given τόδε as having a property, but when we refer to it as τόδε τι, we do not. We identify it as an instance of

---

[15] This is in line with what he shows in the Theaetetus and the Sophist: that only complex discourse can be true or false.

a type. Τι refers to an object like a conventional name does, not like a description.[16]

There is another important difference between the account of an underlying subject that is involved in the predicate logic recommendation and the notion of matter as referred to by a mere τόδε that Aristotle discusses in Metaphysics Ζ 3. Timaeus does not use the word "matter" (ὕλη). The feature placement recommendation leaves no room for matter since it leaves no room for any underlying subject of forms, attributes, or determinations. The predicate logic recommendation leaves room for an underlying subject of properties, but this subject is not matter. For in the precosmos, the only possible referent of a bare "this" is the receptacle of all becoming, and this receptacle is obviously not the matter of which things consist (Cornford 1937, p. 181). For instance, the receptacle is that in which things move, and if it is really supposed to stay perpetually the same, the receptacle itself cannot move. Now assume that Socrates is really a perpetually unchanging receptacle x such that this x has the property of being Socratic. When Socrates enters the room, this x does not enter the room. Rather, the property of being Socratic changes its place, as it were, within x, the receptacle. Thus in Timaean newspeak, Socrates does not really enter the room, but parts of the room start Socratizing. If this is a concept of matter at all, it is the one that Aristotle attacks in Physics IV 2 (209b21-32). Further, since the parts of the receptacle cannot move, they may be identified as such in terms of their relative spatial position. This shows that the kind of matter that Aristotle seems to introduce in Ζ 3 is not even supposed to be the same as Plato's receptacle. For Aristotle argues that the spatial location of an item is one of its features, addressed by words falling under

---

[16] Cratylus thinks that all empirical concepts work like descriptions, which have an object to concept direction of fit (cf. Chapter 2, section 2). Hermogenes thinks that they all work like conventional names, which have a concept to object direction of fit. (We call a name and something may respond.) When Plato argues that reference is only possible if language and thought refer to their objects in both ways, he makes the same point as Kant, expressed in Engstrom's words: Cognition has no unidirectional direction of fit (2002, p. 53).

the category of quantity. Therefore, when we denude a thing of its look and feel, we must also think away its spatial location and dimensions. If we do this, not even a receptacle is left over. Aristotle, however, suggests that what is left over is matter.

6.   Conclusion

This chapter has shown that the notion of matter that Aristotle introduces in Z 3 need not be taken to refer to the analysis of change in the Physics. It is already implicit in, or at least compatible with, the expression τόδε τι that Aristotle already uses in the Categories.[17] The two parts of this expression may be taken to represent, respectively, the matter of a thing and its paradigmatic form; not in isolation, but only as parts of the phrase. As I argue in Chapter 1, and show in more detail in Chapters 5 and 6, the paradigmatic form of a thing is not this particular thing but still an instance of its type. The matter of a thing is what makes it "this" (τόδε); the paradigmatic form is what makes it "such" (τι), i.e. an instance of a type; and a further added predicate would make it "so-and-so" (τοιόνδε). This does not mean that matter is a principle of individuation, which would answer the question how Socrates differs from Callias. Rather than making the difference between two instances of one form, it makes the difference between a particular instance (this instance, τόδε τι) and its paradigmatic form (a generic instance, τι). This, however, is not yet a satisfactory answer to the question of what matter is. The answer is given in the next chapter.

In this chapter, I have further shown that according to one reading of the Timaeus, Aristotle and Timaeus make similar points. Timaeus says that one can (and therefore should) only refer to things in terms of their properties. Aristotle, however, distinguishes between properties and substantial forms and argues that we can only refer to things in terms of one of these. When Aristotle writes that we must first take

_____

[17] Rapp also argues that the problem that Aristotle addresses in Z 3 is already present in the Categories (1995b, p. 438). In the Categories, Aristotle refers to particular substances (τόδε τι) as instances of second substances (τι). In Z 3, he justifies this by arguing that there can be no other way of referring to them.

away the properties of a thing and then also its τι, he ignores his own distinction between τι and τοιόνδε. In this sense, the definition of matter in Ζ 3 is formulated in Timaeus' perspective. By showing that it does not lead anywhere, Aristotle also shows that one must distinguish between properties (τοιόνδε) and paradigmatic forms (τι).

Matter as Subject and Attribute

I have argued that the matter of a sensible substance may be taken to be what corresponds to the τόδε in Aristotle's complex phrase τόδε τι. It is important to keep in mind that τόδε refers to matter only as part of this phrase. In the end of Metaphysics Ζ 3, Aristotle argues that if there were anything left after denuding a thing to a bare and isolated τόδε, this thing would indeed be the primary subject of predication. However, he claims that in fact nothing would be left. An isolated τόδε would not refer to anything, not even to matter. Therefore, the notion of matter as something that is left after completely denuding a thing of everything that may be said of this thing is not Aristotelian. As Gill writes, "[h]e has not clarified his own notion of matter by demonstrating its absurdity" (Gill 1989, p. 30).

Aristotle says that τόδε does not refer to an underlying substance because it does not refer to anything separate (Ζ 3, 1029a27-30). In this chapter, I ask in what sense the matter of a thing is inseparable from this thing. Some think that substances or paradigmatic forms are predicated of matter in the same sense in which properties are predicated of compound substances. If this were true, the formula (τόδε + τι) + τοιόνδε would involve two predicative ties of the same kind. In a first step, a form (τι) would be predicated as a property of matter (τόδε), and in a second step, further properties (τοιόνδε) would be predicated of this substance (the τόδε τι). Others think that conversely, matter is an attribute of a sensible substance and that it is inseparable in the same sense in which an attribute of a thing is inseparable from this thing. Still others think that the matter of a thing is a part or ingredient of it and more generally that a thing and its matter are two distinct but somehow closely related items. I argue that the matter of a thing does not bear the paradigmatic form of this thing as one of its properties, nor is it an attribute of this thing, or an ingredient or part of it, or a separate thing.

1. Matter as Subject

By distinguishing between the τι and the other categories (τοιόνδε), Aristotle apparently distinguishes between two kinds of property. First, all predicates that belong to categories other than substance stand for properties of substances. Second, it seems, predicates in the category of substance stand for a property that matter takes on when it makes up a substance (Metaphysics B 1, 995b35; B 4, 999a33;[1] Z 3, 1029a23-24; Z 13, 1038b5-6). It seems that just as the τοιόνδε is predicated as a property of a τόδε τι, the τι is predicated as a property of a τόδε. This appears to be confirmed by what Aristotle writes in Θ 7, 1049a34-36: Where a form is said of a substance, the substance is the basic subject, but where a τόδε τι is said of something, the basic subject is its οὐσία ὑλική. Instead of saying that "man" is said of a composite that consists of matter, Aristotle here apparently suggests that "man" is a property of matter. On the other hand, however, he emphasizes that substances are not predicated of any underlying thing (Z 13, 1038b15).

Witt gives the following argument why "man" cannot be a property of a man's matter (1989, ch. 4 §§4-5). Assume that "man" is an essential property of Socrates' matter. This should mean that his matter couldn't be anything other than a man, which is false. When Socrates dies, we have the matter without the man. Therefore, "man" can only be an accidental property of Socrates' matter (as Loux thinks, 1991, p. 121). But this again cannot be what Aristotle wants, Witt writes, since whether Socrates is a man or not cannot depend on something else's having some accidental property. This would upset the distinction between what something is and how it looks and feels. In the end, what this man is would depend on how his matter looks and feels, and What-questions would boil down to How-questions. However, it is not obvious that this would follow. The answer to the question what a thing is may depend on the answer

_____

[1] Note that the parallel passage in Metaphysics K 2 does not any longer involve the assumption that form or substance are predicated of matter, but rather that matter is only potentially something (1060a20-1) and that composite substances are somehow "form with matter" (1060b24-5). Z 3 might be a turning point concerning this issue.

to the question what accidental properties its matter has, and still, these two might be different questions. After all, one of them is a question about a thing, whereas the other is about the matter of a thing. This makes them different. Therefore, it need not follow that What-questions and How-questions are indistinguishable.

There are two further possible arguments against assuming that in the phrase τόδε τι, a property is predicated of a subject. Both arguments involve claims that I put forward in Chapter 1 and that I defend in more detail later. The first is that in the phrase τόδε τι, something is said to instantiate a type, and to instantiate a type is not the same as to exhibit a property (other than the vacuous property of instantiating this type). I distinguish between types (or essences) and properties in Chapters 5 and 6. The second argument relies on a claim that I further defend later in the present chapter: that the matter of a thing is something that potentially is this thing. Since the realization of a potential is not a property of this potential, and τι is said of τόδε in the same way in which the realization of a potential is said of this potential, it is not said of τόδε as a property. Before returning to this second argument, I will now briefly discuss the suggestion that matter might be an attribute of a substance.

2. Matter as Attribute

Some suggest a different way of explaining how the matter of a thing depends on this thing. It might be inseparable in the same sense in which an attribute is inseparable from that of which it is an attribute. In an instructive paper on the Aristotelian concept of matter, Dancy writes:

> [Aristotle] is suggesting, I think …, that the picture according to which the matter of a thing appears as the subject for its predicates is back-to-front: the material constitution of a thing ought more properly to appear in predicate position, so that, if fire were prime and air second, it would be better to say that air is fiery, as one says that Socrates is pale, ambulating, moving, and so on. (1978, p. 411)

In a more recent paper, Beere claims something very similar (2006, p. 310). According to Dancy and Beere, matter is really an attribute of a substance, so that when Aristotle tells us to take away all the attributes and determinations of a thing, this includes matter. Unsurprisingly, what is left after doing so is not matter but a bare nothing.

Evidence for this view seems to be that Aristotle sometimes speaks favorably of the custom of treating the material constitution of a thing as one of its features. For instance, we call a sphere brazen and a bench wooden instead of calling the brass spherical or the wood bench-like. When we do this we attribute a certain material makeup to a thing rather than attributing a paradigmatic form to its matter. Aristotle even coins a general phrase and says that we do not call a thing τόδε but ἐκείνινον ("thaten") when it consists of "that" (Metaphysics Z 7 1033a5-23; Θ 7 1049a18-b3).

However, it is not clear whether we should attribute to Aristotle the view that matter is a kind of attribute. For several reasons, matter cannot be an attribute of a thing on a par with its other attributes. First, particular attribute instances generally depend for their existence on the substance of which they are attributes. Therefore, if Socrates' matter were one of his attributes, it would have to cease to be as soon as Socrates ceases to be. However, the matter that is left behind when Socrates dies is certainly the same matter that formerly belonged to him. It continues to exist when Socrates ceases to be. To be sure, Socrates' matter is only what it is, matter of Socrates, because and insofar as it belongs or once belonged to Socrates, but this does not mean that it depends on Socrates for its continued existence (Cooper 1973, p. 697-8).

Further, as we have seen, Aristotle says that in some sense the paradigmatic form of a thing may be said of its matter: "while others are said of substance, the latter is said of matter" (1029a23–24). It would not make much sense to say that all attributes are said of a substance and that the substance is said of matter, if matter were one of the attributes. Moreover, although Aristotle qualifies the claim that matter is the basic subject of predication in Z 3, he still concludes that there are three kinds of

underlying subject: matter, form, and the composite (Z 3, 1029a29–30; H 1, 1042a26–31). Therefore, there must be some sense in which matter is a subject and not an attribute.

Finally, Aristotle may not actually approve of the custom of calling things "thaten." There are two conceivable reasons why he might not. The first is suggested by Heidegger in Vom Wesen und Begriff der Φύσις. When a bed is made out of wood, something turns into something else, for the wood and the bed are conceptually and metaphysically independent. The wood could make up a different thing, and the bed could consist of something different (where the relevant differences are differences in kind). In such cases, it is appropriate to call the bed wooden. However, when a tree comes to be, we have no independent term for that out of which it comes to be. That out of which the tree comes to be is neither conceptually nor metaphysically independent from that out of which he comes to be. Therefore, it is not appropriate to call it "thaten." As Heidegger puts it, "only what is something different than wood can be wooden" (Gesamtausgabe 9, p. 253). If Heidegger is right, accidental compounds may be called "thaten," but substances may not.

Gill furnishes Aristotle with a more general aversion against calling certain things "thaten" (1989, p. 123-4). As Aristotle writes, it would be wrong to say that a healthy man comes to be out of a man. By the same token, Gill argues, we should not say that shaped bronze comes to be out of bronze. Now in the first case, we can say that the healthy man comes to be out of the sick man, which is correct, because "sick man" is the privation of what comes to be. In the second case, however, we do not have a convenient pair of words for the form and its privation, so that we could say something of the form "The A that is B comes to be an A that is C" ("The bronze that is unstatued comes to be bronze that is statued"). Sometimes, however, we still want to describe the process by which a statue comes to be as a kind of qualified change, and this we can only do by referring to something that remains the same throughout this change. When we say that bronze comes to be a statue, it is not clear what it is that remains the same (a copper statue might come to be out of bronze). Therefore,

people sometimes say that the bronze comes to be a bronzen statue. This, however, is not a solution since that out of which the bronzen statue comes to be is still not properly referred to, i.e., as the privation of a statue. Calling a statue bronzen is thus only a botched attempt at assimilating "the statue comes to be out of bronze" to "the healthy man comes to be out of the sick man." If Gill is right, Aristotle therefore thinks that calling things "thaten" is altogether misleading.

3. Matter as Potential

Given the linguistic interpretation of denuding introduced in the previous chapter, the most obvious way to make sense of the claim that matter is inseparable is the following. Τόδε does not refer to anything separate because τόδε, when separated, does not refer to anything. It refers to a substance only in combination, only as part of τόδε τι. In isolation, it refers to something only potentially, since it would refer to something if one would add a τι. In this sense, matter is not actually a τόδε τι but only potentially one (1042a27–28; cf. Z 7, 1032a12-b6).

Ontologically speaking, matter is inseparable from what it is matter of because it is this latter thing potentially. For in general, in order to describe a potential as such, one needs to describe the state that constitutes its actualization. In this sense, actuality is prior to possibility (Metaphysics Θ 8, 1049b5). Since matter is what potentially is a typical instance of a form, the paradigmatic form of a thing is prior to its matter (1050a15, 1088b1-2, 1092a3-5). Therefore, the only way to specify what it is to be the matter of a substance is to say what it would be to be that substance. That matter is not separable means, basically, that there is no independent way of characterizing it (as matter; cf. Witt 1989, p. 141).

Aristotle confirms this when he writes that the matter of a substance is nothing but this substance by way of privation (ὑποκείμενον ὡς κατὰ στέρησιν, H 1,

1042b2-3).[2] It is a τόδε τι deprived of its τι. Now as he points out in Z 7, the form of a privation is the form of what is missing and only potentially there (1032b3-4). Hence, the matter of a thing has, as matter, no form of its own (cf. Z 10, 1036a8-9). Its form is the form of what it potentially is, and this is (or is part of) the paradigmatic form of the thing of which it is the matter.[3]

In Metaphysics H 6, Aristotle writes that "the proximate matter and the form are one and the same, one potentially and the other actually" (1045b17-19). Haslanger (1994) considers three interpretations of this statement: matter and form may be "one and the same" by being (1) identical, (2) two aspects of the same, or (3) two parts of the same. She argues convincingly against (1) and (2). However, this would only imply that (3) is correct if the alternative were exhaustive, which it is not. For Aristotle's statement in H 6 might as well have the logical structure of, for instance, "George W. Bush and Barack Obama are one and the same: one formerly, the other presently" (uttered in 2009). It does not follow from this that Bush and Obama are identical, two aspects of the same, or two parts of the same. In the same sense, A may be the same as B, one potentially and the other actually, so that this does not imply that A and B are identical, two aspects, or two parts of one thing. These are not ways in which a potential relates to its actualization.

The matter of a thing is potentially something specific, and the thing is actually something specific, and what they are in these different modes is one and the same (Owens 1963, p. 341 and 363; Lewis 1994, p. 254-5, fn. 20). When substance is said of matter, actuality is said of a potential (Lewis 1991, p. 265). When a potential is actualized it is not combined with its actualization, and when matter is formed it is not combined with its form (cf. Metaphysics H 6, 1045b9-12). Therefore, the matter and

_____

[2] In Physics I 9, Aristotle emphasizes that matter and privation are not the same (192a3-4). What he means is: The matter of a thing cannot be the same as a lack of form because it may also be there when the form is present.

[3] This is why Aristotle says, in Physics II 2, that for a different form, there is different matter (194b9).

the form of a natural thing are not two numerically different things (Frede 1994, p. 175). The matter of a natural thing is not one of its components, parts, or constituents, nor is it one of its "ingredients" (as Fine 1992 has it, p. 36). As Rorty says, "[I]t isn't really a part of the substance at all—it is just the substance potentially" (1973, p. 416).[4]

In particular, the relation of a potential to its actualization is not a relation of a subject to its properties. One may say of a potential that it is actual, but this is not to attribute a property to this potential (cf. Rapp 1995b, p. 316). The statement "This matter is Socrates" does not have the logical structure of a property attribution such as "This man is pale." It says of a potential that it has been realized, and the realization of a potential is not a property of this potential or of that which has the potential (cf. Z 13, 1038b4-6; H 2, 1043a5-7). Realizations are not properties, any more than existence is one. The statement "This matter is Socrates" has the logical form of statements such as "Dinner is ready" or "The book is written." In such statements, we do not say something of something else. We say of something that this thing is actual or complete.

If there is a sense in which form is said of matter, it is not the sense in which something is said of something else. Form and matter are both what they are by virtue of the same form; one is actually this form, the other is it potentially. Therefore, when a paradigmatic form is said of matter, a thing is said of itself, as it were: The actuality of a thing is said of something that potentially is this thing. Now in the same sense in which existence is not a property of an existing thing, actuality is not a property of an actualized potential. Therefore, it is misleading to suggest that in the phrase τόδε τι, a property is attributed to a parcel of matter, and it is also misleading to think of matter and form as distinct but related things.[5]

––––––––––––––––––––

[4] Rorty also identifies the matter of a thing with its genus (p. 411) and with the "ensemble of accidents in a particular substance" (p. 418). I do not follow him in these respects.

[5] This theme will re-appear more forcefully in Chapter 7, concerning the relation between an efficient cause and its effect.

4.   A Note on Material Constitution

This leads to a corollary concerning the problem of material constitution. This problem is usually stated by taking two items, such as a lump of clay and a statue, and asking how they must be related if one of them is the matter of the other. If the above is right and applies in this case, this way of stating the problem is questionable. Thomson (1998), for instance, considers three possibilities concerning a clay statue and its matter. First, the two might be strictly identical. This is not possible because the clay may exist without having certain properties that the statue must have in order to be what it is. Second, being a statue might be a property of the clay. This is not possible because the clay might be gradually replaced, and the very same statue might also exist without the clay. However, a property could not survive the destruction of that in which it inheres. From this, Thomson concludes that the clay and the statue must be two things capable of independent existence, and she goes on to ask what the relation between them is when one of them constitutes the other. She does not consider the possibility that has emerged above: that the statue and the clay are not strictly the same but that one of them is actually what the other potentially is. To be sure, not every neglected alternative is a way out, but this one might be worth considering.

Types as Classes

The formal cause of a natural thing is what this thing would come to be as a result of its natural development. When a natural thing is fully developed, its formal cause is what it actually is. The formal cause is not this particular thing but still an instance of its type, and an instance of this type is also what this particular thing is. When it is not fully developed, either because the development is not yet complete or because something went wrong, its formal cause is what it would be if it were fully developed.

Aristotle also refers to the formal cause of a thing as the "account of the τί ἦν εἶναι" (Physics II 3, 194b27). In Metaphysics Ζ 4, he writes that only such things have a τί ἦν εἶναι whose account (λόγος) is a definition (1030a6-7). In Posterior Analytics II 11, he says that the τί ἦν εἶναι of a thing is that which is signified by its definition (94a34-35). Therefore, the τί ἦν εἶναι is what is signified by its definition. Now since what the definition of a thing signifies is also its essence, the essence of a natural thing is also its formal cause.[1] Thus for a natural thing, the following three questions amount to the same:

(1) What is its formal cause?

(2) What would it be if it were fully developed and all went well?

(3) What does its definition signify?

Aristotle emphasizes that no particular thing can be defined in its particularity (Metaphysics Ζ 15, 1040a7). Therefore, the formal cause of a thing cannot be a particular thing. This is also clear from the fact that the formal cause of a thing is what

---

[1] Formal causes are essences of things, not of processes. Further, because nothing can be a cause of itself, the essence of a thing that coincides with its essence is not a formal cause.

this thing or its matter potentially is. The object of a potential is not a particular thing, any more than the object of an ability is a single action. Potentials and abilities are realized by particular things and actions, but what they are potentials and abilities for must be more than only one such particular thing or action. The object of an ability is a type of action or activity. For instance, that I understand French does not mean that I understand only one particular utterance but that I am able to understand a potentially unlimited number of particular utterances of a certain type. The same is true for the potential to be a natural thing. The object of this potential is not only one particular thing (the particular thing that comes to be when the potential is realized), but a type of thing. Within the scope of the phrase "potential to be ___," the object of the potential is general. Since the formal cause of a natural thing is what this thing or its matter potentially is, it is therefore a general type of thing rather than a particular instance of this type.

The aim of this and the following chapter is to say as much as possible about types that is at once Aristotelian, correct, and intelligible to us. I do not extract from Aristotle an account of what types are. Rather, I give an account on independent grounds. I mark this change of strategy by using the term "type" instead of the Aristotelian terms "formal cause," "paradigmatic form," and "essence." Like everywhere else in this book, I focus on types of natural things rather than types of artifacts and immaterial things. I discuss several versions of the idea that such types are a special sort of set, arguing that at least some types cannot be identified with extensionally defined sets and that there are also problems with defining types in terms of simple intensions (finite lists of non-relational features). After briefly touching upon the notion of a polytypic class, I follow Gasking and construe types as referent classes or clusters. It will turn out that this matches the procedure preferred by botanists and zoologists: In order to define a type, they single out a particular instance and specify a relation that all other instances of the type must have to this instance. Identifying and classifying an item as an instance of a type will thus be shown to involve two tasks. First, one must be able to identify the "focal" instance of the type, either by direct reference or by description. Second, one must be able to

specify a relation of sameness in type so that one can determine whether a given item is an instance of the same type as the focal instance. I argue that in order to answer the second question one must be able to describe the expected result of successfully replicating a type. This connects to what I say about Aristotle: that the question of what the definition of a type signifies is the same as the question of what the result of its typical development would be.

1. Sets and Classes

One possible answer to the question of what types are is the following: The type that a thing instantiates is a set of which it is a member. For this to work in the case of natural things, one should not identify sets with their current extension. There are several distinct types of extinct animals, such as Tyrannosaurus rex and Apatosaurus louisae, and if all types with the same extension were identical then these types would be identical. But clearly, Tyrannosaurus rex and Apatosaurus louisae are two different types. Also, it does not help to define types in terms of their past and present extension, for the same problem may also arise for species that are not extinct. For instance, there are good reasons to distinguish between the genus Tarsius and the family Tarsiidae, so that Tarsius and Tarsiidae are different types even though both have, and always have had, the exact same extension.[2] Now if two types with the same extension were identical, the family Tarsiidae would be identical to its only genus. This is known as Gregg's Paradox (Buck and Hull 1966).

A further reason for not identifying types with their extension is that types of living beings constantly change with respect to their membership. New instances are born, others die. If sets were identical with their current extension, and if types were such sets, no type would last very long. Again, one might attempt to avoid this problem by defining types not in terms of their past and current extension, but in terms of the entire extension they had, have, and will have. For one thing, however, this would still not solve Gregg's Paradox. There may be good reasons for distinguishing a type from

---

[2] For the record, Tarsiers are small lemurs with big eyes.

a subtype even if they have always had and will always have the same extension. For another, until is extinct or the end of time has come (in which case types will be extinct), all we can do is to provisionally include future members in the set before they actually exist. This would require some sort of projection: We need to be able to determine, for a newly born instance or one yet to be born, whether it will belong to the type or not. This we can only do by applying substantive criteria, of the form "All items that will meet criteria C will be instances of T." However, this already involves the resources for a much more promising solution to the problem: to define types not in terms of their extension but in terms of a finite list of non-relational features that a thing must exhibit in order to qualify as a member of the type. I refer to such finite lists of non-relational features as <u>simple intensions</u>.

Sets that are defined and identified in terms of simple intensions are sometimes called classes (Gasking 1996, p. 144). They can gain and lose members, as long as the criteria of membership do not change. One possible worry about simple intensions is that after all, they tend to rest on extensions. For instance, Buck and Hull suggest solving Gregg's paradox by defining sets in terms of lists of features, but then they go on to identify these features with extensionally defined sets of items that have them (1966, p. 109 fn. 7). This two-stage procedure—to define a class in terms of features and a feature in terms of an extension—does not help much. Consider again the difference between Tarsius and Tarsiidae. Presumably, they differ in that the family, Tarsiidae, is defined in more general terms than its subtype, the genus Tarsius. Therefore, the corresponding simple intensions differ. The list of non-relational features that defines Tarsiidae is shorter. However, if features are defined in terms of their extension, two feature lists with the same extension should also be identical. Therefore, if features are defined in terms of extensions, the two feature lists only look different; in fact they are as identical as the two sets {a,b,c} and {a,a,c,b}. Therefore, one should not define features in terms of their extension. However, any other way of defining features may as well be used to solve the original problem, how to define types non-extensionally. The introduction of features does not get us any closer to a solution of this problem, and this means that we don't need it.

There are a few further possible problems with classes as types. Even though classes may change with respect to their extension, they are still defined in a way that precludes changes in the defining intensions. The defining feature list must remain the same if the class is to remain the same. However, types of living beings evolve, and this means that the defining features of their instances may change. Further, many biologists believe that once a species is extinct it cannot be brought back into existence (Hull 1976, p. 184). This means that even if living beings with the same features were to re-emerge, they would constitute a different type. An account that identifies species with classes (defined by simple intensions) leaves no room for this intuition. If something possesses the features that define a class, it is a member of it, regardless of whether and for how long the class has been empty before.

In order to avoid this latter problem, one would have to define the class in terms of relational features. For instance, one might demand that something is a beaver only if it has the features of a beaver and is a descendent of the first population of beavers on earth. Again, however, this would involve the resources for a more promising solution: non-simple intensions, which also include relational features.

Above all, there might be no set of non-relational features such that all and only the instances of a type share them. In particular, this situation can arise if the type is what Wolterstorff calls a norm-kind (1975, p. 129). Norm-kinds admit of atypical instances. If all instances of a type may lack some of the type's typical features, it may happen that even though there is one set of typical features, many of which are exhibited by each instance, there is no set of features that all and only the instances actually share. Each instance might be atypical by lacking a different feature, such that for each defining feature there is at least one instance that lacks it.

Whether any of these are actual problems depends very much on the logical structure of the feature list and the features on it. In order to deal with temporarily empty classes, we should turn away from simple intensions to intensions that include relational features. For instance, a criterion for type membership might be that all instances of type T must be descendants of the first Ts that existed (cf. Kitcher 1984,

p. 311). When a species is extinct, nothing can ever again have this relational feature, so that nothing will ever again qualify as a member. This would also solve the problem of evolving species. When a species evolves, there is at least one relational feature of its instances that does not change: all are still descendants of the same original population. The third problem, that there might be no set of features that all and only the instances of a type share, can be solved by weakening the requirement that all features on the list must be exhibited by all and only the instances of the type. We may give the feature list a more complex logical structure: We might define it in terms of a disjunctive feature list, or a conjunction of disjuncts, or whatever other construction might be needed to capture all members of a species.

2. Polytypic Classes and Clusters

Following Beckner (1968, p. 22), one may define the following two notions:

A class C is <u>polytypic</u> with respect to a list of features F iff
(1) every two items in C share a sufficiently large number of features from F, and
(2) each feature in F is exhibited by a sufficiently large number of members of C.

A class C is <u>fully polytypic</u> with respect to a list of features F iff
(1-2) it is polytypic with respect to F and
(3) there is no feature in F that all members of C share.

Polytypic classes have an intension with a complex logical structure, such that their members need not all exhibit the same set of features. Note that it depends on the feature list whether a class is polytypic, fully polytypic, or not polytypic at all. For instance, the class of beavers is polytypic, but not fully polytypic, with respect to the feature list "has sharp teeth, has fur, has lungs." It is fully polytypic with respect to the list "has fur, is male, is female." It is not at all polytypic with respect to "is triangular, has eight legs, has won the Nobel Prize." Therefore, that a class is fully

polytypic with respect to a certain list of features does not imply that there is no feature that all and only its members share. It only means that no such feature is on the list.

Unfortunately, the notions of a polytypic and fully polytypic class can only serve a limited purpose. Once we are given the entire extension of a class and a list of features, it is fairly easy to determine whether the class is polytypic with respect to these features or not. The problem is, however, that we often do not know the entire extension of a class. In cases where we do not already know the extension of a class, the information that this class is polytypic with respect to certain features does not enable us to determine its extension. For instance, in order to decide whether a newborn animal is an instance of a type, we would want a list of questions that we can answer in order to tell whether it belongs to the relevant extension or not. However, it is not possible to answer this question on the basis of a list of features and the information that the type is polytypic with respect to these features. If the class is polytypic, we do not know what to do with the items on the feature list that do not apply. That a feature is lacking might mean that the thing that lacks it is not a member of the class, or it may mean that it is one of the few members of it that lack the feature. To be sure, in real life we still know what to do in most cases, but the notion of a polytypic class is not what helps us do it. All it does is express the fact that we cannot decide by simply checking certain features.

Further, as long as only the number of shared features is important, it is hard to define polytypic classes so that they match given types of natural things. It is hard, for instance, to come up with a list of features such that the class that is polytypic with respect to these features includes all and only tigers. As Ben-Yami remarks, "[i]f in order to be a tiger an animal has to have only most of the properties we ascribe to tigers, then lions are tigers" (2001, p. 158). Therefore, most classes that are polytypic with respect to the features that typical tigers have will also include lions. On the other hand, it is easy to make up a feature list with respect to which fish, cats, and kitchen tables form a polytypic class: "has legs, has lungs, has no fur." There is no

clear correlation between polytypic classes and types. Therefore, it would not be enough to identify a type with a class that is polytypic with respect to some list of features. There is probably more than one such class, and most of these classes include members that do not belong to the type we began with. In order to get a grip on the class of instances of a type, we need more than the information that this class is polytypic with respect to a list of features.

A good way of defining types seems to be in terms of relational features of its instances.[3] Gasking suggests defining classes in terms of a focal instance of a type and a relation that holds between the focal instance and the other instances. According to Gasking, a <u>referent class</u> is a class that is defined by specifying a focal instance and a manner in which all other members of the class must be related to it (1996, p. 146). A <u>cluster</u> is a class, any arbitrary member of which could serve as a focal instance (p. 155). A cluster may thus have as many different definitions as it has members: Each member might be taken to be the focal instance, such that the cluster could be defined as the referent class of items that are appropriately related to this instance.

A lot depends on the relation in terms of which a referent class or cluster is defined. For instance, it may not be the case that all instances of a type are directly related to all other instances by a relation such as similarity or parenthood. All beavers are related to some other beavers by being their offspring, but they are not directly related to all other beavers in this way. However, it is easy to define a suitable relation. Gasking calls it a serial relation. A is <u>serially related</u> to B in manner R if there are a series of R-related items that connect A and B. On this basis, he defines biological species as chain-clusters of forms (1996, p. 178). A <u>chain cluster</u> is a cluster that is defined by a serial relation, and a <u>form</u> is a non-empty class that is defined in terms of a simple intension. A chain cluster of forms is thus a class C of classes, all and only the instances of which share certain features, such that any one of the classes in C could be taken as a focal instance, to which all other classes in C must be

---

[3] To be sure, Beckner also does this. He suggests defining biological species in terms of interbreeding and descent (1968, p. 60).

suitably related. Presumably, a biological species is a class of populations that are definable in terms of simple intensions, each of which may be taken as a focal instance P, such that all other populations in the species are serially related to P in some suitable way. By defining species as chain clusters of forms, Gasking can account for the evolutionary change of intensions. When a species evolves, it is possible to specify a relation that holds between populations that possess the old set of defining features and more recent populations with different defining features, such that the integrity of the species is preserved. Further, if this relation is serial, more than one such change may take place without necessarily leading to a new species.

3.   The Type Specimen Method

One reason why we classify items into types is that we can bundle information about them. For instance, we know that all (living) beavers have lungs. We know this about beavers in general, and we do not need to confirm it in each individual case. The main advantage of classifying something as a beaver is that we can apply our general knowledge about beavers in particular cases. This is a beaver, beavers have lungs, and therefore, this must have lungs (cf. Heuer 2008, p. 48 and 190). Such an inference would not be valid for features with respect to which a class is fully polytypic. Here we still have to confirm in each case whether the item has the feature or not; it does not follow from the fact that it belongs to the class. The same is true for referent classes. We may know the non-relational features of the focal instance and know what relation the other instances must have to this instance, but this need not imply much about the non-relational features of the non-focal instances (depending on the nature of the defining relation).

However, it is still possible to bundle information about the members of a polytypic class. Suppose that the list of features with respect to which a class is fully polytypic is fairly long and that all members of this class possess at least 75% of these features. If this is the case, we can save a lot of space and effort by classifying things as members of the class. Instead of describing each instance of the class in every

respect, we can start by listing, once and for all, the features with respect to which the class is polytypic and then confine ourselves to noting which of these features are absent in particular cases. This will render our descriptions of particular instances much shorter (cf. Farber 1976, p. 94). The same can be done for referent classes. We may give a full description of the focal instance and then only note for each of the remaining instances how they differ from it. We cannot simply infer the features of the particular instances from our general knowledge about the type because there are exceptions. Still, we can facilitate description and bundle our knowledge about the particular instances by relying on a general account and confining ourselves to noting the exceptions.

This is in fact what taxonomists in biology do. When zoologists describe a new species, they arbitrarily choose a focal instance, describe it, and assume that a relation can be specified that holds between this instance and all other instances. The focal instance is called "type specimen" or "holotype." For instance, the type specimen of Tyrannosaurus rex is part of the dinosaur exhibition at the Carnegie Museum of Natural History in Pittsburgh. It is the fossil labeled CM 9380 (formerly AMNH 973). According to the International Code of Zoological Nomenclature (ICZN), the name of an animal species is tied to its type specimen.[4] The type specimen of a species is (in general) the specimen on which the original description of the species is based, modulo some conventions about what counts as original here. Once a type specimen is assigned, the name must go where the type specimen goes. Should cats turn out to be robots controlled from Mars, as Putnam suspects (1975 vol. 1, p. 238), then the question whether Catus felis is a species of animals or robots would exclusively depend on one alternative: whether the particular Swedish cat that Linnaeus described under this name in 1758 was a robot or not. If this cat was a robot, the name "Catus felis" applies to robots, regardless of the nature of any other animal that we call that name.

In the present context, two things are especially noteworthy about the type

_____

[4] ICZN ch. 13, §61, as of 2000. See http://www.iczn.org.

specimen method. First, the type specimen need not at all be typical for its species. To be sure, the idea of assigning a type goes back to a tradition within which it was assumed that types should be defined in terms of their most typical instance and some kind of similarity relation, such that the least typical instances would turn out to be most distant from the type.[5] This is of course the most economic way of bundling information: Assuming that the typical features of a type are the features that most instances have, the most typical instance makes for the best type specimen. However, it is not always easy to figure out which specimen is most typical, especially when a species is first described; and once a type specimen has been selected, it is best to stick to it whether it is typical or not. Therefore, the only features of the type specimen that actually matter in modern taxonomy are its spatiotemporal location, its name tag, or its number in a collection, and these are precisely the features that it does not share with any other instance. In modern taxonomy, the type specimen need not satisfy, let alone incorporate, any standards of typicality. It is subject to such standards and may fail to meet them as much as every other instance of its type.

The second thing to note is that for many known and well described species no type specimen has ever been named, while for many others the type specimen is now lost or destroyed. For many common types such as the European beaver, the domestic cat, and other species described by Linnaeus (1758), no type specimen has ever been preserved. There once was a type specimen, to be sure, since the type specimen is the instance of a kind on which the first officially published description is based, and Linnaeus based his descriptions on observations of some particular beaver and cat. But these particular animals were never collected, preserved, or labeled, and they are long dead. There is no way to refer to them other than "the animal Linnaeus described." All we know is their so-called type locality, which is Sweden. This makes it difficult to literally apply Gasking's method of defining a referent class. Since the type specimens of species like Catus felis and Castor fiber are nowhere to be pointed

---

[5] Whewell writes that a type is "an example of any class ..., which is considered as eminently possessing the characters of the class" (1847 vol. I, p. 494).

at, they do not make for good focal instances. Therefore, a definition such as "x is a European beaver if and only if x is related in way R to the type specimen" would be of no avail.

It is remarkable that this does not bother taxonomists much. If they had a problem with missing type specimens, one would expect them to retrospectively assign some wherever possible. In fact, however, this is made almost impossible by the current codes of nomenclature. The type specimen may only be re-assigned if it is <u>really</u> important,[6] and in most cases it is not really important. The type specimen only serves a rather abstract semantic function: It separates the description of a species from its name. The name does not attach to whatever satisfies a certain description but rather to whatever the first properly published description of a species was a description of. If the original description turns out to be wrong, it does not any longer matter what it says. All that matters is what it purported to describe.

In theory, an animal or plant is thus of type T if and only if it is of the same type as the type specimen of T. If one wanted to determine beyond doubt whether a given living being A is of type T or not, one would therefore have to accomplish two tasks. First, one would have to locate the type specimen for T, if there is one. Second, one would have to check whether A and the type specimen are of the same type. The aim of the biological codes of nomenclature is to help complete the first of these tasks. These codes establish rules and procedures concerning the description, collection, and preservation of type specimens, such that when all goes well, it is always possible to locate the relevant type specimen. The second question is the more difficult one, however, and the codes for nomenclature do not provide the means for answering it. They do not tell us when two items are of the same type.

On the face of it, it may seem that when we describe a species as a referent class, we need not mention any non-relational features of their instances. All we need to do is to name a type specimen and a relation that all other instances bear to it, and thanks

---

[6] Winston (1995, p. 301) cites a case where the type specimen turned out to be a composite of three different animals. In this case, it had to be replaced by a new one.

to the codes of nomenclature, we need not even rely on any substantive description of the type specimen (other than, e.g., "the item labeled CM 9380"). However, as I will show momentarily, non-relational features play an indispensable role when it comes to specifying the relation of sameness of type, which all instances of a type must bear to the type specimen. I focus on two criteria for sameness of type that have been proposed for biological species: the morphological and the biological species concept.

4.   Two Species Concepts

Gasking defines morphological species as reference classes of items that are serially sufficiently similar to a focal instance (1996, p. 150). Presumably, the relevant kind of similarity is similarity with respect to morphological, non-relational features. It is sometimes argued that such morphological criteria are mistaken because they do not always pick out the same groups of items as other criteria, such as the biological species concept (Mayr 1996, p. 269). This is of course a circular argument; one might as well argue that the biological species concept is wrong because it does not pick out the same groups as morphological criteria. However, there are also independent arguments against morphological criteria for species membership, some of which should sound familiar by now.

To begin with, at least some of the more straightforward notions of similarity involve lists of shared features. Tversky, for instance, defines similarity in terms of intersecting sets of features, such that two items are more similar to each other than two other items if they share more (non-relational) features than the other two (1977, p. 330). This notion of similarity will not do in the present context. We are here concerned with types that may not be definable in terms of a finite list of non-relational features that all their instances share. If we define a type as a referent class of items that are similar to a type specimen, and if similarity is defined in terms of shared features, we are back where we started. For there are only two possibilities. Either there is after all a specific set of non-relational features that all and only the instances of the type in question share, and then we might as well define the class in

terms of a simple intension. Or, on the other hand, we can define the type as the set of things that share more than a certain number of features with the type specimen. This would amount to defining the type as a class that is fully polytypic with respect to the features of the type specimen. That this would not work, as a definition, has already been shown. Most classes of things that share a sufficiently large number of features with the focal tiger are not classes of tigers, and we still need to know how to pick out the right one of these classes. It follows that one can only effectively define a type in terms of Tversky's notion of similarity if one can also define it in terms of a simple intension.

Moreover, relations other than morphological similarity seem to be more important in many cases. Consider a hypothetical case of two reproductively disconnected biological species, A and B, and let A evolve. Call its earlier stage A1 and its later stage A2. Now it is possible that according to all conceivable morphological criteria, instances of A2 are more similar to instances of B than they are to instances of A1. According to morphological criteria, A2 and B would then constitute one species, and A1 another. According to our intuitions, however, A1 and A2 are one species, and B is another. This is, again, a somewhat circular argument. If morphological criteria were right, A2 and B would indeed constitute a single species and there would be no problem, except that our intuitions would tell us otherwise. However, we do in fact prefer putting items in one biological species when they are related by reproduction and evolution and not when they are merely similar. For at least some types, that is, we prefer non-morphological criteria. The biological species concept aims to provide such criteria.

Using Gasking's vocabulary, one might state the biological species criterion as follows: Two living beings belong to the same species if they serially potentially interbreed. Two organisms interbreed when they sexually produce living offspring, they potentially interbreed if they are able to do so, and they serially potentially interbreed if there is a chain of organisms connecting them, each of which is able to interbreed with the next one in the chain. This is very close to what is usually referred

to as the biological species concept, and there are at least three problems with it. The most obvious problem is that a species concept that is defined in terms of interbreeding organisms does not apply to asexually reproducing organisms. Mayr notes this and suggests that it is best to use the biological species concept for sexually reproducing organisms and other species concepts for others (1996, p. 266).[7] This solution may be ad hoc, but it is a solution. A second problem with the biological species criterion is that in order to determine species membership, one has to determine whether organisms can interbreed, and one would think that the best way of doing so is to perform breeding experiments. However, this is hardly ever done by taxonomists. Instead, it is in most cases simply assumed that two organisms can (serially) interbreed, either because they obviously belong to the same species or because they are genetically (that is morphologically) similar.[8] Further, nothing would guarantee that breeding experiments establish species boundaries where we want them, for it is known that animals that do not interbreed in nature may do so in captivity (Dobzhansky 1935, p. 350). The third problem is perhaps less obvious but most important in the present context. The biological species concept is defined in terms of sexual reproduction, and reproduction is a success term. In order to say whether two organisms potentially interbreed, one must first say what would happen if they would successfully interbreed, and in order to say this, one must say what the

-----

[7] Franklin-Hall argues that although the biological species concept does not apply to bacteria, there is a sense of "sex" in which bacteria do sexually reproduce (2008, ch. 1). If she is right, the biological species concept does not apply to all sexually reproducing beings but only to some of them.

[8] The DNA of a living being may be causally more fundamental than its other morphological features, but it remains one of its morphological features.

result of successfully interbreeding would be like.[9]

There does not seem to be one defining feature that strictly all processes of reproduction share, except the one that they involve some interaction among members of a species that leads to the generation of a further member of the same species. In order for this to be successful, the offspring should at least be alive (for a while). However, there is no such thing as simply being alive. To be alive is in each case to lead a life of a certain kind, and different kinds of living beings lead different kinds of lives. It is rather unlikely that there is a feature that all living beings necessarily share so that one could determine whether something is alive without knowing what kind of life it would lead if it were alive. Further, even if there were such a feature, it would probably not be enough for the offspring of a beaver, for example, that it merely has this feature. Suppose, for instance, that all and only living beings feed upon negative entropy (Schrödinger 1945, p. 71). I very much doubt that a being would qualify as an offspring of a beaver if it results from the copulation of two beavers and does nothing more specific than feeding on negative entropy. Whatever life is, it involves some kind of homeostasis, i.e. maintenance of the same form throughout time; and in order to say whether a form remains the same, one must know what this form is. Therefore, a description of what it means for a given living being to be alive is already a rudimentary description of its life form (cf. Thompson 2008, p. 50-1).

This explains, to some extent, why the biological species criterion works. Part of the reason why interbreeding organisms belong to the same species is that we speak of interbreeding only if two members of one species manage to produce a third

---

[9] Putnam notes that "the characteristic of being cross-fertile with a lemon presupposes the notion of being a lemon" (1975, p. 240). The type specimen method eliminates one way in which this notion is presupposed, for one can refer to the type specimen without describing or classifying it as a lemon. It does not eliminate the second, more subtle way in which being cross-fertile with a lemon presupposes the notion of being a lemon: In order to determine whether sexual reproduction is successful in this case, one must know what lemons are.

instance of the same species. Sameness of species is already presupposed. This also means, however, that the biological species concept presupposes at least a part of what it is supposed to provide. It presupposes necessary conditions under which organisms belong to the same species.

We have reached the following result, concerning both the morphological and the biological species concept. In order to effectively define a species as a referent class of items that are morphologically similar to a focal instance, one must produce a morphological description of its members. In order to define a species in terms of sexual reproduction, one must provide a rudimentary description of the expected outcome of successful reproduction. These results are two instances of the following general pattern.

If we define a type as a class of items that are related to a focal instance by a relation R, the unity of this class hangs on the defining relation, R. If R involves breeding, reproduction, copying, or any other kind of replication, then its definition involves a reference to criteria of success. A specification of these criteria, in turn, must involve a description of the result of successful replication, and this description must involve non-relational features. Therefore, it may seem as though our definition ultimately reduces to one in terms of these non-relational features. It may seem that instead of defining the type as the class of items that are related to one another by a relation that holds between items only when they have certain features, one might as well simply define it as the class of items that have these features.

Note, however, that non-relational features play a very different role in clusters and referent classes than in classes that are directly defined in terms of simple intensions. It is not assumed that all and only the members of a referent class or cluster actually share certain features. The features that define successful replication occur in a counterfactual context. Instead of describing what is the case, we describe what would be the case if all went well. Two organisms interbreed if they engage in an interaction that would, <u>if successful</u>, lead to an item with certain features. When we define a species in terms of copying or reproducing, we do not actually say that all its

instances must have certain features; we only say that the process from which they result qualifies as successful to the extent that they have certain features. Obviously, not all reproduction is successful reproduction, and therefore not all instances of a type need have the features that a result of successful reproduction would have to have. They must possess these features only in order to qualify as <u>typical</u>. That they instantiate their type means that they qualify as typical under certain conditions.

Accordingly, there are two different thresholds involved in replicating a type: one above which the replicas are good copies and one below which they are not copies at all. These two thresholds may be quite independent from one another. To be sure, they might only differ by degree, so that in order to qualify as an instance, an item must satisfy fewer of the same requirements that it would have to meet in order to be fully typical. Kant, for one, assumes that type membership and degree of typicality are both determined with a view to the same focal instance (which he refer to as the "Normalidee," Critique of Judgment §17, B 56). This, however, need not be the case. That something is a result of a process that is governed by certain standards does not at all imply that this thing actually meets any of these standards. Therefore, the criteria for sameness of type may differ considerably from the standards by which we measure the degree to which the instances of this type are typical.

This is a very important point, so I may be excused for dwelling on it. Whether something is a T and how good a T it is are two different questions. The criteria for being subject to a standard are not the criteria for satisfying this standard. Not every T need be a good one, and a very bad T may look much less like a T than something that is not a T. Even a perfect T may look very much like something that is not a T. What matters is ultimately not how something looks but what standards of success govern its coming to be. Accordingly, there are two quite different ways of not being a T. First, something is not a T if it is not the result of a kind of process that typically results in a T. Second, something is not a T if it results from a process that typically results in a T but this process failed (miserably).

Note further that when the defining relation involves reproduction, the relevant

standards of success need not be fully satisfied by any one instance of the respective class. Reproduction is not just copying; its success is not proportional to the resemblance between parents and offspring. The parents of a living being may actually be much less typical than their offspring. This implies what we already know: The standards of typicality that apply to a type are not set by its type specimen, and the instances of this type are not typical to the extent to which they resemble the type specimen. They are typical to the extent to which the process by which they came to be satisfies the criteria of success that are associated with the type that the type specimen instantiates.

5. Conclusion

The features of the result of successfully reproducing a type need not be features that all instances of this type share; there need not even be an actual instance that possesses all of them. The reason why I say this in so many words is that it leads us right back to Aristotle (or even further, to Plato). The result of successful reproduction is Aristotle's formal cause. It is not a particular instance of a type but rather what such an instance would be if it would develop according to its nature and all went well. Its features are not the features that all instances of its type share. Therefore, an account of the formal cause of a thing is not an account of the actual features of this thing. It is an account of the features that a typical instance of its type would have.

Therefore, that something is an instance of a type need not mean that it has certain actual properties (other than the vacuous property of instantiating this type). Something is of a type to the extent to which it is subject to certain standards of typicality, not to the extent to which it satisfies them (cf. Wolterstorff 1975, p. 128). Whether something is subject to a standard and whether it satisfies them are different questions, but in order to answer any of them one needs to know the standards in question. Therefore, an account of a type of natural things must in any case involve a specification of standards of typicality. What Aristotle calls the essence or formal cause of a thing provides such a standard of typicality. When we describe the formal

cause of a natural thing, we do not describe its actual features but rather specify the standards of typicality that apply to it. This is why Aristotle refers to the formal cause as a paradigm (Physics II 3, 194b26). It is what the thing would be if it were fully developed and typical. That the thing has this formal cause means that it is subject to the standards that are prototypically satisfied by a paradigmatic instance of its type.

Note that the description of a standard of typicality is at the same time a description of what satisfies these standards. Therefore, the description of the essence of a type is in fact a description of its typical instance or instances. It is only a small step from here to saying that the formal cause of a thing, which provides an account of its essence, is a fully typical generic instance of its type. The formal cause has all the properties that its typical instances have.[10] This is why Aristotle refers to types as second οὐσίαι: They are not particular but generic substances (cf. Stekeler-Weithofer, unpublished, § 2.2).

The account of the essence or formal cause of a natural thing takes the form of a set of generic propositions, which do not state what always, usually, or most often happens but what happens in typical cases.[11] There are circumstances in which there are actually no typical instances of a type, such that the account of its essence is not a description of any of its real instances. Nonetheless, the description of the typical case must be based on knowledge that we have about real cases. If no or only a few typical instances actually exist, an argument must be given as to why the respective standards still apply.

In this chapter, I have shown that the best way to give an account of a type of

---

[10] Hugly and Sayward say that whereas token sentences are perceptible, "[t]he sentence itself is imperceptible" (1981, p. 181). This is misleading. The description of the type "sentence" is certainly a description of something that is visible. The properties of a typical instance are also properties of the type itself.

[11] Cf. Moravcsik (1994) and Johnston (1981, p. 291). For more on generic propositions, see Carlson and Pelletier (1995), (Thompson 2008, ch. 4), and Foot (2001, p. 27-37).

natural things is to specify the standards of success that govern the process by which its instances come to be. Further, the best way to specify these standards is to describe what satisfies them, i.e. what would typically result from this process. In the next chapter, I explore the idea that the formal cause (or essence) of a natural thing is not a property but a paradigmatic instance of its type or a range of such instances.

Essences vs. Properties

In the beginning of the previous chapter, I pointed out that according to Aristotle, the definition of a natural thing defines the typical result of its natural development. I have then argued, on independent grounds, that types of natural things are best defined by describing the result of successfully replicating their instances. In this chapter, I explore some implications of this idea regarding the notion of an essence. I begin by asking whether the essence of a thing should be identified with an explanatorily fundamental property of this thing. I deny this, because the essence of a thing is not explanatory in the right way and because it is not one of its properties. Then, after briefly comparing essences to sortal universals and types to natural kinds, I further explain the difference between essences and properties. I do this by drawing a distinction between acts of identifying an item as an instance of a type and acts of describing an instance of a type as having certain properties. This distinction is a cognate of the distinction between casting and stating that I have introduced in Chapter 2.

Aristotelian essences are often taken to be properties or sets of properties that all instances of a type necessarily share. More specifically, they are sometimes described as properties such that everything that has such a property in one possible world must have it in all other possible worlds in which it exists (Barcan Marcus 1971, p. 198; Brody 1973, p. 359). As an account of Aristotelian essences, this is wrong for a number of reasons.[1] One simple reason is that some necessary properties are what Aristotle calls propria (ἰδίαι), but propria are not essences (Topics I 5, 102a18-30; cf. Topics V). Therefore, for something to be an essence, it is not sufficient that it is a necessary property of all instances of a type. Charles (2002) therefore claims that an Aristotelian essence is a special sort of necessary property that "is the one cause of all

---

[1] See also Kung (1977) and Matthews (1990). Witt (1989, ch. 6) also contrasts Aristotelian and modern essentialism.

the kind's derived necessary properties" (p. 203). I argue against this suggestion in different stages. I begin by showing that an essence cannot be a single necessary property of a thing that explains all its other necessary properties. In addition to explanatorily fundamental properties of a thing, any explanation of its further features must refer to external constraints. Then I point out that although essences and explanations are mutually interdependent, essences are not explanatory. They are, rather, what is explained. Finally, I argue that essences are not explanatorily fundamental properties because they are not properties at all. I support this latter claim by showing that essences must be more fundamental than properties because we must be able to identify a thing in terms of its essence before we attribute a property to it.

1.   One Property to Rule Them All

Charles does not only assume that Aristotelian essences are explanatorily fundamental properties but also that the essence of a thing must be one single property of this thing that explains all other properties that instances of its type must have (2002, p. 203).[2] He extracts this from two passages in Posterior Analytics II 6. In the first, Aristotle argues that an account of what a syllogism is cannot be used as one of the premises in a syllogism and that the same is true for essences (τὸ τί ἦν εἶναι, 92a11-19). Charles takes this to mean that a definition must be based on some kind of deduction that does not involve this definition (p. 189). In the second passage, Aristotle writes that the object of a definition must not be an accidental unit, such as "musical and literate man" (92a29-33). From this, Charles infers that the demonstration on which a definition is based must start from one basic feature that explains all the other necessary features of what is defined (p. 195). It should be obvious that both passages admit of alternative readings. What Aristotle wants to say

—————————————

[2] Against this, Lennox (unpublished) points out that Aristotle often explains the typical properties of a living being not in terms of one underlying property but rather by reference to a certain way of life (βίος).

in the first passage might just be that definitions belong to the framework within which something is demonstrated and not to the content of a demonstration. An essence may be something by means of which we can get to know and explain facts, but it is not a further fact that could figure as a premise in such explanations. In the second passage, Aristotle basically says the same as in Metaphysics Z 4 (1030a10-11): A formula by which a thing is said to be something else is not a proper definition. By assumption, being musical and being literate are independent features of a man: He might only be musical or only be literate or none of them or both. When two features occur in the definition of a thing, they must not be independent in this way. This does not imply that there must be one single feature that explains all other features.

According to the view that Charles attributes to Aristotle, an essentialist should be able to explain all necessary properties of a type by referring to exactly one property that all and only its instances must have. It is no wonder that by these standards Aristotle ultimately fails to be an Aristotelian essentialist (p. 336). For Aristotle often refers to several independent factors in order to account for the necessary properties of a thing. If Charles were right, only one of these factors could be an essence, and Aristotle would thus fail to stick to his essentialist agenda. He would not explain all necessary features of a type by exclusively referring to its essence. However, the problem with an essentialist agenda, as Charles describes it, is that it cannot possibly be accomplished. This can be shown as follows.

Some types can only be instantiated by items that fully satisfy all relevant standards of typicality, which is to say that their instances cannot be more or less typical, and the relevant standards of typicality do not differ from the criteria for type membership. The instances of such types coincide with their essence because they cannot possibly fall short of it. Things that coincide with their essences cannot be explained by referring to their essence since nothing can be explained by reference to itself (Metaphysics Z 17, 1041a14-15). Material things, on the other hand, do not coincide with their essence (Metaphysics Z 11, 1037b3-4) and may therefore be

explained by reference to their essence. Their essence can be a cause. As Aristotle writes, the formal cause of such things can figure as the middle term in a demonstration (Posterior Analytics II 8, 93a6-8).

Things that coincide with their essence do not make for good targets of an essentialist agenda because they cannot be explained in terms of their own essences.[3] On the other hand, for things that do not coincide with their essence, the essentialist agenda can also not be completed. One cannot explain exclusively in terms of their essence why they may fall short of their essences. This may seem like something that does not require an essentialist explanation, but in fact many typical features of living beings can only be explained by reference to the fact that they may fall short of their essence. Therefore, that they may fall short of their essence is essential for them. For instance, beavers typically build dams. This belongs to their nature. One may explain it by referring to their essence: In order to be what they are, they need a certain water level, and building a dam is a way of achieving this. However, this is not a full explanation of why beavers build dams, for there would be no need for dams if the water level were already high enough. The explanation why beavers build dams must therefore involve two facts: Beavers need a certain water level and often the water is not high enough. However, the second factor in this explanation is necessarily an external constraint. It cannot belong to the essence of a beaver that the water level is often not high enough. Nonetheless, without such external constraints, the typical behavior of beavers would be pointless. Therefore, the typical behavior of beavers cannot be explained exclusively in terms of the essence of beavers.

2.   Essence and Explanation

The essence of a thing cannot be a single one of its necessary properties that explains all other necessary or typical features of this thing. Still, essences might be

---

[3] Whether this is an inescapable conclusion does not actually matter in the present context. Here we are still primarily concerned with natural things and processes, which do not coincide with their essences.

explanatorily fundamental properties. Charles writes that "[w]hat it is to be something (as captured in the definition) and the basic relevant cause are one and the same" (p. 200). And indeed, Aristotle seems to say that essences are explanations in Posterior Analytics II 2. The essence of a thing is what this thing is (τί ἐστι), and an explanation of a thing tells us why it exists (διὰ τί ἐστιν). In II 2, Aristotle writes that in certain cases, the questions τί ἐστι and διὰ τί ἐστιν amount to the same (90a15; cf. II 8, 93a3-4).

However, this does not imply that the essence of a thing explains why this thing exists or why it is what it is. As I point out in the Introduction, the question διὰ τί need not be a request for an explanation. In general, διὰ τί asks for something that accounts for a thing and makes it possible or intelligible, and asking "Why?" is only one way of doing this. Other ways of asking διὰ τί are: How did this happen? Who did this? What's the point? What does it take to be this kind of thing? That is, the question διὰ τί ἐστιν may ask for an explanation of why something exists, but the sense in which it does so need not be the sense in which it is the same as the question τί ἐστιν. When we say what a thing is, we answer the question διὰ τί by giving an account, but this account need not explain why the thing exists. For instance, the question of whether the moon is eclipsed may be answered by pointing out what an eclipse is—namely the interposition of the earth between the sun and the moon—and then showing that the earth is indeed so interposed (II 8, 93b7-9). This answer does not explain why the earth is interposed between the sun and the moon. Therefore, since an eclipse is an interposition of the earth between sun and moon, it also does not explain why the eclipse occurs. Still, it says on account of what the eclipse is occurring, for it justifies the assertion that it is occurring. In this sense, the essence of a thing may explain what this thing is without explaining why it is what it is.

Eclipses are processes and not things, and it might be misleading that Aristotle refers to them in his examples. As will be shown in Chapter 8, essences of causal processes must be complex, such that one may explain or predict their later stages by referring to their earlier stages. One of the parts of such a complex essence is a cause

of the other part. Thunder, for instance, may be defined as noise in the clouds that is caused by the extinction of fire in the clouds. This definition does not only specify the form of thunder, it also provides means by which one may causally explain why thunder occurs. It is caused by the extinction of fire in the clouds. In this sense, the definition of thunder is analogous to a syllogism that shows that thunder occurs (II 10, 94a13-14). However, this does still not mean that the essence of thunder is explanatory.

Aristotle writes that in one sense, the essence of thunder includes its cause, and in another, it is what may be explained by reference to this cause (II 10, 94a11-14). Thus in one sense, the essence of thunder is causally complex: It is noise in the clouds as caused by the extinction of fire in the clouds. In another sense, the essence of thunder is only noise in the clouds, and the extinction fire in the clouds is its efficient cause. In neither case is the essence of thunder the same as its efficient cause. The only thing that is an efficient cause here is the extinction of fire in the clouds, but this is either only a part of the essence of thunder or not its essence at all. The essence of thunder may have an efficient cause or an internal causal structure, but this does not mean that it is the basic efficient cause of the necessary features of thunder. When we show that thunder occurs by showing that noise is caused by the extinction of fire in the clouds, we show that something occurs that fits the definition of thunder (van Fraassen 1980a, p. 28). Here, the essence of a thing is its "basic relevant cause" in the obvious sense: It is its formal cause. As such, it answers the question διὰ τί. It is not a cause in the sense that Charles has in mind, which would efficiently or "teleologically" cause the necessary features of a thing (p. 293).

Charles sometimes says, more moderately, that essences and explanations are mutually interdependent. This is indeed so, and it can be confirmed as follows. As Aristotle says, we do not take ourselves to really know a thing unless we know its cause (II 11, 94a20). This is also true for essences. Where the essence of a thing admits of an explanation, we do not really know this essence unless we know its explanation; and since we cannot explain a thing in terms of itself, this explanation

must be given by reference to a cause other than the formal cause (which is the essence itself). If we speak of the essence of thunder in the second of the two senses distinguished above (as noise in the clouds), the extinction of fire in the clouds is not a part of it but only its cause. Still, the essence of thunder, noise in the clouds, cannot be understood without knowing this cause. Not all noise in the clouds is thunder; thunder is only such noise in the clouds that is caused by the extinction of fire in the clouds. Conversely, the causal explanations in question must involve essences (Charles 2002, p. 217). The reason for this is that explanations of essences must hold true at a generic level. There may be cases in which extinction of fire in the clouds accidentally causes noise. (An airplane catches fire, Superman extinguishes it, and the passengers applaud.) Such noise is not thunder because it is only accidentally caused by the extinction of fire in the clouds, and the opposite of accidentally is essentially. (Applause is not essentially caused by the extinction of fire.) This implies that the explanation of an essence must be an explanation in terms of further essences. It does not imply that essences are explanations or explanatorily fundamental features.

3.   Essences, Properties, and Essential Properties

Charles assumes that all instances of a type necessarily share a set of properties in the first place (2002, p. 195) and that their essence is one of these necessary properties. However, as I argue in the previous chapter, there need be no non-relational property that all and only the instances of a type share. Still, all and only the instances of this type must share exactly one essence. Therefore, essences should not be taken to be (non-relational) properties.

There are several ways in which essences differ from properties. It is sometimes said that essences and properties differ in that they provide answers to two different questions (cf. Loux 1991, p. 21). The first of these questions, to which essences provide an answer, is what a thing is. The second question, to which properties provide an answer, is how a thing is or what it is like.

Further, essences differ from properties in that they may stand in for what they are

essences of, whereas properties cannot. Aristotle points out that the definition of a thing, which states its essence, may be predicated of all instances of its type (Topics VII 5, 150a37-b1). Any reference to thunder may thus be replaced by the phrase "noise in the clouds (as caused by the extinction of fire)." This, however, is not true for properties and their definition. We cannot replace all occurrences of "white" by the definition of whiteness (cf. Categories 5, 2a27-34). Socrates is white, white is a color of a certain kind, but Socrates is not a color of this kind.[4]

Moreover, it has been argued that essences cannot be properties in the sense that they attach to given things, because the essence of a thing is constitutive for it, so that without the essence there would be no well defined thing in which it could inhere as a property in a subject.[5]

Finally, the essence of a thing is what results from its natural development, and as I point out in Chapter 1, no thing develops into a property. Therefore, Aristotle says in Physics II 3 that the formal cause of a thing is the whole and its composition (τό ὅλον καὶ ἡ σύνθεσις, 195a20-21), such as the syllable that is composed of letters, a thing that consists of matter, or something that is composed of elements.

That essences are not properties does not mean that there are no essential properties. Essential properties are properties that belong to essences. Since there need not be a set of properties that all and only the instances of a type share, it follows that essential properties are not properties that all and only the instances of a type share. Essential properties are not even properties that all (if not only) the instances of a type must share. Therefore, essential properties are not necessary properties. A fortiori, they are not necessary properties of a thing that explain further properties of it.

A first clue how the essential properties of a thing differ from its ordinary properties may be found in a remark that Aristotle makes in Categories 5. He writes

---

[4]According to Dancy, "white" applies to Socrates only paronymously (1975, p. 361).

[5] Cf. Woods (1967, p. 237-8), Modrak (1979, p. 375), Frede and Patzig (1988, Einleitung, p. 44), Witt (1989, p. 121-2), and Rapp (1995b, p. 311).

that substance (οὐσία) does not admit of more or less (3b33-34). Porphyry says the same about differentiae: "Differences in their own right do not admit of more or less, whereas accidental differences, even if they are inseparable, take augmentation and diminuition" (Isagoge §3, 3a43-45, tr. Barnes). Boethius further claims that propria, such as two-leggedness and the ability to laugh, do not admit of degrees. Let me here focus on Boethius.

Differentiae and propria are essential properties. Humans are essentially rational, capable of laughter, and two-legged. According to Boethius, this means that there can be no degrees to which they are rational, capable of laughter, and two-legged. Now of course, some humans are more rational than others, and some have only one leg. Boethius maintains that still, all humans are equally rational and two-legged because they all have the same potential of being so (in Isagogem ed. prima II 22, CSEL 48, p. 122). However, when Boethius claims that humans have the potential to have two legs, he cannot seriously mean that a one-legged human has the power to grow an extra leg. All he can mean is that the one-legged human is two-legged in the sense that she still is an instance of a type of animal that generally has two legs. This must therefore be the meaning of "potential" here: Humans are potentially what humans in general are (e.g., two-legged). What humans in general are is also what each of them essentially is.

This also means that the extent to which humans are actually (as opposed to essentially) rational and two-legged does admit of degrees. The essential properties of a thing are not its actual properties. That humans are essentially rational and two-legged means that humans in general are rational and two-legged, and "general" has a normative sense here. It means that all humans are subject to the same standards of rationality and health. Some humans may not be rational, but to the extent they are not they fail to meet standards of rationality that equally apply to all humans. When humans fail to have two legs, they fail to meet standards of health.

Now for a thing that is subject to standards, there may be differences in degrees at three different levels. First, a thing that is subject to a standard may satisfy this

standard more or less well. This happens when humans are more or less rational. Second, a standard may apply to something to a greater or lesser extent. The standards of syntax apply to prose more rigidly than to poetry. If humans are essentially rational, this is not true for humans and the standards of rationality. These standards apply to all of them to the same extent. Third, the standard itself may leave room for degrees, such that there is a continuum of equally good ways of meeting it. This might well be true for humans and the standards of rationality and health, for there may well be different ways of being perfectly rational or healthy. When Boethius says that for humans, rationality does not admit of degrees, he excludes only the second case. All humans are equally subject to the standards of rationality. It is still open whether all humans actually are rational and whether there is only one way of being perfectly rational.

Aristotle, Porphyry, and Boethius may not have realized that the third kind of degree is as admissible as the first—that is, that a standard may allow for a continuum of different ways of satisfying it. They seem to think that there can only be one way of fully meeting the standards of typicality for a type. This leads to the idea of a single archetype for each kind and to consequences such as the following: Where the male and the female form of a type systematically differ, one of them must be defective (De Generatione Animalium I 3, 737a27-29). However, such consequences do not actually follow from the idea that essences are or provide standards of typicality. There is nothing wrong with standards of typicality that leave room for degrees in the third sense. For instance, female lions are typically larger than male ones. Therefore, an account of the essence of a lion that leaves room for the possibility that female lions are as typical as male ones will also have to leave room for differences in size. Still, this does not mean that different lions are subject to the relevant standards to different degrees, and this is the sense in which essential properties and essences do not admit of degrees.

We see now more clearly that essences are not explanatorily more fundamental properties, because essences are not properties at all. Instead of being more

fundamental properties, essences are more fundamental than properties. Further, the essential properties of a thing need not be its actual properties. They need not be properties of this thing; they are properties of its essence. They may be properties of the thing but only insofar as the thing matches its essence or coincides with it. This also means that the fact that essences and explanations are mutually interdependent plays no role in distinguishing between the essence of a thing and its properties. In order to distinguish between essences and properties, we do not need to order the properties of a thing according to their explanatory value. We need to understand how an essence can be something other than a property. Before turning to this question, I discuss two related notions: the notion of a sortal and of a natural kind.

4.   Sortals and Natural Kinds

The distinction between essential and actual properties that I have just drawn is not the same as the distinction that is sometimes made between sortal and characterizing universals.[6] Sortal universals (or sortals, for short) are traditionally taken to be universals that satisfy three conditions. First, they supply a principle for distinguishing and counting their instances (Strawson 1959, p. 168). The universal "white item" is not sortal because it is not possible to count all white items in a given area. "White mouse" is sortal, because it is possible to count all white mice in a given area. Second, the instances of a sortal are such that neither their parts nor that of which they are a part are instances of the same sortal (p. 205). This is often referred to as the mereological criterion. Again, "white item" is not a sortal because the parts of white items are white items, and white items are often parts of white items. On the other hand, "white mouse" is, because white mice are not parts of white mice. The third way of distinguishing between sortal and characterizing universals is the most general one: The sortals that apply to a thing answer the question what it is rather than how it is (Feldman 1973, p. 269). I refer to this as the what-criterion.

_____

[6] Cf. Wiggins (2001, p. 9). Wallace also writes that the notions of a sortal and of a second substance are "different but closely related" (1965, p. 8-9).

As Feldman points out, these three criteria are independent of each other (1973). In particular, there is no good reason why the question of what a thing is should always be answered by a universal that does not apply to its parts. We may ask of a portion of milk what kind of milk it is and be told that it is goat milk. This is an answer to the question of what it is, but it also applies to its parts and possibly to that of which it is a part. According to the what-criterion, "milk" is a sortal universal: Milk is what a thing is, not how it is. According to the countability test and the mereological criterion, however, "milk" is not a sortal. One may count glasses or liters of milk, but not just milk, and the parts of milk are (usually) also milk (cf. Rapp 1995b, p. 214).

This means that Strawson's division between sortal and characterizing universals is not exhaustive. He writes that the phrases "is an instance of" and "is characterized by" are "properly followed by, respectively, the designation of a sortal universal and the designation of a characterizing universal" (1959 p. 172). Since "milk" may properly follow "is an instance of" but not "is characterized by," it might thus seem to be a sortal universal. Later on, however, Strawson adds that terms like "milk" ("snow, water, coal, and gold") are neither characterizing nor sortal universals (p. 202). In any case, whether milk instantiates a sortal universal or not, it seems to have an essence. One may identify something as an instance of milk as opposed to attributing a property to it. Since the distinction between sortal and characterizing universals is tied to the countability test and the mereological criterion, it is therefore better to avoid this terminology in the present context. What we need is a notion of essence that is also applicable to such things as milk, gold, and water.[7]

Something very similar is true of the notion of a natural kind. There are many different accounts of natural kinds out there, some of which are of no use in the present context. In order to distinguish between useful and useless notions of natural

_____

[7] Aristotle furnishes stuff like gold with an essence (Matthews 1990, p. 254; Loux 1991, p. 32). Also, as he remarks in Topics I 7, different samples of water are the same in kind (103a14-18), so there must be kinds of stuff like water, and essences corresponding to them.

kinds, we have to say what essences are without relying on any one of these accounts. Therefore, the notion of a natural kind is not of much help. Hacking (1991) distinguishes three kinds of natural kinds. Mill-kinds are classes of things that share indefinitely many properties. Peirce-kinds are classes of things that are subject to the same set of laws and universalities. Leibniz-kinds are defined in terms of the deep structure of their instances. The basic idea behind all three conceptions is that the instances of a natural kind share a large number of properties for some specifiable reason (Platts 1983). For Mill-kinds, this reason may remain unknown. All we know is that for some reason or other, the instances of a natural kind do share an indefinite number of properties. However, since the number of these properties is indefinite, one cannot define a natural kind by listing all of them. There must be a finite subset of these properties by means of which instances may be identified as such, and ideally the properties in this subset will cause or explain many of the other properties (Mill, System of Logic IV,vii,2, Collected Works VII, p. 714). Peirce and Leibniz focus on specific ways of accounting for the features that instances of a kind share, either by reference to general laws or by reference to a shared microstructure. Russell adds to this that the properties that all instances of a natural kind share should not merely be conceptually related (1948, p. 335). It is not enough to say that the instances of a natural kind share an indefinite number of properties for some specifiable reason, because it is easy to define indefinitely many logically connected properties that all members of any arbitrary collection of items share. (The property of being self-identical while 1+1=2, the property of being self-identical while 1+2=3, etc.) For a natural kind, at least some of the many properties that its instances share must be conceptually independent from others, so that there is no merely conceptual explanation why all of them attach to the instances of the same kind. There must be a causal explanation of this fact, and like many other causal explanations, this causal explanation should support induction. One may therefore generally think of natural kinds as groupings of items that support induction (Goodman 1983, p. 122-3; cf. Griffiths 1996, p. S5).

In the present context, the notion of a natural kind is useful only if natural kinds

are not confused with necessary properties, as done by Cocchiarella (1976, p. 203). Also, the essence of a natural kind should not be identified with a set of necessary and sufficient properties that all instances must have, as Wilkerson has it (1993, p. 1). For, as I have argued, we should allow for types that cannot be defined in terms of simple intensions (i.e., non-relational features that all and only their instances share). Further, in the present context, there is no good reason to postulate that natural kinds are eternally unchanging (Ruse 1987, p. 229) or that they cannot possibly admit of borderline cases (as in Fales 1979, p. 351).[8]

Above all, there is no reason to suppose that all natural kinds must share a certain microstructure. Putnam is often taken to say as much (Ben-Yami 2001, p. 162; Rapp 1995b, p. 41). He writes that when we explain what water is by pointing at a sample of water, we define water as "whatever bears a certain equivalence relation ... to the piece of liquid referred to as 'this' in the actual world" (Putnam 1975 vol. 2, p. 231). The exact nature of this equivalence relation depends on the kind of kind in question. For water, it does seem to involve sameness of microstructure, so that something is water if and only if it has the same microstructure as "this." In other cases, however, the instances of a kind may be tied together by different relations of sameness in kind.[9] That sameness of kind does not always rest on sameness of microstructure is Ghiselin's valid point about biological species (1981, p. 271). What the instances of a biological species share might simply be that they are parts of one whole, and

---

[8] See Granger (1987b) for arguments against the assumption that Aristotelian kinds must be eternally unchanging, and Franklin (1986) and Granger (1985, 1987a) about the possibility of borderline cases.

[9] LaPorte (1996) shows that even for chemical types, microstructure is not essential. Johnston (1997) also argues against the idea that sameness of microstructure implies sameness in type.

obviously the parts of one whole need not all share the same microstructure.[10]
Millikan distinguishes between ahistorical kinds, which may often be defined in terms
of shared microstructural features of their instances, and historical kinds, which are
better defined in terms of some other story that connects their instances. In the most
obvious cases, this story is one that involves processes of copying or reproducing
instances of the kind. In still other cases, the instances of a kind may be alike in that
they originate in the same kind of circumstances or serve the same function (2000, p.
18-23).

In any event, the notion of a natural kind is of use to us as long as it is not
associated with the idea that their essences must be a fixed set of properties that all
instances share or a microstructure they all have. This, however, means that many
theories of natural kinds should not be trusted as to what essences are. We should
therefore directly address this question: What are essences if not properties?

5.   Identifying, Classifying, Describing

In order to attribute properties to something, one must identify a thing to attribute
them to. Essences are that by means of which we identify something as an instance of
a type. Once one has identified a thing one may attribute properties to it, and on the
basis of these properties one may classify it as an instance of a more specific (or more
general) type. It is thus not impossible to classify a thing on the basis of its properties.
However, it must be possible to identify a thing without already knowing its

---

[10] Ghiselin concludes, notoriously, that biological species are not kinds but individual
things, of which the instances of the species are parts, and this is where he clearly
stops making sense. Items may very well be of the same kind without sharing the
same microstructure. Therefore, even if the instances of a biological species should
not share the same microstructure, this would be no reason for denying that species
are kinds. On the other hand, there is no reason for denying that kinds are individuals
of a sort: They may be distinguished and counted. This does not mean that they are
not kinds.

properties. If essences are that in terms of which we identify a thing, there must be a distinction between essences and properties such that essences do not reduce to or depend on properties.

This seems to be open to the following objection. It appears that in order to identify a given thing A as an instance of a type T, one needs to know the following two things. First, one must know what properties a thing would have to have in order to qualify as a T. Second, one must know what properties A has in order to check whether these are the properties that a T would have to have. None of these two tasks seems to be possible without knowledge about properties, and hence, properties seem to be more fundamental than essences. This objection can be met by showing how a thing may be identified as an instance of a type without matching its properties against the defining properties of this type. In order to show this, I draw a distinction between two acts: the act of identifying a thing as an instance of a type and the act of attributing a property to a given instance of a type. This is in fact a version of the distinction between casting and stating that I discuss in Chapter 2. Cognates of this distinction are discussed in a number of places in the literature.

Strawson distinguishes between acts of referring to a subject and acts of predicating something of it. He argues that reference (to an empirical object) can be successful only if some particular empirical statement is true (1959, p. 183). In this sense, "a subject-expression ... presents a fact in its own right" (p. 187). This fact, however, is not stated when we refer to the item but only presupposed (p. 192). If it does not hold true, the reference is unsuccessful and therefore neither true nor false. For instance, if we refer to something as an ice cube and there is no ice cube, we have relied on the fact that there is an ice cube, but we have not said anything false about anything. We have failed to refer. Predication, in contrast, does not presuppose the truth of an empirical fact. It states a fact, and it is false if the fact does not hold. If we say of an ice cube that it is pink and it is not, we have referred to it and said something false. Strawson further points out that in order to say something substantive about a thing, the fact on the basis of which this thing is identified must

differ from the fact that is stated of it (p. 210). That is, in order to say something substantive and true about a thing, we need to know at least two things about it: that in terms of which we identify it (it is an ice cube) and that which we state of it (it is pink). Elsewhere, Strawson draws a distinction between a referential and predicative position in a categorical sentence and writes that "the singular term is used for the purpose of <u>identifying</u> the object, of bringing it about that the hearer (or, generally, the audience) knows <u>which</u> or <u>what</u> object is in question, while the general term is not" (1961, p. 398). He further argues that what is said of a thing is always of a higher logical type than this thing, so that properties are said of things, metaproperties are said of properties, and so on. Thus "pink" is of a higher logical type than "ice cube," and "bright" (as said of a color) of a higher logical type than "pink."

Donnellan distinguishes referential from attributive uses of a description. He writes that a speaker who uses a description referentially in an assertion uses it "to enable his audience to pick out whom or what he is talking about" (1966, p. 285). Martinich (1977) observes that when a description is used referentially its job may as well be done by any other description that is taken to be true of the referent. The person referred to as the first African American president of the United States may equally well be referred to as the first African American president of the Harvard Law Review. On the other hand, when a description is used attributively, its job may be done only by a (roughly) synonymous description. To describe Obama as the first African American president of the United States is not to describe him as the first African American president of the Harvard Law Review.[11]

––––––––––––––––––

[11] Martinich thinks that this defines the distinction that Donnellan draws. He argues that since a name may be used in a context where it may not be replaced by other references to the same item, there must be an attributive use of names, and he thinks that Donnellan cannot admit this. However, that referential terms may be replaced by arbitrary coreferential terms is not the feature that defines referential use. It is only a corollary of the definition. That a corollary of a definition is also true of something other than the defined thing does not affect the definition from which it follows.

Belnap draws a related distinction between conditional assertion (for all x: if x is T, then x is F) and restricted quantification (for all Ts x: x is F; 1970, p. 7).[12] A conditional assertion is about all items in the universe of discourse, and it attributes a predicate F to an item under the conditions that a predicate T applies to it. In contrast, a restrictedly quantified assertion is only about items that satisfy a certain description T. In a restrictedly quantified sentence, we do not state two things of a thing but only one; that it is a T is not stated but presupposed. The term T that restricts the quantification in "for Ts x: x is F" is not used in order to attribute a property to a given thing. It is used, as it were, to give a thing.

Strawson, Donnellan, and Belnap express the same idea. The distinctions they draw may all be mapped onto the distinction between the casting part and the stating part of a Sellarsian sentence that I draw in the second chapter (which is not to be confused with Austin's original distinction between casting and stating). When we cast an item as a T, we use the term T referentially in order to bring this item into view. We do not thereby state of this item that it is a T. The casting term is not said of the thing but only functions as a tool to get hold of it. In contrast, when we state of the item thus given that it is F, we use the term F attributively.

Millikan draws a related but different distinction (2000, ch. 3). She argues that an act of classifying a thing as a T involves the attribution of properties to it and that the act of identifying and keeping track of a thing must be prior to both classifying and describing it (p. 41). Now when we cast an item as a T, we seem to classify it as a T. If this is true, and if identifying is prior to classifying, identifying must also be prior to casting. It seems that there must be a way of identifying and tracking an item that is even more fundamental than an act of casting it as an instance of a type.

Millikan's distinction between identifying and classifying corresponds to Austin's original distinction between two directions of fit, as opposed to my distinction between the two parts of a Sellarsian sentence. In a sentence such as "#123 is a

---

[12] In a similar way, Gupta introduces common nouns as terms that may restrict quantification (1980, p. 7).

rhombus," an item is identified by using a label (#123), without yet being described or classified as an instance of a type, and then it is said of #123 that it is of type T (a rhombus). Referring to the item by using its label is an act of identifying it but not yet one of classifying it as an instance of a type. Saying of #123 that it is a rhombus is an act of classifying it, and this act presupposes the act of identifying it. However, we can only distinguish between identifying and classifying in this case because it is assumed that all relevant items have labels, and this assumption is not realistic. Not all items have labels, and the items that do have labels had to be labeled at some point. In an act of labeling them, they could not have been identified by using the label they were going to get, so it must have been possible to identify them without using the label. Therefore, although referring to an item by using a label is one way of identifying it, it cannot be the only and most fundamental way. Millikan has to explain how one can identify and track an unlabeled item without yet classifying it as an instance of a type.

There are arguments to the effect that one cannot identify an item without identifying it as an instance of a type.[13] They are cognates of the arguments discussed in Chapters 3 and 4: An isolated "this" does not refer. These arguments do not show that classification is prior to identification. They only show that to identify an item is always to identify it as an instance of a type. If Millikan is right and classifying an item presupposes identifying this item, to identify an item as an instance of a type must be different from classifying it. We must therefore distinguish between identifying something as an instance of a type and classifying it as an instance of a type. In the first case, the type term is used as a casting term; in the second, something that is already cast by other means is stated to instantiate a type. That is, we must distinguish the following three acts:

---

[13] For instance, Rapp argues that one can identify persisting, individual things only by applying sortal universals (1995b, e.g. p. 426-28) and that one can identify continuous matter (such as milk, gold, and water) only by applying concepts that are in important respects analogous to sortal universals (p. 283).

<u>Identifying</u> an item as a T	=	casting an item as a T.
<u>Classifying</u> an item as a T	=	stating of a given item that it is a T.
<u>Describing</u> an item as F	=	stating of a given item that it is F.

Let me return to the objection raised above. When we cast an item as a T, there must be a reason for casting it as an instance of this type rather than another. There must be specific criteria that a T would have to satisfy, and this seems to mean that in order to cast something as a T we must first check whether it satisfies these criteria or not. In order to do this, however, we need to know the properties of the thing to be cast, and therefore properties must be more fundamental than casting terms (or essences).

Millikan offers the following reply. She points out that in human development the capacity to identify and track objects emerges earlier than the capacity to classify them as instances of types. Further, the latter capacity (to classify objects) involves and presupposes the capacity to identify and keep track of these objects. Therefore, one must be able to identify a thing before being able to describe or classify it. Also, in order to identify and track an item in terms of its features, one would have to identify and track the features of an object, and tracking features of objects does not seem to be easier than tracking objects (2000, p. 66).[14] Moreover, acts of identifying and classifying serve different functions. We classify things into types in order to conveniently bundle, store, and retrieve general information about them. We identify things in order to apply this general knowledge in a particular case or to gather new knowledge about this case (p. 34-9).

This reply only shows that there must be a way of identifying a thing without

---

[14] This also affects Timaeus' recommendation that in a world of constant change, one should refer to features instead of their subjects (cf. Chapter 3, section 5). It is not clear how this recommendation would make things easier, since the features should be as difficult to locate as the things that have them.

describing it, and that in fact there is one. It does not show how it is possible. I submit that this can be shown by employing the distinction between casting and stating, in the sense in which I distinguish between a casting and a stating part of a Sellarsian sentence.

Casting an item is like calling a name.[15] We call a name and someone may or may not show up. When someone responds we have good prima facie reasons for assuming that the respondent actually has the name we called. But this need not be so. The person may respond for unrelated reasons—she might have misheard the name, or she might respond because she knows that we got her name wrong. (I usually respond to the name "Henning.") Likewise, if an item responds to our attempt at casting one, we have good prima facie reasons for assuming that the item responds for the right reasons, so that it may also be described or classified in the terms that we used for casting it. However, it may as well respond for unrelated reasons. In particular, we may successfully cast an item of the right kind even though it lacks features that we would expect of such instances or we are mistaken about the typical features they have. For instance, we may cast a beaver as a quadruped although it has only three legs, and we may successfully cast whales as fish although they are mammals. This is possible because the fact that we may successfully cast an item by using a description does not imply that this description is true of it.

As Donnellan points out, one may successfully refer to a person as the man drinking a martini even though this person is only drinking water (1966, p 287). Note that this does not in turn imply that one can refer to a thing under an arbitrary description (Donnellan 1968, p. 210). I can only intend to cast something as a T if I expect that there is something that can be cast as a T. But whether a reference to an item succeeds does not actually depend on what properties this item actually has. Rather, "the intention to refer to something in using a definite description is a complex intention involving expectations regarding one's audience" (p. 214).

---

[15] Cf. Wittgenstein, Philosophical Investigations §27, tr. Anscombe: "Think in this connexion how singular is the use of a person's name to call him!"

Therefore, even if the reason for identifying a thing as an instance of T may later be described as the fact that the item has a certain property, this does not mean that we have identified the item <u>as</u> having this property. That a description has been successfully used for casting an item does not imply that it is true of the item cast. This shows that there is a difference between casting something as a T and stating of it that it has certain properties. When we identify something in terms of a property we do not identify it <u>as</u> whatever has this property. For instance, that an infant recognizes its mother by her smell does not imply that the infant identifies the mother "as whatever bears that smell" (Millikan 2000, p. 81).

However, there seems to be a difference between descriptions and type terms. When we cast an item according to a description, we do not cast it as whatever satisfies this description. But when we cast an item by using a type term, we do cast it <u>as</u> an instance of this type. In this sense, casting involves classifying. The important thing to note here is that the fact that something instantiates a type need not imply that it actually has any interesting property, other than the vacuous property of instantiating this type, the relational property of instantiating the same type as several other items, and the property, if this is a property, of being subject to certain standards of typicality. When we cast an item as an instance of a type, we cast it in terms of an essence, and the essence of a thing is not a property that it actually has. Even more, not even the essential properties of a thing are properties that it actually has. They are properties only of the essence of this thing. When we cast an item as a quadruped, we do not imply or presuppose that it has four legs. All we presuppose is that if it has only three legs one must be missing.

This means, again, that essences figure in a counterfactual context. The essential properties of a thing need not be its actual properties. Rather, they are properties that one would expect it to have given that it is of a certain type. When we cast an item as a T, we do not state of this item that it satisfies a certain description. Instead, we make it available as an instance of a certain type, and this justifies certain expectations as to how it will behave. The item may fail to meet these expectations in two different

ways. First, it may be a T but lack certain features that we associate with being a T. By casting it as a T, we have made it likely that it would have certain properties; it does not have them, but it still is a T. This may happen if the thing is an atypical T or if our expectations as to how Ts should behave are mistaken. Second, the item that we have successfully cast as a T may turn out to be of a type other than T. In this case, we have raised the wrong expectations.

I have distinguished between these two kinds of expectation (or thresholds) in Chapter 5. I have argued that the criteria for being subject to a standard are not the criteria for satisfying it. When we cast an item as an instance of a type, we do not suppose that it satisfies the standards that apply to it and thus not that it actually has certain features. We only suppose that it is subject to certain standards, so that it qualifies as typical to the extent to which it has certain features. This supposition—that it is subject to certain standards—must be justified in some way, but the facts on the basis of which it is justified do not define the type to which the standards belong.

This is another reason why types are best defined in terms of standards of typicality and not in terms of actual properties. The attributive use of descriptions concerns facts. The referential use of criteria concerns standards and expectations rather than actual facts. Essences and essential properties are that in terms of which we refer to a thing in order to then be able to say something about it. Therefore, essences and essential properties function as standards that apply to what is referred to, and not as actual properties that inhere in it.

The answer to the objection raised above is thus that we do not need to know a thing's actual properties when we cast it as an instance of a type. To cast a thing as an instance of a type is not to state that it has certain properties but to raise certain expectations as to what properties it will probably have and to set up a standard according to which whatever properties it has qualify as typical or not. If the thing fails to meet our expectations or is highly atypical, we might consider casting it as an instance of a different type instead. But then it might also be of a different type even though it satisfies our expectations and standards. It may do so because it is a very

atypical instance of a different type. In any case, there are ways to dispute and defend acts of casting but they do not simply reduce to the question of what properties a thing actually has. There may be good reasons why an expectation may be unjustified or a standard inappropriate, but usually the reason is not simply that the expectation or standard fails to be met.

6.   Another Take on Metaphysics Z 13

Although the distinction between referential and attributive uses of a descriptive term is not Aristotle's, it may be used in order to clarify a controversial issue concerning Aristotle's paradigmatic forms. On the one hand, Aristotle writes that paradigmatic forms are οὐσίαι and that no universal can be an οὐσία. On the other hand, the paradigmatic form of a thing is what its definition signifies, and no particular individual can be defined (as a particular individual). Further, Aristotle writes that the instances of a species have one paradigmatic form in common and that what is common to many is universal. He thus seems to give two conflicting answers to the question of whether paradigmatic forms are universal (cf. Lesher 1971, Rapp 1996).

Modrak (1979) draws a distinction between two kinds of universal—substance types and properties—and suggests that when Aristotle speaks of definitions of substances he refers to substance types, whereas when he denies that substances are universal he means that no substance can be a property. Similarly, Woods argues that there is a difference between being a universal and being predicated universally and that Aristotle only denies that οὐσίαι are predicated universally, not that they are universals (1967, p. 229). These suggestions take us some way in the right direction. They still fail, however, because they rely on a distinction between two kinds of universal, or being a universal and being predicated universally, and no such distinction can be found in Aristotle. Also, they still assume that paradigmatic forms are universals of a sort, and this is what Aristotle seems to deny in Z 13.

Modrak distinguishes between kinds of universals, and Woods distinguishes

between being a universal and being used in a universal way. I suggest that we should instead distinguish between two different ways of using paradigmatic forms. When we use a paradigmatic form in order to classify a given item, it functions as a universal in that we use it in order to say something general. When we use a paradigmatic form in order to identify a thing, we refer to this thing as an instance of a general type but we do not thereby say something general of it. In themselves, paradigmatic forms are neither universal nor particular, but they may be used in both ways.[16]

This idea, that paradigmatic forms might in themselves be neither universal nor particular, is not entirely new. Owens writes that Aristotle's forms make possible the cognition of both the individual and the universal and that they must therefore be prior to both (1963, p. 392). Frede and Patzig also note that the type term "man" may either be understood in a way so that it says something general about a particular item or so that it refers to an individual item as an instance of something general (1988, Kommentar p. 14). Further, Rapp writes that when we use a sortal predicate in order to identify an item, we do not emphasize what is common to many but single out a particular item as an instance of a kind (1995a, p. 85; cf. 1995b, p. 451). Elsewhere, he applies this to Metaphysics Z 13: When an essence is said of a thing in order to identify it as an instance of a type, it is not said of this thing as something else, nor is it attributed to more than one subject (1996, p. 177-82; 2007, p. 169).

That the distinction between universal and particular is not one between kinds of term but one between different ways of using a term may perhaps be further supported by a reflection on the "literal" meaning of the Greek expressions for "universal" and

_____

[16] Cf. Kant, Logik Jäsche §1, note 2 (Akademie Ausgabe IX, p. 91): there is no distinction between universal and singular concepts, only a distinction between a universal and singular use of concepts.

"particular" (καθόλου and καθ᾽ἕκαστον).[17] The word καθόλου is composed of κατά with a genitive, and in general κατὰ + genitive indicates a downward or inward movement, starting from something higher or from an encompassing whole. In καθ᾽ἕκαστον, κατά is used with an accusative, and κατά + accusative generally indicates a movement toward some particular thing or through a number of such things, such that saying something καθ᾽ἕκαστον is to address a single item, or one of several items at a time, by going through them. Both καθόλου and καθ᾽ἕκαστον may thus, tentatively, be taken to refer to the same movement, starting from an encompassing whole, toward its particular parts. They differ by emphasizing different aspects of this movement. When we say something καθόλου, the emphasis is on the whole from or within which we start. When we say something καθ᾽ἕκαστον, the emphasis is on the particular items toward which we move.

The vocabulary that Austin introduces in How to Talk (1979) may be used in order to systematize this distinction. Austin distinguishes four ways of using an utterance of the form "Item A is of type T." In Chapter 2, section 2, I have discussed two of them:

(1a) <u>Stating</u>: Given an item A, to answer the question about this item as to what its type is.

(2a) <u>Casting</u>: Given a type T, to answer the question about any of the available items as to whether they fall under it.

These acts differ in their direction of fit. In an act of stating an item is given, whereas in an act of casting a type term is given. In both cases it is said of the item

---

[17] This is highly speculative. What I call the "literal" meaning of καθόλου and καθ᾽ἕκαστον is probably not the meaning in which Aristotle or any other Greek has ever consciously used these phrases. The above is only a piece of speculative etymology, and as such it cannot be used as a starting point for an argument or demonstration (only as a starting point for further speculation). Nothing else in this book depends on it.

that it instantiates a type. Since the distinction between καθόλου and καθ᾿ἕκαστον does not concern the direction of the movement from something general (the type) toward something particular (the item), but rather concerns the emphasis on its parts, it may be drawn in terms of what Austin calls the "onus of match." We should therefore consider the two acts with the opposite onus of match:

(1b) <u>Placing</u>: Given an item A, to answer the question about any of the available types as to which type it instantiates.

(2b) <u>Instancing</u>: Given a type T, to answer the question about this type as to whether any of the available items instantiates it.

The distinction between καθ᾿ἕκαστον and καθόλου corresponds most clearly to the distinction between casting and instancing. Casting (2a) is about something particular. When we cast an item as an instance of a type, we use a general term in order to refer to something particular. We use a given general type term in order to get hold of something particular. The result of casting is a reference to a particular thing. In this sense, casting proceeds toward something particular (καθ᾿ἕκαστον). When we instance a type (2b), on the other hand, we say of this type that an item instantiates it. We do not primarily pick an item by using the type term, but we say of the type term that an item instantiates it, however we get hold of this item. Here the emphasis is on the type and we proceed from something general (καθόλου).

Stating and placing are somewhat difficult to describe in terms of the distinction between καθόλου and καθ᾿ἕκαστον. When we state something of an item (1a), we say of a given item that it instantiates a certain type. Since the item is already given, we do not proceed from something general toward something particular. Rather, we start with something particular and end up with something general. When we place an item (1b), we say of a type that a given item instantiates this type. Again, we start with something particular and end up with something general, but this time the emphasis is on the general.

	(a) emphasis on particular	(b) emphasis on general
(1) from particular to general:	<u>stating</u> from the particular	<u>placing</u> toward the general
(2) from general to particular:	<u>casting</u> toward the particular (καθ᾽ἕκαστον)	<u>instancing</u> from the general (καθόλου)

This is a complicated set of distinctions. All we need here is the distinction between casting (2a), on the one hand, which proceeds καθ᾽ἕκαστον and is the clearest case of an employment of a general term toward something particular, and all other possible uses of sentences of the form "Item A is of type T" on the other hand, which may all be said to be about or toward something general. We need this distinction in order to understand what Aristotle says in Z 13: that paradigmatic forms cannot be universal, even though the object of a definition can only be universal. Paradigmatic forms refer to particular substances when they are used in order to cast them. If I am right, Aristotle means even though paradigmatic forms may also be used καθόλου in other contexts, they cannot be used in καθόλου in this context. To use a term καθόλου is to proceed from something particular toward something general. We do this when we describe or classify substances, but not when we identify them as instances of general types. To single out a substance is to speak καθ᾽ἕκαστον. In this sense, primary substances are not καθόλου: A reference to them is not the result of predicating a paradigmatic form καθόλου.

In this chapter, I have argued that essences are not properties. Essences provide standards of typicality that apply to instances of a certain type, and to describe these standards is to describe what satisfies them. Essential properties are properties of what satisfies the standards in question. Therefore, the essential properties of a thing are not its actual properties but the properties that it would have if it would fully satisfy the standards of typicality associated with its type. When we identify an item as an

instance of a type, we do not thereby attribute a property to this thing. All we do is to imply that the thing in question is subject to certain standards of typicality, so that it qualifies as typical to the extent to which it has the relevant essential properties. It may lack some of these essential properties, in which case it will be atypical. To cast an item as an instance of a type is thus to raise certain expectations as to how the thing that is cast will behave.

Is Causation a Relation?

The efficient cause is unlike the other three Aristotelian causes in that there is such a thing as efficient causation. Pace Charles (1991), there is no such thing as "teleological causation," nor is there material or formal causation. Material, formal, and final causes do not give rise to effects, except accidentally, by also being efficient causes.[1] The efficient cause is called efficient because it is the only cause that has effects.[2] It would therefore be wrong to associate material, formal, and final causes with relations between causes and effects. It is not obviously wrong to think of efficient causation as a relation. Aristotle also refers to the efficient cause as that which produces what is being produced (τὸ ποιοῦν τοῦ ποιουμένου, Physics II 3, 194b31), and as he emphasizes in Nicomachean Ethics VI, production (ποίησις) differs from action (πρᾶξις) in that it is directed at an end that differs from the act of achieving it (1140b6). If the efficient cause is a producing cause, it therefore aims at something that differs from the act of aiming at it. This means that it must be possible to divide all causal processes into three parts: an efficient cause, something that this cause does so that an effect results, and an effect that differs from what the efficient cause does.

In this chapter, I ask how distinct these three parts actually are. It is often assumed that cause and effect are two independent and separable entities and that causation is a relation between them. I argue that to the extent to which cause and effect are

---

[1] Bronze does not produce a statue (Metaphysics A 3, 984a24-25), and the final cause does not produce anything (De Generatione et Corruptione I 7, 324b15).

[2] As far as I can see, Aristotle uses αἰτίατον in four places (Posterior Analytics I 9, 76a20; II 17, 98a36 and 98b3; Metaphysics K 8, 1065a11). In all of them, he discusses causes of all kinds, which means that αἰτίατον should not be translated as "effect." The αἰτίατον of a cause is whatever it is a cause of. The efficient cause is the cause whose αἰτίατον is an effect.

conceptually separated they cease to be, or to be seen as, cause and effect. As cause and effect, they are in fact two aspects of one and the same causal process (just as the matter and the form of a thing are two aspects of this thing). The examples that Aristotle gives of efficient causes are cases where one thing does something to another thing, so that the causal process of which cause and effect is a part involves two (numerically) different things. In this chapter I mainly consider simpler cases, where a process undergone by a thing causes a process undergone by the same thing. Cases where one thing does something to another thing may be described as causal processes undergone by pairs of things.

That we think of causation as a relation between distinct items seems to be mostly due to Hume; it is not part of the Aristotelian tradition. It is revealing that Sosa and Tooley give the following reasons for neglecting Aristotelian accounts of causation: "Aristotle did not think of causation, as present-day philosophers do, as a relation between events or states of affairs"; and "Aristotle was apparently unaware that there are very serious difficulties concerning the concept of causation" (1993, p. 31). These are actually good reasons in favor of reading Aristotle. The reason why he is unaware of difficulties may well be that he does not have them, and he might not have them because he does not think of causation as a relation between events or states of affairs. I therefore try to think of causation as something other than such a relation. Again, I will not extract my arguments from Aristotle but briefly indicate in the end how they match the results of Chapter 1.

1. Causation as a Relation

Hume thinks that if causation is anything at all it must be a relation between two different items: a cause and an effect. So does almost everyone after Hume. It is considered a matter of debate whether causation is a relation between distinct events (Davidson 1980, ch. 7), facts (Mellor 1995, ch. 13), or properties (Fales 1990, p. 52). Further, it is debated what this relation exactly is. It has been suggested that two items stand in a causal relation if there is a law of nature that somehow ties them together

(Davidson); if one of them is counterfactually dependent on the other (Lewis 1986); if one can manipulate one of them by manipulating the other (Gasking 1996, ch. 7); if some underlying mechanism can be found that connects them (Glennan 1996); or if there is some transfer of conserved quantities among them (Dowe 1995).

The assumption on which these debates are based—that causation is indeed a relation between a cause and an effect—has only rarely been seriously questioned. One good reason for this is that in all causal processes one may distinguish between a cause and an effect, which must stand in a certain relation to each other. Given a suitable interpretation of R, a, and b, it may therefore seem as though one can refer to any causal process by using the formula R(a,b).[3] It seems impossible to find instances of causation that are not relations in this sense. This is why Mellor, for instance, prepares his argument against the assumption that causation is a relation between events by postulating that for R(a,b) to be a relation, both a and b must exist (1995, p. 156). He concedes that causal processes can always be represented as some R(a,b); his only problem with this is that not everything that may be represented in this way is a relation, because a or b may represent the absence of something. David Lewis does the same when he argues that causation is not a relation. He points out that the void, for instance, has causal properties and that relations require relata (2004, p. 281). Since the void is nothing, and nothing cannot be a relatum, he concludes that causation is not a relation. I do not find this argument very compelling. It does not seem to be the case that generally relations cannot involve absences. Consider the relation "better than." Is it not true that in certain circumstances, five dollars are better than nothing? Or that lack of poison is usually better than lack of water? It does not follow from any of this that "better than" is not a relation.

One thing that motivates the assumption that causation, if it is a relation at all, must be one among existing things might be that causal relations would have to be backed up by some kind of real connection in the world. Real connections, it is said,

_____

[3] This only seems to be so. That a process has parts that are related to each other does not imply that this process is a relation among its parts.

cannot hold among absences. Causation may involve absences, it is then argued, and therefore it cannot be a relation. I want to rest my argument on the opposite case. It is precisely <u>because</u> instances of causation are real goings on in the world that they are not relations.

Relations are the results of relating items; in order to get a relation between two items, one need not do more than put these items together. This is why relations can be identified with sets of ordered n-tupels. If causation were nothing but some kind of relation between distinct items, one should be able to independently identify causes and effects, put them together in the right way, and thus get an instance of causation. I argue that this is not possible. I am not objecting to the idea that one might consider the causes and effects that are involved in a given causal process in isolation. As far as I am concerned, one may split up causal processes into as many separate parts, respects, and moments as one pleases. One may also put them together again, and what one gets by doing so is something that I am willing to call a causal relation. However, it is one thing to be able to split something up into parts, another to be able to construct it out of independently given parts. I believe that although one can split up causal processes into separate and independent parts, one cannot get them back as causal processes by merely relating these parts to each other. This is what it means to deny that causation is a relation.

There are some authors who also seem to deny that causation is a relation. Hanson, for instance, argues that the causal chain analogy is misleading (1958, ch. 3). However, not every causal relation is a chain, and therefore he does not in general deny that causation is a relation. Also, Taylor does not make a general case against the claim that causation is a relation; he only denies that it should in any case be construed as a relation between two events (1966, p. 16). The same is true for Mellor. He seems to deny that causation is a relation. However, all he wants to say is that causation is not a real relation between particular objects or events (1995, p. 168). These authors do not deny that causation is a relation. They argue that it is not a relation of a certain kind (such as a chain), or not one between entities of a certain

kind (such as particular objects or events).

Fair, Salmon and Dowe might qualify as allies in my case, since they argue, modulo some variations, that causation is a process by which some persistent quantities are transmitted from one item to another (Fair 1979; Salmon 1998; Dowe 1995). If processes are not relations, causation is not a relation. However, these authors do not actually deny that causation is a relation. Fair, for example, has no difficulties calling causation "a physically specifiable relation of energy-momentum flow" (1979, p. 220). If energy flow were a relation, causation would also be one.

Chakravartty also claims that one should conceive of causation "not as a relation between discrete entities" (2007, p. 104). However, he ultimately characterizes causal processes as "systems of continuously manifesting relations between objects with causal properties and concomitant dispositions" (p. 114). I am not sure what to make of the idea of a continuously manifesting relation. What Chakravartty seems to have in mind is something that I would not call a relation but a continuous process. In any case, his account appears to involve a hidden circle. In order to understand what causation is according to Chakravartty, one needs to understand what causal properties are, and causal properties are properties that confer dispositions for behavior on objects (p. 108). Actualizations of dispositions, however, are conceptually prior to dispositions, and therefore, in order to explain what causal properties are, one needs to say what kind of behavior the objects in question are disposed to engage in. The answer is, presumably: some kind of causal process. Causal properties are thus precisely those properties that dispose things to take part in causal processes. Therefore, what Chakravartty effectively says is this: Causal processes are systems of continuous processes ("continuously manifesting relations") among objects with properties that dispose them to take part in causal processes. This is, of course, circular.

2. Hume's Argument

Hume is quite liberal about what he calls philosophical relations: They can be any

"circumstance, in which, even upon the arbitrary union of two ideas in the fancy, we may think proper to compare them" (Treatise I,i,5, Selby-Bigge p. 13). There need be no real connection for two ideas to be related, and no positive existences need correspond to these ideas. There can be philosophical relations among absences. But there is one thing that Hume will not allow: There are no relations that hold between one thing and this very same thing (in the same respect). Philosophical relations can only hold between distinct items (cf. Beebee 2006, p. 31). Even the relation that Hume refers to as identity is not the relation of a thing to itself but a relation between different occurrences of the same, at different places and times (Treatise I,iii,2, p. 74). This is why the relation of identity is common only to all beings "whose existence has any duration" (I,i,5, p. 14). A thing can be identical to itself only in the sense that its distinct stages are stages of the same thing. Hume thus denies that there are any reflexive philosophical relations, which might not be a good idea. All transitive and symmetric relations must also be reflexive, and therefore Hume cannot allow for transitive and symmetric philosophical relations. At any rate, if causation is a philosophical relation in Hume's sense, it must be irreflexive and asymmetric. It can only hold between distinct items.

The claim that causation is, if at all, a relation between distinct items has important consequences. For according to Hume, whenever two items are distinct they are also separable. Hume writes that "there are not any two impressions which are perfectly inseparable" (I,i,3, p. 10). This is sometimes called the Separability Principle (Garrett 1997, ch. 3). It seems to amount to the claim that if two items are in any way different from one another they are also separable. Now if cause and effect must be two distinct items, or two distinct stages of one item, and if any two different entities are perfectly separable, then it follows that cause and effect must be perfectly separable from each other. This leads straightforwardly to Hume's claim that "the effect is totally different from the cause, and consequently can never be discovered in it" (Enquiry IV,i, §25, Selby-Bigge p. 29).

As far as I can see, there is only one passage where Hume comes close to

defending the assumption that causation is a relation. He writes:

> Let us therefore cast our eye on any two objects, which we call cause and effect,
> and turn them on all sides, in order to find that impression, which produces an
> idea of such prodigious consequence [as the idea of causation]. At first sight I
> perceive, that I must not search for it in any of the particular <u>qualities</u> of the
> objects; since, which-ever of these qualities I pitch on, I find some object, that is
> not possest of it, and yet falls under the denomination of cause or effect. And
> indeed there is nothing existent, either externally or internally, which is not to be
> consider'd either as a cause or an effect; tho' 'tis plain there is no one quality,
> which universally belongs to all beings, and gives them a title to that
> denomination.
> The idea, then, of causation must be deriv'd from some <u>relation</u> among objects;
> and that relation we must now endeavour to discover. (Treatise I,iii,2, p. 75)

If this is an argument to the conclusion that causation must be a relation, then it
proceeds as follows. Hume assumes that our idea of causality must rest either on an
impression of a quality that all causes and effects share, or on an impression of a
relation. He then argues that there is no quality that all causes and effects share and
that therefore our idea of causation must rest on the impression of a relation.

The premises of this argument are questionable. There are in fact several attributes
that all entities share (some of them are known as transcendental attributes); and it is
not obviously wrong to think that these attributes might be the very attributes that
make everything that exists a cause or an effect of something else.[4] There might be
ways of construing what Kant says in some such way. I need not elaborate this point
here; I am merely casting doubt on the premise that there are no qualities that all

---

[4] Hume asks for observable qualities, and arguably, transcendental attributes are not
observable qualities of individual things. But then, according to Hume, causal
relations are also not observable in particular instances.

causes and effects share. There may well be such qualities.

It is also not at all clear why one should only consider qualities and relations as possible candidates. If something other than a quality or a relation could give rise to the idea of causation, one could not infer that it must be a relation from an argument that shows that it is not a quality.

Most importantly, Hume argues here that causation is not a quality because we cannot discover any such quality by inspecting particular causes and effects. Later on, he will claim that we also cannot discover any causal relations by inspecting particular causes and effects. If his argument against causal qualities is valid then it should also count against causal relations. It should follow that after all, causation is not a relation any more than it is a quality. On the other hand, if we can discover causal relations by considering entire sets of causal sequences instead of inspecting particular causes and effects—which is what Hume suggests regarding causal relations—then we should also be able to discover causal qualities by the same means. In the end, Hume gives us no good reason for thinking that causation is a relation rather than something else.

3. Drowning

So far I have only rejected Hume's argument for the assumption that causation is a relation. That he does not provide good reasons for thinking that causation is a relation does not mean that there are no such reasons. I rest my positive case against the assumption that causation is a relation on a consideration of the following example of a causal process given by Hume:

> The idea of sinking is so closely connected with that of water, and the idea of suffocating with that of sinking, that the mind makes the transition without the assistance of the memory. (Treatise I,iii,8, p. 104)

This is supposed to be a case in which one is justified to expect a certain effect,

given its cause, on the basis of past experience. We are accustomed to expect sinking when we see water, and our idea of causation rests on this habit. As Hume tells us, the causal relations between water, sinking, and suffocating are not necessary, since nothing about water logically implies sinking and nothing about sinking logically implies suffocation. The first man on earth "could not have inferred from the fluidity and transparency of water that it would suffocate him" (Enquiry IV,i, §23, p. 27). However, Hume says, there is still a non-accidental connection between water and suffocation, about which we may learn from experience. Once we have learned our lesson, we come to expect sinking when we see water and suffocation when we begin to sink.

The problem is that this cannot possibly be what we learn about the causal properties of water. It would simply be wrong to assume a straightforward causal connection between water and sinking, since we deal with water all the time without any risk or even possibility of sinking into it. We take a walk when it rains, water the plants, brush our teeth, have a drink, take a shower, and so on; no sinking is to be expected in any of these cases. Further, it would not be much more reasonable to expect suffocation on any encounter with water, even if this encounter is a case of sinking. Someone who expects suffocation whenever she sinks into water does not have a reasonable causal belief but a serious problem. That is, if there is any habit that a reasonable person should develop concerning the causal properties of water, it cannot be the habit of expecting sinking and suffocation when encountering water. When Hume writes that water has "suffocated every human creature" (Enquiry VI, §47, p. 57), he must be imprecise or exaggerating.

However, when trying to spell out what Hume should have said so that one may safely take him literally, it turns out that cause and effect are not as separable as he thinks. For suppose I step into a puddle. Should I expect to sink? No. Suppose I am lying in a bathtub, with my entire body under water. Should I now expect to sink? No. Suppose that I am swimming in a deep lake; I am a good swimmer, and not tired. Again, there is no good reason to expect that I will sink. Suppose, finally, that I find

myself in a deep lake, unsupported, and unable to swim. Now I should expect to sink. But this situation, where I find myself in this lake, unsupported, and unable to swim, is a situation where I already _am_ sinking.[5] I need not any longer expect it; it is already happening. The same is true for suffocation. There is no reason to expect suffocation whenever my head is under water; I need only expect suffocation when this impairs my respiration for a long enough time. And when it is doing so, I am already suffocating.

This shows that there are good reasons to expect an instance of sinking only when at least part of this sinking is already taking place, or when some other process is going on of which sinking is a later stage. Likewise, there are good reasons to expect suffocation only when some process is already going on of which suffocation is a later stage. What I may expect in both cases is simply that the causal process that is currently going on will continue to go on.

The suffocation that is caused by sinking in water, the kind of sinking that lets us expect suffocation, and the encounter with water that lets us expect both, are in fact all the same instance of drowning. I may reasonably expect suffocation upon an encounter with water to the extent to which I think of this encounter with water as the beginning of a drowning or a process that involves drowning, of which suffocation would typically be a later stage. If there is any relation between cause and effect here, it is very much like Hume's relation of identity: It holds between different stages or occurrences of the same occurrence. Here, it holds between different stages of one instance of drowning.

Note that the problem with Hume's formulation is not that water does not _necessarily_ lead to suffocation. It is not that the relation between water and suffocation is contingent. I do not suppose that something is only a cause if nothing can stop it from producing its effect. Even if one sinks into water deep and long enough to make it reasonable to expect suffocation, something might intervene. As long as suffocation has not come to an end it can be interrupted. The part of

---

[5] Cf. the similar point about teleological processes in Rödl 2003, p. 107.

suffocating that has not happened need not necessarily happen. We all hope this. But still, as long as an instance of drowning is going on, one is entitled to expect that it will continue to go on. If one were only entitled to expect suffocation when it cannot any longer be prevented, there would be no use for lifeguards. So the problem is not that water does not necessitate sinking and suffocation. It is that most encounters with water do not at all license the expectation that one will sink or drown.

The argument I have given above might appear to resemble one famously put forward by Russell (1912, p. 5-8). Russell's argument, however, depends on the assumption that all causes must necessitate their effects, which I have just rejected. He argues, roughly, as follows. A cause and its effect are either (a) temporally contiguous or (b) not. If (a) cause and effect are contiguous, then at least one of them must be temporally extended. If the cause is extended, only its very last stage can actually be causally efficacious. If the effect is extended, only the first bit of it is actually the effect. On the other hand, if (b) cause and effect are not contiguous, then there is a temporal gap between cause and effect, which means that something might intervene. If something might intervene, Russell argues, the connection is not necessary and thus not a causal one. The argument concerning the first horn of this dilemma (a) is not convincing, but we need not fix this here. I am interested in the second horn (b): the argument that cause and effect cannot be separated because then something might intervene. Russell assumes that in order to be a cause of suffocation, water would have to necessitate suffocation. My argument does not depend on this assumption. Causes need not necessitate their effects; they need only make it reasonable to expect them. A mere encounter with water does not even make it reasonable to expect suffocation.

Collingwood gives a version of Russell's argument that is a bit closer to mine. He writes that if an event A is to cause another event B in the relevant sense of "cause," then A cannot stop anywhere short of B (1940, p. 314-5). Thus, if A causes B at all, B

must be happening as soon as A is happening.[6] Collingwood, however, assumes that a cause must include strictly everything that leads to its effect, so that the effect must inevitably follow if the cause is present. Again, I do not and need not share this assumption.

I also do not think that whenever one may reasonably expect drowning, drowning must already be going on. What is going on might be any kind of process of which drowning is a later stage. Suppose, for instance, that I see a stroller with an infant in it rolling toward a deep lake, and assume that there is no fence or lifeguard or anything else that would seem to prevent the stroller from falling into the water. In such circumstances, I am entitled to expect that a human being will drown, although what is taking place is not yet the beginning of an instance of drowning. It is the beginning of a larger process of which drowning is a later stage, just as sinking in water may be the beginning of a process of which suffocation is a later stage. So I do not claim that whenever one is entitled to expect something it is already happening. I claim that whenever one is entitled to expect the occurrence of a particular (natural) process, it must be a part of a causal process that is already going on. This is only true for particular occurrences of natural processes. One may expect, in general, that stones will sink when thrown into water, but this is not to expect a particular occurrence. Further, one may expect an action to be performed on the grounds that it has been promised, but actions are not natural processes.[7]

4. Three Objections and Replies

I maintain that a cause and its effect are two parts of one and the same causal process.[8] I have argued for this claim by pointing out that it is not reasonable to

––––––––––––––––––

[6] Cf. Physics II 3, 195b16-18, where Aristotle says that an actual cause and its effect must be simultaneous. This is so because the effect is the actualization of the cause.

[7] Cf. Chapter 8, end of section 5, concerning actions.

[8] It remains to be said what causal processes are. I do this in the following chapter.

expect sinking and suffocation upon every kind of confrontation with water and that a description of an encounter with water that justifies this expectation is in fact a description of the beginning of a drowning. This way of putting my argument gives rise to the following three objections.

First, many authors believe that causal connections must be contingent. However, if the kind of encounter with water that entitles us to expect suffocation is in fact the beginning of an instance of drowning, then cause and effect seem to be conceptually related, and not only as a matter of contingent fact. That drowning causes suffocation sounds a lot like an analytical truth. If causal connections must be contingent, this means that there can be no causal connection between drowning and suffocating.

Second, it may seem as though the degree to which one is entitled to expect a certain effect depends on the way in which the cause is described. But whether one is justified to expect suffocation should not be a matter of how one describes things. It should be a matter of how things are.

Third, what is present cannot be a part or an aspect of what is not present. However, if the cause of suffocation is the beginning of an instance of drowning, it is present before the end of drowning (suffocation) has occurred. Now if the cause of suffocation is the beginning of an instance of drowning, and suffocation may still be prevented at the time when it occurs, it might turn out that what seemed to be the beginning of a drowning was actually not the beginning of a drowning. There might be no drowning for it to be the beginning of. Then, however, Russell and Collingwood seem to be right: If and as long as suffocation may still be prevented, sinking cannot be taken to be its cause.

As for the first objection, it is often taken for granted that causes cannot logically entail their effects in the same respect in which they cause them. For instance, Schlick argues that the amputation of a limb should not be taken to cause the loss of a limb because, as a matter of definition, it is the loss of a limb (1979, p. 244). In a similar vein, Mackie claims that the movement of a cricketer's bat should not be said to cause a drive, because again, as a matter of definition, it is a drive (1974, p. 288). The idea

that causes and effects cannot be logically or conceptually related is often attributed to Hume. Dardis, for one, believes that "Hume taught us that causal connections are not necessary connections" (1993, p. 577), and that therefore nothing that logically entails X can be causally relevant for X. I think this is an overstatement. Hume taught us that causal connections, if there are any, need not be necessary connections, not that if there were necessary connections they could not possibly be causal. It should be clear that when we (successfully) refer to an event as the cause of X, this logically entails that it is the cause of X and there can be no principle demanding that just because something is referred to as the cause of X, it cannot be causally relevant for X. Rather, if anything at all is causally relevant for X, the cause of X should be so. As Davidson says, "there is something very odd in the idea that causal relations are empirical rather than logical" (1980, p. 14). Of course, Dardis does not want to say that when we call the cause of X "the cause of X," it must therefore cease to be the cause of X. What he actually wants to say is something more reasonable. His point is that under a description that does not imply anything other than that the effect will occur, a cause is not <u>explanatorily</u> relevant. Simply to refer to "the cause of X" is not to give a causal explanation of X. It should be clear, however, that no demand for explanatory relevance can make it so that the cause of X is actually <u>causally</u> irrelevant for X. In any case, I am here concerned with causation, not with explanation, so I need not care about explanatory relevance.

There is another reason why some postulate that causes should not be logically related to their effects. Many accounts of the causal relation begin by claiming that an effect of X is something that invariably follows upon or is counterfactually dependent on X. This, however, also applies to cases where something is logically implied without being caused. For instance, whenever an event E occurs, it logically follows that E occurs, but it does not sound right to say that every event must cause itself. Now according to a simplistic counterfactual dependence view this would be true: E would not have occurred, had E not occurred; and therefore, it seems that E caused E (Mackie 1974, p. 32). If one wishes to fix this problem while sticking to the main idea of a dependence account, it is a promising move to postulate that there must be no

logical relation between causes and effects (Le Pore and Loewer 1987, p. 635). This move effectively rules out the problematic cases. However, if there are causes and effects that are not only causally but also logically related, such a contingency principle will throw them out, too. If the principle that causal connections must be contingent is only needed for fixing internal problems of dependence accounts (such as the problem above), it need not be valid in general. If it causes trouble, it should be rejected.

I conclude that the first objection has no basis. It is true that in the case of drowning discussed above, cause and effect are conceptually related. The cause is correctly described as an encounter with water that leads to suffocation, and the effect is correctly characterized as suffocation caused by such an encounter with water. To describe the cause as a cause is to relate it to its effect. There is no valid and general principle to the effect that such conceptual connections must preclude causal ones.

It is of course true that causal relations are not the same as conceptual relations. There is no "Denknothwendigkeit" in causation, as Ostwald puts it (1903, p. 524). No law of thought dictates that there must be instances of drowning; no such law tells us where they occur, and none tells us that all instances of drowning must inevitably take their entire course. But this does not mean that there can be no conceptual relations between causes and effects besides the nonconceptual ones. No law of thought tells us whether there are any triangles, where they are, or whether three lines must combine to form a triangle. This does not mean that we can make no armchair discoveries about triangles. Why should we then think, as Brand seems to, that we cannot possibly make any armchair discoveries about causation? (1975, p. 164) Just like triangles, drownings have essential properties: properties that follow from their definition, or from a general account of the type of process they instantiate. We can look up in a dictionary what it means for someone to drown, and from this we can infer how drownings begin and end. We can thus know some things about the causes and effects of drownings without leaving our armchairs.

Heymans, to whom Ostwald replies, suggests that whenever we discover causal

connections we discover that two things that seemed to be independent facts are ultimately only "two different sides of a singular fact "(1902, p. 483). If this is true and all things in this universe are in some way or other causally connected, all things may ultimately turn out to be different sides of one singular, contingent fact. If we knew this fact and all its different sides, we would also know all things. But of course, we do not.

Let me move to the next objection, that the possibility of giving conceptually related descriptions of a cause and its effect does not imply that they are parts or aspects of the same. We can describe, the objection goes, an instance of sinking into water as the cause of suffocation, and an event of suffocation as the effect of sinking into water; but this does not imply that the sinking and the suffocating are parts or aspects of the same. The descriptions overlap, but the facts may still be perfectly separable (cf. Lewis 1986, p. 174). Mackie defends Hume along such lines. He concedes that it may always be possible to find descriptions of a cause and its effect such that there is a conceptual relation between them, but "[w]hat matters is that there are no logically necessary connections between the events themselves, or between any intrinsic descriptions of them, however detailed and complete" (1974, p. 17).

I argue that causes and effects turn out to be parts or aspects of the same instance of causation. It does not matter how one describes them. One may describe them as causes and effects, i.e., in a way that reveals that they are parts of the same; or one may give two entirely unrelated descriptions of them that do not reveal this. The first is a description of them as cause and effect; the second is a description of them as something other than cause and effect. It is conceivable that the intrinsic descriptions of a cause and its effect are entirely unrelated to one another. If this is true, then their intrinsic descriptions do not reveal that they are in fact parts of the same, and thus they do not reveal that one is the cause of the other. I have no problem with this, and nothing I say above implies the contrary. What I point out above is only that an event such as sinking into water is the cause of suffocation only insofar as it is the beginning of a drowning. This fact is reflected in the fact that in order to describe

sinking as a cause of suffocation, one must describe sinking and suffocating as parts of the same instance of drowning. They do not cease to be parts of the same process if one does not describe them in this way. The only thing that happens then is that one fails to see them as cause and effect. Mackie argues that when two descriptions are conceptually related, it does not follow that their objects are conceptually related. I agree, and add: When two descriptions, intrinsic or not, are <u>not</u> conceptually related, it does not follow that their objects are not parts of the same. That we can describe a cause and its effect in isolation does not imply that they are, as cause and effect, distinct items. It only implies that we can describe them in a way that does not at all reveal how they belong together.

The third objection was that, as long as no one suffocates, what is happening is not yet an instance of drowning, and when the beginning of a drowning occurs, its end may still be prevented. This seems to mean that after all, no drowning might occur, so that nothing can be the beginning of it. This argument involves the assumption that nothing can be a part of a process which has not yet come to its end, and this assumption is false. Suppose that Billy goes overboard and will drown if no one helps. Suppose further that Suzie comes to help and saves Billy from drowning. The question is: Was Billy drowning, although in the end, he did not actually drown? Even though it might seem odd to some philosophers, we must answer that yes, Billy was drowning but then did not drown. For suppose the contrary: that unless someone actually has drowned, no drowning is taking place. This would mean that unless Billy is already dead, he couldn't possibly be drowning. But of course, once Billy is dead, he cannot be drowning either. Since he is either dead or not, and in both cases cannot be drowning, he can then never be drowning. This is more absurd than to say that someone was drowning but actually didn't drown (cf. Rödl 2005, ch. VI). Drownings typically end with someone suffocating. An instance of drowning that is interrupted is an atypical one in that it does not in fact end with anyone suffocating. But still, it is an instance of a kind of process of which suffocation typically is a later stage. This is the reason why its earlier stages are correctly treated as causes of suffocation, even if they might fail to cause suffocation.

5. Conclusion

To sum up, the beginning of a causal process may be taken to be an efficient cause of the rest of it. Sinking is the cause of suffocation because it is the beginning of an instance of drowning, of which suffocation (typically) is the end. At the time at which sinking is causally efficacious, this drowning is already going on. If one separates the sinking from the rest of the drowning by describing it in a way that conceals the fact that it is a part of an ongoing instance of drowning, this connection is lost. As long as sinking and suffocation are taken to be two separate processes, the unity of the drowning of which they are parts cannot be regained by simply relating the sinking to the suffocating. Since an act of relating two distinct events to each other does not yield an instance of causation, causation must be more than a relation between distinct events.

This can also be brought out by considering deviant causal chains. Even if all the distinct stages into which one might split an instance of drowning actually occur, they might be linked to one another in the wrong way. Suppose that someone proposed a definition of drowning as a process that begins with someone's getting wet and ends with this person's being dead. This is obviously not a good definition, for there are many cases in which someone first gets wet and then is dead without there being a drowning. Someone might get stabbed in a shower. It would be less obviously wrong, but wrong nonetheless, to define drowning as a process that starts with someone getting wet, then involves that person suffocating, and ends with her being dead. As a description this may do, but as a definition it would still not be enough. (Someone might get strangled in a shower.) No matter how many discrete steps are inserted, this kind of argument could be repeated. As long as two processes are described in separation, and not as parts of the same larger process, they might as well be parts of a deviant causal chain. In order to see how one event causes the next, one needs to fill the gaps so that the events cease to be separate.

I have shown in this chapter that Hume offers no good reasons for thinking that causation must be a relation, and I have argued that in one case that he discusses,

causation is not a relation between separate and distinct items. Consider the list of causal verbs that Anscombe gives in Causality and Determination:

> A small selection: <u>scrape</u>, <u>push</u>, <u>wet</u>, <u>carry</u>, <u>eat</u>, <u>burn</u>, <u>knock over</u>, <u>keep off</u>, <u>squash</u>, <u>make</u> (e.g. noises, paper boats), <u>hurt</u>. (1981, p. 137)

If I am right, one should not think of any of these as a relation between distinct items. If pushing has the same logical structure as drowning, it is wrong to split it up into distinct stages and then say that the pushing is really only a relation between such stages. First, something gets lost on the way, for not all cases where the movement of one thing causes the movement of another thing are also cases of the first pushing the second. There can always be deviant causal chains. Second, even if one could split up an instance of pushing into separate parts, the problem can be reiterated for each of these parts. The temporal parts of a pushing are, presumably, causal processes, and if all causal processes are relations between distinct stages, there will have to be stages of them, and relations between these stages, and so on. To say the least, I do not think that the strategy of splitting up things into parts and then putting these parts together again leads anywhere interesting.

This is, in a way, Hume's central insight. When we consider causes and effects in complete separation from each other, we cannot see them as cause and effect. Taken in isolation, no cause can reveal any of its causal powers. Contrary to Hume, I infer from this that causation cannot merely be a relation between separate items. Rather, causations are processes, of which causes and effects are different sides.

This chapter has been about Hume's idea that causation must be a relation between distinct items, rather than Aristotle's efficient cause. I have not said much about Aristotle's efficient cause. In Chapter 1, I argue that although Aristotle often speaks of the efficient cause as an agent that produces something, this agent can only be a cause of what it produces insofar as it acts. Therefore, the efficient cause is a thing that acts,

and its (proximate) effect is an action of this thing.[9] If causation is not a relation between distinct items, it follows that what an agent does in order to cause something is not distinct from what this agent causes by doing it. To be sure, the effect of what the agent does is something further, but it is not, as effect, distinct from its cause. It is the continuation of the process that the agent has started. What the agent does and what it causes are causal processes, and causing one by doing the other is a further causal process, of which the first two are parts.

---

[9] Aristotle usually cites cases where an efficient cause has more distant effects, such as a process undergone by a different thing, or a thing that results from this process.

Causal Processes

I have argued that causation is not a relation between distinct items. Although one may split up instances of causation into separate and distinct segments, one should not expect to be able to put these segments together again and thereby get back what one started with. This would work only if causation were nothing but a relation between separate items to begin with, which it is not. Characteristically, one can separate a cause and its effect from each other only to the extent to which one abstracts from the cause being the cause of the effect. To consider a cause in isolation is to consider it insofar as it is not a cause of anything, and it is no wonder that Hume could not discover the causality of a cause by abstracting from its being a cause.

This does not mean that there are no causal relations. It is in many cases enough to know how two factors are related, counterfactually or probabilistically, without knowing much about the causal processes that connect them. Some such dependencies are rightly called causal ones, and the question as to under what circumstances a relation is a causal one is important. However, this question is not the same as the other important (and more fundamental) question that I am going to tackle in this chapter: under what circumstances a process is a causal process. I consider three possible ways of answering this question. First, one may distinguish between causal and non-causal processes by relying on what scientists tell us about causal processes. This approach is in the same situation as definitions of matter on the basis of physical theories about matter. Scientists can tell us many important things about causal processes, but we cannot simply use what they tell us in order to define what they are talking about. A second possible way of distinguishing between causal and non-causal processes is to ask what the verb "cause" means and how it relates to specific causal verbs such as "push" and "scratch." My discussion of this approach leads to the result that the verb "cause" has no content of its own. One may describe all causal processes by using this verb, but in the sentences that result the verb

"cause" turns out to be entirely redundant. In a third attempt, I approach the question what causal processes are by asking why we are interested in such processes. This leads to a very general account of what it means for a process to be causal. Roughly, it turns out that a process is causal if it instantiates a type of processes that admit of a certain kind of complex description.

Like the previous chapter, this is not a chapter about Aristotle. Its result is an account of causal processes that is compatible with Aristotle's notion of an efficient cause as outlined in Chapter 1.

1.   Causal Processes

It seems that not everything that happens is an instance of causation. What is the difference between a causal process and a non-causal one? Strawson writes that causal processes are instances of "action and transaction which are directly observable in the particular case and which are properly to be described as causal in so far as they are varieties of bringing something about, or producing some effect or some new state of affairs" (1985, p. 120-1). However, as Hume would have objected, "bringing about" is clearly a causal notion, and so it should not be used in order to explain what causation is. It is too close to the explanandum. We are left with the idea that causal processes are actions or transactions, which have a certain general structure that remains to be specified.

Non-causal processes have been referred to as "illusive sequences" (Frankl 1907); Salmon calls them "pseudo-processes" (1998, p. 287). They only look like causal processes but are none; something is missing. For instance, when a movie of one ball pushing another one is projected on a screen, several causal processes are going on, but the pushing is not one of them. Whatever one may do to the image of the first ball's movement, it does not affect the image of the second ball's movement. Had the image of the first ball's movement not been in the movie, the second image could still be in it. There is no transfer of energy or momentum from the first to the second. The images of different stages of a movement form a sequence, but although the stages of

the original movement were parts of a causal process, the <u>images</u> of these stages are not. The sequence is illusive. Its elements are not linked in the right way; they are only images of something that is linked in the right way. The question is: What has reality got that the movie doesn't?

An idea that keeps re-emerging in the literature in one form or another is that in a causal process, something must persist. Let me name a few arbitrarily selected instances of this view. Meyerson thinks that the crucial principle is that <u>causa aequat effectum</u>: In some sense, cause and effect have to be the same (1926, p. 19). Russell tells us that what he calls a causal line "may always be regarded as the persistence of something—a person, a table, a photon, or what not." He continues: "Throughout a given causal line, there may be constancy of quality, constancy of structure, or gradual change in either, but not sudden change of any considerable magnitude" (1948, p. 477). Mackie writes that "a singular causal sequence instantiates some pure law of working which is itself a form of partial persistence" (1974, p. 229). Ehring maintains that causal processes are characterized by particular property instances that persist throughout the duration of the process (1997, p. 121).

Another prominent idea is that causal processes involve some kind of transfer. For instance, Aronson claims that "what underlies the use of mechanical transitive verbs is a process of transference which occurs between the cause and effect objects" (1971, p. 422). Fair writes that the connection between cause and effect consists in a "flow of energy" from one to the other (Fair 1979, p. 231). These formulations are still problematic because "transference" and "flow of energy" are causal locutions. Transferences and flows are instances of causation, and it does not help much to define causation in terms of one of its instances. The crucial step that Salmon takes in this situation is to simply identify transference with a kind of persistence. He thus takes up both ideas: Causal processes involve a transfer of some quantity, but this transfer is nothing but the continued manifestation of this quantity. As Salmon uses the term, "transfer" ceases to refer to a species of causation (1998, pt. III). However, he is not always clear about what is being transferred, that is, continuously

manifested. He writes, for instance, that "causal processes constitute the mechanisms by which causal influence is transmitted" (1998, p. 293). It is unclear to me how a process can be a mechanism, but more importantly, Salmon characterizes causal processes as processes by means of which <u>causal influence</u> is transmitted, and this is quite circular.

Dowe gives a clearer exposition of basically the same idea. According to his account, a causal process is "a world line of an object which possesses a conserved quantity" (1995, p. 323). (The context makes clear that "which" refers to the object, but presumably there is not much of a difference between an object and its world line.) Causations are thus histories or partial histories of objects that possess some conserved quantity. A conserved quantity is "any quantity which is universally conserved, and current scientific theory is our best guide as to what these are" (ibid.). I take it that what Dowe says about causation is correct. Science tells us that some quantities are universally conserved; that is, they are conserved throughout all physical processes. This means that whenever a process is causal, then it is the world line of an object that possesses a conserved quantity, and vice versa.

The problem I have with Salmon and Dowe is that their accounts seem to have the methodical order reversed. That certain quantities are universally conserved is something that scientists have found out <u>about</u> causal processes.[1] As Fair writes: "For a large class of cases, physics has discovered that there has been a transference of the physical quantities, energy and momentum, from the cause to the effect" (1979, p. 228). Scientists have discovered this about causal processes by studying them. In order to study causal processes, rather than something else, they already had to be able to tell causal processes from non-causal ones.

---

[1] This point is compromised if the advanced sciences are not actually interested in causation (Russell 1912; Norton 2003, p. 21). If they are not, scientists do not any longer make or rely on any discoveries about causation. This would make it even more obscure how the discoveries of advanced science could be of any use in an explanation of what causation is.

Salmon and Dowe's criterion works in much the same way as a pregnancy test. We know what pregnancy is. Scientists investigated instances of pregnancy and determined that there is something that we did not yet know that they all have in common. On this basis, they developed a test that reliably tells us, in cases where it is not obvious, whether someone is pregnant. Although this test is very useful, no one thinks that it can tell us what it <u>means</u> to be pregnant. A pregnancy test may tell us when someone is pregnant but it does not tell us what pregnancy is. For instance, that pregnancy tests test for the presence of chorionic gonadotropin does not mean that pregnancy consists in the presence of chorionic gonadotropin. The same holds true of Dowe's criterion. Scientists usually know causal processes when they see them. They have found out that all such processes have something in common: Certain conservation laws apply to them. On the basis of this discovery, one may develop a test that tells us, in cases where it is not obvious, whether a process is a causal one or not. But this is only a means of probing for causal processes; it need not tell us what they are. And in fact it doesn't. There must be a more fundamental way of characterizing causal processes. Otherwise, the statement that scientists have discovered that conservation laws apply to causal processes would end up being a tautology.

Moreover, Dowe's criterion does not seem very practicable to me. In order to find out whether something is a causal process, we are supposed to ask whether this process is the world line of an object and whether this object possesses conserved quantities. But in most cases, finding out whether something is a causal process is either simply the same as finding out whether it is the world line of a physical object, or it is easier. Again, there must be ways of understanding what causation is that do not depend on the notions of a world line and of conserved quantities.

2. "Cause" as a Dimension Word

A second possible way of answering the question what causal processes are is to ask what the verb "cause" means. Aronson writes that "cause" is a dimension word

for transitive verbs (1971, p. 417). He takes this term, "dimension word," from Austin, who introduces it as follows: A dimension word is "the most general and comprehensive term in a whole group of terms of the same kind, terms that fulfil the same function" (1962, p. 71). For instance, "real" is a dimension word since it is the most general word in a group of words of the same kind that fulfill the same function. Other, more specific members of this group are "proper," "genuine," "live," "true" etc. One can take a characterization of something as proper and replace it with a characterization of the same thing as real. Since both terms have the same function, the meaning of the original phrase will not be dramatically changed. However, because "real" is a more general term than "proper," some of the meaning may get lost. One cannot conversely take all occurrences of "real" and replace them by "proper."

A minor problem with Aronson's account is that he restricts causal verbs to transitive ones. This may seem uncontroversial. After all, the examples that Aristotle gives of causal processes (sculpting, counseling, procreation, curing) involve transitive verbs, and many other authors also seem to think that causal processes always involve two things, one of which acts on the other. Dullstein, for one, maintains that all causal verbs stand for types of interactions between an agent and a patient (2007, p. 103). However, I do not think that we should confine ourselves to transitive causal verbs. I prefer accounts that also count such (intransitive) verbs as "falling," "growing," and "sliding" as causal. Once we understand these cases, it will be easy to extend our account to cases that involve more than one thing. Some authors distinguish between immanent and transeunt causation (Mackie 1974, p. 156). Transeunt causation is what happens when one thing does something to another one. Immanent causation is, as Russell has it, "interpreted as the persistence of a thing or a process" (1948, p. 494). I do not want to exclude the possibility of immanent or intrinsic causation because I think it is the more basic case, on the basis of which transeunt causation can be understood. Therefore, I do not agree with Aronson that all causal verbs must be transitive.

The more important problem with the suggestion that "cause" is a dimension word for causal verbs is that "cause" does not in general work in the same way as more specific causal verbs, and therefore is not the most general and comprehensive instance of their type. This is so even if we consider only transitive verbs, such as the following:

> ... bend, break, burn, carry, eat, heat, hurt, irritate, keep off, kill, knock over, melt, prevent, push, remind, scrape, squash, steer, twist, wet, ...

I have compiled this list from Ducasse (1957, p. 422) and Anscombe (1981, p. 137). Let me call it "Anscombe's list." The question is: How does "push," for instance, relate to "cause"? Godfrey-Smith writes that "the essential semantic role of the word 'cause' in languages like English is to collect together a range of more specific relations that we also pick out with words like 'squash,' 'push' and so on" (forthcoming). I do not think that any one of the items on Anscombe's list is a relation, but still, the word "cause" seems to somehow collect them. How does it do this?

It should be clear that "causes" is not simply a genus of which the verbs listed above would be species. Purchasing, for example, is a species of acquiring. Purchasing something is a way of acquiring it; when someone purchases a thing, she thereby acquires this thing. Pushing a thing, however, is not a way of causing this thing. When someone pushes a thing, she does not thereby cause it. Therefore, pushing is not a species of causing. If "cause" is a generic term for causal processes, as Godfrey-Smith suggests, it must be generic in a more complicated sense.

One might think of Anscombe's list as an open referent class, of which "causes" is the focal instance.[2] The verb "to cause" would then collect the verbs on Anscombe's list by being this focal instance, to which all other cases of causation would be related in some way to be specified. Alternatively, Anscombe's list might be thought to be the

_____

[2] Cf. Chapter 5, section 2, for the notions of a reference class and a focal instance.

beginning of a cluster, such that a process is causal if and insofar as it bears this relation to any arbitrary causal verb.[3] However, what ties the instances of Anscombe's list together cannot be that they all resemble the word "causes" or that they all have the same function as this word. For, far from being a paradigmatic item on the list, the word "cause" does not even fit in there. When A bends, breaks, or burns B, A does something to B. When A causes B, B results from A. Bendings are much more similar to breakings than they are to causings.

Consider, in contrast, the following quote from Hart and Honoré's Causation in the Law:

> Preoccupation with the familiar pair of terms "cause and effect" may make us think that there is a single concept of "causation" awaiting our inspection and that the huge range of other causal expressions, "consequence," "result," "caused by," "due to," "lead to," "made," etc. are mere stylistic variants. (1985, p. 27)

What Hart and Honoré list are in fact <u>species</u> of causes, effects, and causal processes. "Effect" is a genus of which "consequence" is a species. I am here interested in a corresponding list of species of causation, which would include "result in," "lead to," and the like:

> ... result in, produce, generate, bring about, create, lead to, make, ...

_____

[3] Skyrms calls causality a cluster concept (1984, p. 254), and Longworth expounds a "broadly pluralistic theory of token causation, according to which causation is a 'cluster concept' with a 'prototypical' structure" (2006, p. iii). However, neither of them has in mind a cluster of instances of causation. Rather, they propose lists of criteria, such that a process is causal if it satisfies enough but not necessarily all of them. What they have in mind are clusters of theories of causation.

I refer to this as "Hart & Honoré's list." It is a list of specific ways of causing something. Not every instance of causing something is an instance of creating this thing, but conversely, every instance of creating something is an instance of causing it. Creating is thus a species of causing. In contrast, the items on Anscombe's list are not specific ways of causing something but rather specific ways of <u>acting on</u> something. Whenever A pushes B, A acts on B, although not every instance of A acting on B is one of A pushing B. Therefore, pushing is a species of acting on. Although causes may be said to act, and agents may be said to cause something, causing is not the same as acting. When an agent acts, she causes something by doing something else. She does not, strictly speaking, cause her own action. She could cause her own action only by doing something that causes it (cf. Ducasse 1932, p. 396-7). In any case, Hart & Honoré's list covers only a small subset of causal processes, which means that only a small subset of causal verbs actually work in the same way as the verb "causes." We still do not know what it is, in general, that causal verbs signify, including the verbs on Anscombe's list.

3.   Aronson's formula

Aronson does not simply stick to his characterization of "cause" as a dimension word for transitive verbs. He develops a more detailed account of the relation between "causes" and the items on Anscombe's list and proposes the following rule:

> ... a sentence of the form
> <u>Noun Phrase</u> + <u>Transitive Verb</u> + <u>Direct Object</u> + <u>Objective Complement</u>
> can be replaced by
> <u>Noun Phrase</u> + <u>caused</u> + <u>Direct Object</u> + <u>Copula</u> + <u>Objective Complement</u>.
> (1971, p. 417)

Aronson immediately concedes that not all transitive verbs lend themselves to this treatment; for instance, "He found my remarks intelligible" does not mean that he in

any sense caused the remarks to be intelligible. But let us assume that such cases can be effectively excluded. Also, let me simplify Aronson's formula a bit. He suggests, basically, that where "___" is a causal verb and "---" its objective complement, we may infer:

A ___ed B ---   =>   A caused B to be ---.

Instances are:

A pushed B to the left   =>   A caused B to be on the left,

A knocked B off the table   =>   A caused B to be off the table,

A painted B white all over   =>   A caused B to be white all over.

There are, to be sure, several cases on Anscombe's list that do not fit in here because they lack objective complements. By further narrowing down our focus to transitive verbs with objective complements, however, we obtain at least some kind of genus-species relation. Pushing something may not be an instance of causing it, but at least pushing something to the left is an instance of causing it to be on the left.

Now that statements that involve causal verbs imply statements that contain the verb "cause" does not tell us much about causal verbs. All kinds of things may imply statements that contain the verb "cause." We would be in a far better position if we could expand the implications above to equivalences. Consider the following attempt:

*A pushed B to the left   <=>   A caused B to be pushed to the left.

This is still not enough because it leaves room for deviant causal chains (whence it is marked with an *). For instance, the description on the right side still fits cases where A did not push B but caused something else to push B to the left. In order to fix

this once and for all, however, one would have to put the same terms on each side of the "<=>":

    A pushed B to the left   <=>   A caused B to be pushed to the left by A.

In general:

    A ___ed B ---   <=>   A caused B to be ___ed --- by A.

One thing that this shows is that in order to capture the full meaning of a statement such as "A pushed B to the left" by using the two-place term "X caused Y," one needs to relate two terms to one another that are logically related. (Something very similar is shown in the previous chapter.) Another thing is that in the resulting expression, the verb "cause" is simply redundant. In the sentence "A caused B to be pushed to the left by A," it does not serve any purpose. It might as well not be there.

Thus it turns out that the verb "cause" can only be redundantly added to a causal verb and cannot stand in for any of its content. This is not trivial, and it tells us something about the meaning of this verb. It tells us that it can be added to causal verbs without at all changing their meaning, and this is at least a partial answer to the question of what makes causal verbs causal. Conversely, a causal verb is a verb to which one may add the verb "cause" in the way specified above, without adding any content. One can do this for all verbs on Anscombe's list: A caused B to be bent by A, A caused B to be broken by A, and so on. Whatever "to cause" signifies, it is something that is already expressed by each of the causal verbs, hence the

redundancy.[4]

To be sure, that the verb "cause" does not add any content to a causal verb also means that it does not help much to replace expressions that contain causal verbs by expressions that contain the verb "cause." To apply the phrase "X caused Y" to all causal processes alike only <u>looks</u> like giving a unified description of all of them. Putting the verb "cause" in these phrases is "no more than an exercise in labeling" (Norton 2003, p. 3). The verb "cause" is only a flag; it has no further meaning of its own. Therefore, analyzing its meaning does not tell us what causation is.[5]

4.   A Note on Diagrams

That the verb "to cause" has no meaning of its own leads to a corollary concerning causal diagrams. It is customary in the current literature to represent causal processes and relations by diagrams that consist of labeled nodes and unlabeled arrows:

   A   ------------>   B
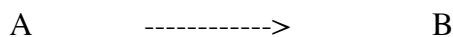
In such diagrams, the nodes (e.g., A and B) stand for events (Spirtes et al. 2000, p. 21); changes in values of variables (Woodward 2003, p. 112); states at times, entities,

---

[4] If the four causes are concepts of reflection (as suggested in Chapter 2, section 6), this is what one should expect. One should not think of causal relations as further empirical objects. Two different representations of a causal process may be related to each other as determinable to determination, but the concept of reflection by means of which this is done does not represent a further object.

[5] Austin observes a similar tendency in action theory: to replace attributions of actions by expressions of the form "A is doing X" (1979, p. 178). Like "is causing," "is doing" is entirely redundant here. (One may just as well say: "A is Xing.") An analysis of the meaning of "doing" would not tell us what an action is.

or properties of entities (Norton 2003, p. 16); or simply everything that can be referred to by a variable (Pearl 2000, p. 12). The arrows are supposed to have a constant meaning: They refer to causal relations among the things referred to by the nodes (cf. Cartwright 2007, p. 21). They may as well bear the label "causes."

In order to describe a causal process by drawing such a diagram, one needs to split it up into different segments that are causally related. In order to do this, one needs to somehow shovel all specific content of the verb "to push" into one or both of the nodes. In a first approximation, one might come up with a diagram in which two movements are related, such that one of them causes the other:

    A is moving   ---causes--->   B is moving

But here, the content of the verb "to push" is almost entirely lost. Having no content of its own, the label "causes" does not make up for any of it. There are many ways in which a movement may cause another movement, and many of them are not pushings. In order to capture the specific content of "to push," one may draw any of the following diagrams:

    A is pushing B   ---causes--->   B is moving
    A is moving   ---causes--->   B is being pushed by A
    A is pushing B   ---causes--->   B is being pushed by A

But here, of course, "A is pushing B" remains quite unanalyzed.

Note that the way in which Aronson states his rule suggests a somewhat different model, according to which things, such as A, cause events, such as the event of B being pushed. If his strategy would work out, one might be tempted to say that all causal processes have the following general form: One thing causes something to happen to some other thing.

A   ---causes--->   B to be pushed by A

This model of causality has a long tradition. As noted in Chapter 1, Aristotle often treats things and agents as efficient causes. The Stoics also generally maintain that a cause is a bodily thing and its effects something that can be said, that is, a predicate or a state of affairs.[6] According to Watkins, this is also the model that Kant still employs (2005, ch. 4). There are a number of objections against thinking of causation, in general, in this way.[7] For one thing, if causes are things but effects are events, and if no thing is an event, then there can be no causal chains. In such a chain, something would have to be the effect of something and at the same time the cause of another. If causes were always things and effects always processes, it would have to be both at once. Further, in most cases where we say that a thing causes something to happen, the thing is also fully present before and after it causes the happening. But then it seems that it cannot be the sufficient cause of what happens. If it were, it would always cause the happening whenever it is present. I mention these objections only to put them aside. I am not endorsing a model of causation according to which things cause events. Aristotle, for one, does not in general propose such a model.[8]

However this may be, it follows that drawing diagrams with only one kind of arrow is a bad way of representing causal processes. A far better strategy is the one

---

[6] Sextus Empiricus, Adversus Mathematicos 9, 211. Cf. Frede 1987, p. 137; Bobzien 1999.

[7] Cf. Chapter 1, section 4, and the end of section 2 above.

[8] Aristotle writes that a sculptor is, as such, only potentially the cause of the statue (Physics II 3, 195b3-6 and b16). Potentials, however, are strictly speaking only the causes of further potentials, not of actualities (195b27-28). In De Generatione et Corruptione II 9, 335b18-20, Aristotle also points out that the forms cannot cause a change if they do not change; accordingly, a sculptor should not be a cause unless she does something that causes the statue.

that Machamer et al. describe (2000, p. 3). In their diagrams, the arrows do not stand for bare causations or causal relations but for specific causal processes, such as the ones on Anscombe's list. For instance:

   A   ---pushes--->   B

   Here, A is not represented as causing B. Rather, something that A does is represented as affecting B.[9] Further, the nodes in such diagrams refer to things and not to processes or events. The items that are related by the causal activities represented by nodes are things, not events. All events are represented by arrows. Therefore, there is no categorial asymmetry between causes and effects, and there can be causal chains. Further, since all nodes refer to relatively permanent items, temporal relations do not hold between the nodes but only among the arrows. For instance, one may depict as follows that when A pushes B, then, a little later, B will scratch C:

   A   ---pushes--->   B   ---scratches--->   C

   This does not at all suggest that whenever A is present, it must be pushing B. There is no causal relation and thus no temporal relation between A and its pushing, or between A and B. There is a causal and thus temporal relation between pushing, represented by one arrow between A and B, and scratching, represented by the arrow that connects B and C. The causal relation between these two processes is represented by the entire diagram, not by a further arrow.

5. Types and Handles

   My question is: How do causal processes differ from other processes? I have

---

[9] Lewis follows this convention: In his diagrams, the nodes always represent things, and the arrows stand for what these things do to other things (e.g., 1986, p. 200).

briefly considered accounts according to which processes are causal if certain conservation laws hold true of them. I reject these accounts not because they would be false or useless, but because they turn the claim that the respective conservation laws apply to causal processes into a tautology. According to such accounts of what causal processes are, no one can possibly discover (or fail to discover) that causal processes are governed by conservation laws. I have then raised the question how the verb "to cause" relates to causal verbs such as the ones on Anscombe's list. It has turned out that "to cause" does not refer to a genus of which pushing, scratching, etc. are direct instances. Therefore, it is also not simply a dimension word for transitive verbs, as Aronson suggests. It is not the most general and comprehensive term on Anscombe's list. If it is a dimension word, then it is one for the items on Hart & Honoré's list, but this list contains only a small subset of causal verbs.

There is still one element of Aronson's suggestion left, which I want to take up now. A dimension word characterizes a range of verbs that have the same <u>function</u>. Let me therefore ask: What is the general function of causal verbs? Paraphrasing Anscombe: We seem to have a special interest in causal verbs. What is it we have a special interest in here? (Cf. 1957, §46)

I can think of three main sorts of things that one generally wants to do with causal processes and relations: explain, predict, make use of. All three of them involve a move from somewhere to somewhere else. It makes no sense to explain A by reference to A; it makes no sense to predict A on the basis of A; and when we make use of A, we want something other than A in return.[10] I take this to be the main reason why people think that causation must be a relation between different items. Whenever we explain, predict, or make use of something, we relate something to something else.

_____

[10] "Make use of" (or "utilize") is a somewhat technical expression here. I want it to correspond to the Latin <u>uti</u>, as opposed to <u>frui</u> (as distinguished by Augustine, De Doctrina Christiana I 4,4, CSEL 80). It stands for the use that one makes of something in order to achieve something further, not to the use one may make of something for its own sake.

However, as I have argued, causal processes should not be taken to be relations. There are causal relations, but such relations are not all there is to causality. Causal relations are backed up by causal processes, such that two items are causally related insofar as they are involved in the same instance of a causal process, and these causal processes are not relations. Therefore, we need to understand how something that is a process and not a relation may still lead us from somewhere to somewhere else. And this is not difficult to understand.

Collingwood writes about one of the uses that we make of causal processes that "the cause of an event in nature is the handle, so to speak, by which human beings can manipulate it" (1940, p. 296). Let me adapt this metaphor and say that in general, what we want from a causal process is that it provides us with two handles. One of these handles is, as Collingwood has it, the cause; the other one is the effect. We may grab one of them, as it were, and the process will then give us another handle in return. For instance, once we know that turning on the light begins with flipping the switch, we can flip the switch and get light as a result. As Schlick emphasizes, this does not merely mean that causal processes may be described by functions that map some input on some output. Non-causal processes, such as the sequence of images in a movie, may also be described by such functions. Schlick writes that in addition to being representable by functions, causal processes must provide a possibility of calculation, and that this "does not mean possibility of finding a function with particular mathematical properties, but it means possibility of applying a function with any properties to such data (or 'events' or 'states of a physical system') as have not been used for its construction" (1979, p. 253-4). We do not only want to map a given input on a given output; we want to be able to calculate, estimate, or achieve an outcome that is not yet given, on the basis of an input that is given. We can do this on the basis of our general knowledge about the type of process or mechanism we are dealing with.[11]

---

[11] I am here arguing that causal processes must be such as to make prediction possible, not that prediction is only possible on the basis of causal knowledge.

Causal processes must be such that one may characterize them as instances of a generally specifiable type, and then use one's general knowledge about this type in order to predict, explain, or bring about something that one has not already relied on in order to identify them as instances of this type.[12] This simply means that causal processes instantiate types. As instances of types, they can be known to possess a considerable number of features beyond those that are currently manifest. To identify a process as causal is to bring it under a specific type or to assume that it can be brought under a type, such that one may know and state more about the process than what one immediately observes.

This might very well be what Kant means when he claims that all processes must proceed according to the law of cause and effect. He argues, in the Second Analogy of Experience, that in order to be able to distinguish between a subjective sequence of impressions and the objective process that it represents, we must assume that the objective process proceeds according to a rule, which marks a certain sequence of subjective representations of it as correct and others as incorrect (Critique of Pure Reason, B 238). This means that the process must be an instance of a type of processes, which may be characterized by stating the rule according to which such processes proceed. The principle that every event must have a cause follows from this as a corollary. Kant argues that all objective processes must be parts of one and the same world history and therefore every process must be a part of a more comprehensive objective process. This more comprehensive objective process must also proceed according to a rule, and this implies that every process must follow upon another process according to a rule.

That a process is an instance of a type means that one may reasonably expect it to take the course (or a course) that is typical for processes of this type. There are in general two good reasons for classifying processes into types and thus expecting something to happen before it has actually happened. First, intentional actions may be

---

[12] Schlick refers to this as extrapolation and prediction. In a parallel passage he adds that the "predicted" outcome need not lie in the future (1979, p. 185).

classified in terms of the intentions with which they are performed. One may know what will happen by knowing that someone intends to make it happen and is capable of doing so. Second, natural processes (which are not intentional actions) may be classified into types according to the principles and motion and rest that govern them. In this case, one may expect something to happen because one knows that something is already happening that is of a type, instances of which typically lead to something further. What is happening now is seen as the beginning of a process, of which what is expected is a later stage. One may know this without knowing anyone's intentions, by knowing the relevant principles of motion and rest inherent in the natural things that are involved in the ongoing process. These principles determine what one may expect under normal circumstances.

Note that the second reason to expect something to happen is more fundamental than the first. One can only intend to make something happen if there is a way of making it happen, i.e., if it is possible to start some process, of which that which one wants to cause typically is a later stage. Intentional actions rely on natural processes. On the other hand, one may intentionally interrupt natural processes and manipulate their course by exploiting other natural processes and links that interfere with the first. When this happens, the outcome may not be predictable on the basis of general laws that apply to natural processes. I am here only concerned with natural processes and the expectations they license in the absence of intentional interference.

6. Conclusion

If the verb "cause" can only redundantly be added to an expression involving a causal verb, one should not expect to be able to say much about its content. It does not have much content, and therefore if we had a good definition of it this definition should also not have much content. Here, I have only tried to bring out why causal processes are important to us. However, what I have said does give us the beginning of an account. Causal processes are such that one may do one thing with them by doing another thing. It remains to be said how exactly they make this possible, but

two things are already fairly clear. First, causal processes must instantiate some kind of rule-governed pattern. They must be instances of types, such that their different stages are connected by general rules. Their stages must be stages of the same. This motivates the idea that cause and effect must be somehow the same, or that at least certain quantities or other features must be conserved throughout a causal process. In fact, that which must remain the same over time is simply the process itself. When a process is causal, we must be able to make sure that throughout some period of time, it is still going on. Accordingly, there must be some way of identifying this process as the same process at different times. In the case of natural processes, this is possible on the basis of the principles of motion and rest that govern it. The second thing that has become clear is that every causal process must admit of two logically independent characterizations, so that one may identify it as an instance of its type by means of one characterization and then use another characterization in order to predict, explain, and make use of something on this basis. In this sense, causal processes must have the structure of what Aristotle calls ποίησις: they must be such that when they occur, something leads to something further.

This might well be all one can say about causation in general.

Basic and Derived Final Causes

In the previous chapter, I maintain that causal processes are processes that instantiate types of processes. In the case of natural processes, the reason why they instantiate types is that they are governed by the principles of motion and rest inherent in the natural things that are typically involved in them. As instances of generally specifiable types, causal processes have essences. Natural processes have essences because they are governed by principles of motion and rest inherent in the natural things that are involved in them. I also claim that causal processes must provide "two handles," that is, it must be possible to identify them as instances of a type on the basis of one set of facts about them, so that one may then expect further facts to be true of them. This means that the essences of causal processes must be complex. For instance, thunder is a causal process, the essence of which is noise in the clouds as caused by the extinction of fire in the clouds. This essence is such that one may distinguish a cause from an effect within an instance of thunder.

In Chapter 1, I argue that the material cause of a natural thing relates to its formal cause in the same way in which the efficient cause of a natural process relates to the final cause of this process. The material cause of a thing is that which potentially is this thing. The formal cause is what the material cause potentially is. Likewise, the final cause of a natural process is what its efficient cause potentially is. That one ball causes another ball to move means that the first movement is the beginning of a process, of which the second movement is a later stage. The first movement is potentially this process, the second movement is (a part of) the realization of this potential. The potential of an efficient cause is fully realized if a process of a certain type has completed the course, or one of several possible courses, that instances of its type typically take. The efficient cause of a natural process is thus potentially a completed process. The final cause of this process is what its efficient cause potentially is, i.e., what the process would be if it were complete in a fully typical

way. To describe the final cause of a natural process is to describe the course that it typically takes.

Sprigge asks:

> Is the distinction between a teleological law and [an] efficient causal law perhaps just a difference in the way in which a certain recurrent pattern in the world is described, certain descriptions expressing an interest in how processes arise or can be made to do so, while other descriptions express rather a dramatic interest in the upshot of such processes? (1971, p. 158)

If final causes relate to efficient causes as formal causes relate to material causes, then the answer is yes. When a ball sets another ball in motion, the movement of the first ball is a process that causes the motion of the second ball. Both are parts of a larger process that naturally leads from one to the other. To treat the movement of the first ball as a cause is to treat it in terms of what it leads to, and what it leads to is its final cause.

In this chapter, I first take up a claim that I make in the first chapter, that the final cause of a natural process is its limit, and argue for it in more detail. Then I discuss three objections to the idea that final causes are essences of natural processes. I conclude with a few remarks on the issue of reducibility.

1.   Final Causes as Limits

A lot has happened since I last discussed the final cause in this book. It will be good to recall some of the things that I have said, and elaborate on them. In Chapter 1, I point out that an account of the typical course of a process will in most cases involve more than just a specification of its result. In Nicomachean Ethics X 4, Aristotle writes that a motion is perfect either when it is considered in its entire duration, or in the moment when it has reached its end (1174a21-22). In either case, the process is only complete when it has taken its entire course. It need not be complete when only

its final stage occurs because its final stage may have occurred for reasons that have nothing to do with the completion of this process. Therefore, the final cause of a process cannot just be its final stage. It must be the entire typical course that leads to this final stage. A process has reached its final cause not when it ends in a certain way but only when it reaches this end in a certain way.

I also say in Chapter 1 that the final cause of a natural process is its essence and that in one sense of "limit," this essence is the limit (πέρας) according to which it proceeds (περαίνειν). Aristotle does not explicitly state that final causes are essences of processes. It can be demonstrated by showing two things: that final causes are limits of processes, and that the limit of a natural process is, in the sense of "limit" that is relevant here, the essence of this process.

In De Motu Animalium 6, Aristotle justifies the claim that all movements of animate things have a natural limit by arguing that all animals move or are moved naturally for the sake of some good (700b14–35). He writes that living beings are responsible for the motion of everything, except the motion of the entire universe and of things that are merely pushed by others (700b11-13). Because the motions of all ensouled beings have a limit, it follows that all active, sublunar motions have a limit (cf. Metaphysics B 4, 999b9-11). When Aristotle says that all active, sublunar motion has a limit, he might simply mean that all such motion will eventually stop. However, when he argues that the movements of living beings are always for the sake of something, the reason he gives for this is simply that the limit of their movements <u>is</u> that for the sake of which animals move. Aristotle writes:

> For all animals both impart movement and are moved for the sake of something, so that this is the limit to all their movements: the thing for-the-sake-of-which (ὥστε τοῦτ' ἔστιν αὐτοῖς πάσης τῆς κινήσεως πέρας, τὸ οὗ ἕνεκα). (700b15-16, tr. Nussbaum)[1]

---

[1] Cf. also Metaphysics α 2, 994b15-6.

In Physics I 2 (185a33-b2) and III 5 (204a29-30), Aristotle writes that only quantities can lack a limit, and that "unlimited" cannot be said of things but only of their accidents. The reason is that, as he says in Metaphysics Δ 17, the essence of a thing is its limit, and therefore, nothing that has an essence can be unlimited, except accidentally. According to Δ 17, the limit of a thing is the first outside of which there is no part of this thing, and the first inside of which there is all of it (1022a4-6). This implies that the limit of a process is not simply the point where it stops, for this point does not include all of the process. The limit of a process can only be the entire course that it takes.[2] In one of the senses that Aristotle lists, the limit of a thing is its essence (1022a7-9).

The essence of a thing or process is that in terms of which it may be identified as an instance of a type. The limit of a thing or a process is thus what makes it one definite thing or process. Since the spread of fire has no limit, as Aristotle says in De Anima II 4 (416a15-17), it is not a unified process. It has no essence. Further, the essence is what is defined in a definition, and in order to define a process, one must specify more than its end point. If final causes are limits and limits are essences, final causes can be taken to be essences of processes. The final cause is not final in the sense of being the last thing in a sequence; it is final in the sense of giving a process definite shape.[3]

There are further passages in which Aristotle identifies the limit of a process with

---

[2] Nussbaum writes that in 700b13-14, Aristotle uses πέρας in the sense of "final cause," whereas in 700b31-32, it stands for the point where something stops (1985, p. 332-3). Corcilius objects that in both passages Aristotle refers to the same, namely the completion (Abschluss) of a movement (2007, p. 29-6 fn. 18). I agree with Corcilius that Aristotle means the same in both places and argue that this is neither the purpose nor the endpoint of a movement. In both passages, he is referring to the typical course or essence of a process.

[3] In De Primo Principio cap. II, Duns Scotus acknowledges this when he calls that which relates to the final cause as the effect to the efficient cause the finitum.

its final cause. In De Generatione Animalium I 1, he tells us that "nature flees the limitless (ἄπειρον), for the limitless is without end (ἀτελές), and nature seeks an end (τέλος)" (715b14-16). The end that nature seeks when it flees the limitless should be the limit. That nature seeks an end means that all natural processes have a final cause, and as I have argued, the final cause of a process is the course that would be typical for it to take. This is in fact what Aristotle also says in Physics VIII 1: That which occurs by nature always follows a certain order (252a11–12).

In Metaphysics Θ 6, Aristotle claims that actions that have a limit (πέρας) are always means to an end and never ends in themselves (1048b18-19). On the face of it, this seems to mean that the end of an action is not its limit, and that to occur for the sake of an end is not the same as to have a limit. But that an action is not an end does not imply that it has none, and the end it has may very well be the same as its limit. If the end is the same as the limit, then Aristotle is claiming that actions that have a limit are not their own limits (they might be a limit of something else). In Chapter 6, section 1, I point out that causes cannot coincide with what they explain. This is also true of the causes of processes. One cannot explain a process that coincides with its final cause in terms of this final cause. Ἐνέργειαι, which coincide with their completion, may have essences, but these essences are not their causes. Κινήσεις do not coincide with their end, since as long as they occur, their end has not been reached. By the same token, they do not coincide with their limit, since as long as they occur, the pattern according to which they occur has not been completed.

Further, Aristotle says in Physics II 2 that all those changes are for the sake of something that might be due to thought (196b21-22). The reason for this is that one could not possibly intend to do something entirely unspecific. We can only intend to perform an action when we know in advance what type of action it will be (where a failed instance of a type still counts as an instance of this type). Hence, the things that might be done intentionally must at least have some typical and determinate shape according to which they are supposed to proceed when performed. If the final cause of a process is that according to which it proceeds, all and only such processes could

possibly be intended that have a final cause. This, however, does not imply that a process is for the sake of something only when it is intended with a further end or purpose in view. Rather, we may only purposely intend to realize processes that already have a limit.

I conclude that the final cause of a natural process is indeed its limit (and thus its essence). Now in general, a final cause is that for the sake of which something happens (τὸ οὗ ἕνεκα). If final causes are limits and thus essences of processes, it should be possible to say that a process happens for the sake of its limit or essence. And this seems to be possible. The Greek ἕνεκα is often used in phrases such as "as far as ... is concerned" or "regarding ... ." [4] In this sense, one may say of a process that it proceeds with a view to (ἕνεκα) its limit. This only means that it proceeds according to a pattern. That natural processes proceed with a view to a limit might still sound odd. It seems that all one could say is that natural processes are identified, classified, or described with a view to their limit or essence. However, as Aristotle writes in Metaphysics Δ 17, the limit of our cognition is also the limit of the cognized object (1022a9-10). If one reads "limit" as "end point" here, Aristotle would say that where our knowledge ends, the object of our knowledge must also end. This is obviously wrong, since there may be things about the object of our knowledge that we do not know. Alternatively, in line with the reading suggested here, the limit of our knowledge of a thing might be the paradigmatic form with a view to which we know this thing, and this is exactly what Aristotle says immediately before: The limit of a thing is its "for the sake of which," its substance, and its "what it was to be" (καὶ τὸ οὗ ἕνεκα, καὶ ἡ οὐσία ἡ ἑκάστου, καὶ τὸ τί ἦν εἶναι, 1022a8-9). What Aristotle means is therefore that the essences with a view to which we know and identify things are also the essences with a view to which the things are what they are. In this sense,

---

[4] For instance, Xenophon writes in Memorabilia 4,3,3 that without the light we would be like the blind, as far as our eyes are concerned. The Greek text has ἕνεκα for "as far as ... are concerned": ἕνεκά γε τῶν ἡμετέρον ὀφθαλμῶν. Plato uses ἕνεκα in a similar way, for instance in Gorgias 493E and Politeia 329B, 548D-E, and 582D.

that for the sake of which a thing exists is that with a view to which we understand it. Therefore, that with a view to which we identify processes as instances of types, their limit, is also that according to which they proceed. It is the course that processes of this type typically take, and as such, the limit of a process is the same as its essence.

Let me now address three possible objections to the claim that final causes are essences of natural processes. The first is that Aristotle often identifies the final cause with the best or the good for a certain type of thing, and there seems to be a considerable difference between the essence of a process that a certain type of natural thing undergoes on the one hand, and what is best for that type of thing on the other. Second, one may object that if the final cause of a process is its own essence, the final cause of kicking a ball will be kicking a ball, and the final cause of building a beaver dam will be building a beaver dam. This, however, would not at all be informative, and it seems that usually the final cause of something is something else. The final cause of kicking a ball should be something like scoring a goal, and the final cause of building a beaver dam should be the rise of the water level, the ultimate final cause whereof should be the survival and reproduction of a beaver. It seems that this important explanatory role of final causes gets lost in my account. Finally, it might be argued that it is very common to speak of final causes of things, for instance of an organ or artifact. It seems that the final cause of a thing is not the same as the essence of a process.

2.   The Typical and the Best

As for the first objection, Aristotle sometimes says that the final cause of a thing or process is what is best for it. I claim that the final cause of a natural process is its typical course, or what is the same, its limit and essence. These apparently conflicting statements can be reconciled if it can be shown that the essence of a process is what is best for it and for the natural things that are typically involved in it.

When Aristotle says that the final cause is something good, one might take him to mean that to have a final cause is to benefit something or someone. However, there

are a few passages in Aristotle that seem odd when read under this assumption. In Physics II 8, Aristotle supposes that things either happen as a by-product of something or by chance, or else for the sake of something (199a3-5). It is not immediately obvious why everything that does not happen as a by-product or by chance should yield some benefit. However, if the final cause of a process is its essence, it is clear that what happens by chance or as a mere by-product does not have a final cause. Chance happenings and by-products have no essence, and no standards of typicality apply to them.[5]

Further, in Physics II 8, Aristotle backs up the claim that plants develop for the sake of something by pointing out that we can think of misdeveloped plants (199b9-13). Again, the rationale for this move can only be that if there is a possible difference between a good and a bad development of a plant, then its development must occur according to a certain standard. If there is such a standard, Aristotle infers, there must also be something for the sake of which the development occurred. He may safely draw this conclusion if that for the sake of which it occurred is the standard. This is also why Aristotle can claim that accidental features of a natural being are not for the sake of anything (De Generatione Animalium V 1, 778a29–32). That a feature is accidental means that no standards of typicality apply to it. If the final cause is a standard of typicality, something that is accidental cannot have a final cause.

Moreover, Aristotle argues in Physics II 8 that although rainfall regularly leads to plant growth, this is not its final cause (cf. Furley 1985). He does not want to deny that rainfall is beneficial to plants. Rather, he denies that the growth of plants imposes any natural limit on rainfall. If the growth of plants were the final cause of rainfall, it would be typical (in the normative sense of this word) that rainfall makes plants grow.

---

[5] For a different reading see Charles (1991, p. 113). Charles claims that Aristotle argues ad hominem and that the assumption that everything that has no final cause occurs by chance is one that his opponent makes. That it must be Aristotle's own assumption is emphasized by Wieland (1992, p. 260-1).

But there is nothing atypical about rainfall that does not make plants grow. Hence, making plants grow is not its final cause.

So far, I have argued that some passages in Aristotle make more sense if one identifies the final cause of a process with its limit or essence rather than a good that it brings about. It remains to be shown in what sense what is typical is nonetheless what is best. For Aristotle does often say that the final cause of a process is "the best." In Physics II 3, he writes:

> And there are the things which stand to the rest as their end and good; for what the other things are for tends to be best and their end. It may be taken as making no difference whether we call it good or apparently good. (195a23-25, tr. Charlton)

That for the sake of which things happen tends to be the best (βέλτιστον ... ἐθέλει εἶναι). If that for the sake of which a process happens is the essence of this process, there must a sense in which the essence of a process is the best. And there is indeed such a sense. If the essence of a natural process is a fully typical instance of its type, it is rather natural to say that the essence of a process constitutes the <u>best-case scenario</u> for its type. It is what the process would be if all went well.

It may be helpful in this context to distinguish a relative sense of "good" and "best" from an absolute one (cf. Nicomachean Ethics VII 12, 1152b27). Something may be said to be good or best for something else, as when we say that a certain amount of basalt in the soil is good for certain trees, or that it is best to keep cats indoors. The soil is good relative to the goodness of trees, and if keeping cats indoors is good at all, it is good relative to the good of beings other than cats. On the other hand, something may be good in itself, as when we speak of a decent copy of a book, a good man, or a flourishing plant. This is the absolute sense of "good." The relative

sense of "good" derives from the absolute one.[6] When something is good for something else, the first helps making the second better. This second thing may again be better in a relative sense, i.e. better for some further thing; but eventually, this chain has to end. In the end, things are good in a relative sense because they contribute to the absolute goodness of something. This latter thing is often a living being (cf. Physics II 7, 198b8-9).

For a tree to be absolutely good is to be a good tree, whereas for a cat to be absolutely good is something different, namely to be a good cat. The standards by which living beings are absolutely good are the standards of typicality that apply to their life form. When we say what it is to be a good tree or a good cat, we describe paradigmatic instances of these kinds. Here, what is best is just what is typical. Hence, when Aristotle says that a natural process occurs for the sake of the best, and "best" has the absolute sense, he may well mean that it occurs with a view to the typical.

I maintain that the final cause of a natural process is its limit. In Nicomachean Ethics II 6, Aristotle writes that the bad is a form of the unlimited (ἄπειρον) and the good of the limited (πεπερασμένον, 1106b29-30). In IX 9, he adds that life is good for a man because it is definite (ὡρισμένον), and definiteness belongs to the nature of goodness (1170a20-21). This means that something is good insofar as it is definite, and conversely, something is indefinite insofar as it is bad. Also, if a definition states what is definite, it must also state what is good in this sense. One reason why this is so is that the essence of a thing provides a standard of typicality, and there are only standards of typicality, goodness, and perfection; there are no separate standards of

---

[6] As Wieland puts it, "βέλτιστον ... means the inner perfection of a thing ... . ἀγαθόν or βέλτιστον is thus only a predicate that essentially applies to every τέλος as such" (1992, p. 264). In line with Wieland, Gotthelf argues that "the very same states (or activities) which are ends will also be goods, but they will not be ends in virtue of being goods" (1988, p. 115; cf. p. 117-8).

atypicality, badness, and imperfection.[7] Even if there is more than one essence for a type, the standards of typicality and goodness always circumscribe a limited range of cases. Life and development cannot happen in any arbitrary way. On the other hand, there is no limit to the ways in which something may fail to be good. Decay and death, for instance, do not as such involve a specific sequence of steps; they can happen in any arbitrary way (Matthen, forthcoming, §5).

This answers the first of the three objections raised above. That the final cause of a process is the best for the sake of which it occurs does not mean that it is not the essence of this process. That a process proceeds according to a limit means that it is an instance of a type, such that standards of typicality apply to it. Such standards of typicality are standards of absolute goodness, so that one may say that a process occurs for the sake of the good when it proceeds according to standards of typicality.

3.   Derived Final Causes

The second objection was that in general, the final cause of a process seems to be something other than this process or its own essence. The third was that Aristotle assigns final causes not only to processes, but also to things and properties. These two objections are best dealt with at the same time.

---

[7] Aristotle sometimes speaks of entire species as deficient (Granger 1987b). In most cases, he does so by using a standard of goodness that applies to all species of a genus.

One might object that there are also standards of typicality that apply to diseases and deformations. Doctors classify such things and they are surprised if they encounter atypical malfunctions. However, although diseases are typically something bad, they are so only in a relative sense. They are bad only because they cause something else not to meet certain standards of goodness. The standards by which diseases themselves are classified as typical or atypical are not standards of badness. They are standards of goodness for diseases, according to which the typical instances qualify as good in an absolute sense, however bad they may be for what they afflict.

It is true that I have not said enough about the final causes of things and properties. So far, I have introduced and defended only what I take to be the basic and primary notion of an Aristotelian final cause. In this basic sense, the final cause of a natural process is its own essence, or a paradigmatic instance of its type. This account is incomplete. Since causal processes have complex essences, they involve parts, and the final cause of a complex process can also be taken to be a remote final cause of its parts. In this way, one may say that one process occurs for the sake of another process. Indeed, this is one of the most common uses of the term "final cause." (I maintain that in spite of being by far more common, it is not more basic than the use elucidated so far.) Further, there are also final causes of things and properties, and things and properties can be final causes. As Moravcsik notes, "there is no separate ontological category in Aristotle that would collect all of the elements that can play the role of this explanatory factor" (1974, p. 9). States, activities, organisms, artifacts, actions, and natures may be final causes (Charles 1991, p. 102-3).

I therefore introduce two ways in which one may talk about final causes in a derived sense. First, one may relate a process not to its own essence but to the final cause of another process to which it contributes or of which it typically constitutes a part. The second derived sense is one in which final causes belong not only to processes but also to things, capacities, and features. It depends on the primary sense, in which a final cause is the essence of a process. Things, capacities, and features have final causes only insofar as they may be involved in processes, and it is easy to define their final cause in terms of the essence of these processes. Let me introduce some definitions.

> Def. 1. A <u>remote final cause</u> of a natural process P shall be the essence of another natural process Q where both P and Q are undergone by the same natural thing, and P is typically involved in Q. Where it is necessary to emphasize the difference, final causes that are not remote shall be called <u>proximate</u>.

Something is a remote final cause of a natural process only if it may be described as one of its <u>typical</u> consequences. For instance, even though biting chips off a tree regularly causes wood to be left over, this need not be its remote final cause, since it may not be typical for the movements that the beaver makes. If beavers would eat all remaining wood, so that no wood is left over, they would not therefore fell trees in atypical ways (assuming that wood leftovers play no further role in tree fellings).

A remote final cause of a process may often be identified with its <u>purpose</u>. In contrast to ἐνέργειαι, movements (κινήσεις) cannot be complete while they are occurring. They do not coincide with their own essence (cf. Metaphysics Θ 6). As long as they are going on, there must be some part of their typical course yet to be taken. Therefore, one can always describe them in a way such that the part of them that is presently going on occurs for the sake of something further. This further thing may be called their purpose.

There are two different kinds of remote final cause. Some remote final causes are such that the processes that have them may also be re-described as having these remote final causes as proximate ones; for others, this is not possible. That is, sometimes when a natural thing is involved in a process A and this is a typical part of its undergoing a process B, we may say that undergoing A, in the present circumstances, is the very same as undergoing B (Anscombe 1957, §26). For instance, when a beaver is biting chips off the trunk of a tree in the course of felling it, we may say that for a beaver at this time in these circumstances, biting into the tree is the very same as felling it. There will be a whole series of processes for which this holds true. The beaver is moving its jaws and by doing so, it is biting; by biting, it is felling the tree; and by felling the tree, it is building a dam. So far, all the steps that are involved may coincide. We may point at the beaver and say that moving its jaws is, in these circumstances, the same as biting, which is the same as felling the tree. However, as Anscombe remarks concerning the case of intentional action, there occurs a break when we try to extend this series beyond a certain limit (§23). Although the beaver is felling the tree as part of the larger task of raising the water level, it is not entirely

accurate to point at it while it is felling the tree and to say that right now, it is raising the water level. It is rather only doing something that contributes to raising the water level. Although it is not correct to say that for a beaver in these circumstances, moving its jaws is the very same as raising the water level, our definition of a remote final cause still applies. We may still say that for a beaver in these circumstances, raising the water level is a remote final cause of moving its jaws, given that both biting and raising are performed by the same beaver and that one of these processes typically contributes to the other.

> Def. 2. An <u>external final cause</u> of a process P, in contrast to an <u>internal</u> one, shall be taken to be the proximate or remote final cause of a natural process Q to which P typically contributes, where Q and P are <u>not</u> undergone by the same thing.

For instance, the external final cause of the secretion of milk in mammals is that another living being of the same type may be nourished. Also, what the stomach of a living being does contributes to the growth and maintenance of this living being, and hence, the latter is one of the external final causes of what the stomach does.

A very common application of the notion of an external final cause is the attribution of a function to an organ or artifact. This is a context in which final causes are attributed to <u>things</u> rather than processes. Still, what happens here may be described in terms of the basic notion of a final cause, as the essence of a natural process. There are at least two different senses in which things may be said to have final causes. First, one may identify the final cause of a thing with the internal final cause of what it typically does. In this sense, the final cause of a beaver is what beavers typically do (biting into trees, building dams, reproducing). This kind of final cause may be called the <u>ergon</u> of a beaver. Accordingly, the ergon of a beaver's stomach is what the stomach of a beaver typically does. In some cases, what a thing typically does has an external final cause; i.e., there is a further natural process that is typically undergone by a different thing, to which the ergon of the first thing typically

contributes. In some of these cases, the ergon of the first thing is also its <u>function</u>.[8] The essence of the process to which the function of a thing contributes is not the function of this thing. It is its purpose. For example, the function of the digestive system of a beaver is the proper metabolism of the organism that contributes to the beaver's self-maintenance, and the function of a beaver's dam is to block the water, since this contributes to creating a deep-water pond and thereby also contributes to the survival and reproduction of beavers. The survival of the beaver is a purpose both of its digestive system and of the dam it builds.

Further teleological notions may be introduced along similar lines. <u>Capacities</u> may be said to have a final cause insofar as their actualizations have a basic or derived final cause. More generally, items other than processes (objects, agents, properties, states of affairs) may be said to have a final cause insofar as they are typically or intentionally associated with something that has a final cause. Moreover, there is also a derived sense in which a thing may be said to <u>be</u> the final cause of a process, namely when it essentially figures in the typical course of this process, for instance as its typical result. In this highly derived sense, a mature organism may be taken to be the final cause of its development.

I also present these definitions in order to demonstrate how one may define the usual teleological vocabulary without referring to intentions, benefits, or intelligent design. The definitions that I have proposed employ only two conceptual resources:

---

[8] In De Motu Cordis, Harvey explains what the function of the heart is: It pumps blood. He often uses <u>actio</u> and <u>functio</u> interchangeably (e.g., in ch. 14: hinc fieri actionem sive functionem cordis, quam pulsu peragit). In ch. 17, he distinguishes between the action and the function of the heart: Its action is contraction, and its function is to thrust the blood out into the arteries (tr. Whitteridge, p. 129). The action of the heart is thus the proximate internal final cause, and its function a remote but still internal final cause of what it does. Harvey uses the term "final cause" only once in De Motu Cordis (ch. 8, p. 76), in a context where it may well mean "ultimate cause." What he has in mind here is clearly a remote and external final cause.

the idea of an essence or standard of typicality that applies to processes, and the notions of involvement (of a thing in a process) and contribution (of one process to another one).[9] What is commonly called teleology is not so much a matter of attributing goals, purposes, and benefits to things. Rather, it is a matter of the mereology and analysis of the natural processes (and their essences) in which these things are involved.

Let me illustrate some of the derived senses of "final cause" by discussing a passage from Physics II 3. Aristotle writes:

> And also, then, when another thing moves, whatever things come to be between the mover and the end, such as thinning, purging, drugs, and tools, are means for health. For all these are for the sake of the end. But they differ from each other as

---

[9] Section 18 of Heidegger's Sein und Zeit may be read as a treatise about the way in which things are typically involved in processes and processes typically contribute to further processes. Heidegger distinguishes three kinds of involvement or contribution. First, in an everyday context, we encounter things as being involved in processes that they typically undergo and uses that one may put them to. This is their "Bewandtnis." Second, what things do, or what we do with them, contributes to further processes. This is what Heidegger calls the "Um-zu" of a Bewandtnis. Finally, the nexus of things that are involved in processes, and processes that are involved in further processes, can be made sense of within the context of a life form. In the context of Sein und Zeit, this is the human life form (Dasein). Heidegger also calls it the "Worumwillen" (the for-the-sake-of-which). He writes: "Das Worumwillen bedeutet ein Um-zu, dieses ein Dazu, dieses Wobei des Bewendenlassens, dieses ein Womit der Bewandtnis" (p. 87). That is, roughly: We make sense of things in terms of the processes in which they are typically involved (their Bewandtnis, ergon or function); we make sense of these processes in terms of further processes in which they are typically involved (their Um-zu, which is a remote or external final causes); and we make sense of these further processes in terms of our own life (the Worumwillen).

being operations or tools. (194b35–195a3, tr. Coughlin, modified)

As Aquinas notes, the items that Aristotle lists here may be taken to constitute a series: Tools are used for preparing a drug, the use of which causes purging, which causes thinness, which causes health (In octo libros Physicorum II 3, Marietti no. 181). The movements that the doctor makes when she prepares the drugs have their own, proximate final cause. As intentional movements, they typically proceed toward some immediate end. These movements also have remote final causes, which are purging, thinning, and curing. Preparing the drugs is complete when the drugs are prepared, but as a matter of typicality or intention, it contributes to a larger process that includes purging, thinning, and curing among its later stages.

I distinguish two kinds of remote final cause: In some cases, when an item does A in order to do B, what it does may be equally well referred to as A or B. In other cases, even though A typically contributes to B, doing A and doing B cannot be taken to be the same. Preparing the drugs is a remote final cause of the hand movements that the doctor performs in the first of these senses, since in the given context, making these movements is the same as preparing the drug. However, curing the patient is a remote final cause of preparing the drugs only in the second sense. Both preparing the drugs and curing the patient are carried out by the same agent, but in the given circumstances, preparing the drugs is not the same as curing the patient. Further, preparing the drugs, purging, thinning, and curing are external final causes of the movements that the doctor's tools undergo. They are among the purposes of the tools. Finally, the tools have a function insofar as their typical (and in this case intended) use is the preparation of drugs. Their function is what they do, insofar as it contributes to something that doctors do with them. What doctors do with them is their purpose.

Some will have noticed that in the passage quoted above, Aristotle takes the notion of a remote final cause as given and introduces proximate final causes by saying that the steps toward this end may also be said to have a final cause. In Physics II 8, he also writes that "where there is an end, the preceding steps may be taken to be

for the sake of that" (199a8-9). I proceed in the opposite direction. I start with a basic notion of a final cause, as an essence of a process, and derive the notion of an ergon, a function, and a purpose from this basic one. I take this to be one of the cases where the things that are more intelligible to us are not more basic in themselves (cf. Topics VI 4). It is easier to understand what a purpose is than what the essence of a natural process is, but in fact, essences of processes are more fundamental than purposes. Aristotle begins with what is easier to grasp; I start from what is more fundamental.

My general answer to the second and third objection is that Aristotle does indeed use the derived notions of a final cause more often than the basic one. When he does so, he uses teleological vocabulary that is less basic in itself but more familiar to us. Further, there is a simple reason why the derived notions of final causality are more familiar to us than the notion of a proximate and internal final cause. The proximate internal final cause of a natural process is, after all, its own typical course, and in most cases we will already have specified this course when we refer to this process at all. It will hardly be informative to be told that the final cause of felling a tree is felling a tree, or the final cause of biting is biting. It is much more informative to be told that for a beaver, the final cause of biting into a tree is felling it, and the final cause of felling it is building a dam.

The remote final causes of a thing or process are also often the only ones that are of explanatory value. As Aristotle says in Metaphysics Z 17, one cannot explain anything by referring to this same thing, and therefore one cannot explain a tree felling by referring to its proximate final cause. By referring to this final cause, one may explain why tree felling involves biting into trees, but this is to relate a part of tree felling to its remote final cause. As far as actual and ongoing natural processes (κινήσεις) are concerned, this is always possible. Natural processes do not coincide with their essence, and therefore they can be picked out in a way such that their own essence may be explanatorily related to them. For instance, when a beaver is actually felling a tree, it must be doing something that may be described as something other than felling a tree, which is typically involved in felling a tree. It may be biting into

the trunk right now. Therefore, one may relate what they are doing, under a more narrow description, to the essence of what they are doing, under a wider description. Further, there will be some part of biting (e.g., moving the jaws) that can be described as having biting into the trunk as its remote final cause. And so on. Moreover, natural processes may differ from their own essence by taking an atypical course. Again, they will differ from their typical course, and therefore one may meaningfully relate them to their essence.

This makes it fairly easy to answer the second objection. The objection was that final causes are not essences of processes because if they were, a process would only occur for the sake of itself. The answer is that it is indeed odd to say that a process occurs for the sake of its own essence (i.e., its proximate and internal final cause), if this process has already been picked out as having this essence. It is odd to say that beavers fell trees for the sake of felling trees. But this does not make it false. It is true that when beavers fell trees, they act in order to fell trees. Also, it may be informative to say that the aim of felling a tree is felling a tree, in case this aim fails to be achieved. Further, there are plenty of ways of picking out processes such that attributing their proximate final cause to them is less odd. A process may be picked out in an accidental way so that its essence is not yet clear. For instance, one may say that beavers produce wood leftovers for the sake of felling trees. "Producing wood leftovers" is an accidental description of what they do, the essence of which is that for the sake of which they do it: to fell trees. Hence, that it is often odd to explicitly attribute a proximate and internal final cause to a process has a good reason, but it does not imply that final causes are not essences of processes.

Something similar is true for efficient causes and causal processes. In Chapters 7 and 8, I discuss cases of immanent causation, where a process undergone by a thing causes a process undergone by the same thing (cf. Chapter 8, section 2). In the most basic sense, the efficient cause of a process is something that potentially is this process, so that what it is an efficient cause of is not a further process undergone by a different thing. Aristotle does not speak of efficient causes in this basic sense. In this

sense, they are as unfamiliar as proximate final causes, and they are so for the same reason: When we ask what the efficient cause of a certain kind of process is, we are usually not satisfied with the reply that it is "a potential instance of this kind." We were asking for a description of the cause that goes beyond the mere fact that it is the cause of some certain thing. The notions introduced in this chapter may equally well be applied to efficient causes. A <u>remote</u> efficient cause of a process is the efficient cause of a further process that typically contributes to the first. An <u>external</u> efficient cause of a process is the efficient cause of a further process undergone by a different thing, which typically contributes to the first. Put in these terms, we are usually not interested in the proximate efficient cause of a process, but in its remote and external efficient cause. Further, a <u>thing</u> may be taken to be the efficient cause of a process if it is typically involved in something that potentially is this process. A sculptor is an actual cause of a statue only if and insofar as she acts as a sculptor. In his examples, Aristotle usually refers to things that are remote and external efficient causes of processes or developments. These cases are more familiar, but less basic.

4.   Reducing Final Causes

Natural processes are processes for which there is such a thing as a typical course (according to the relevant principles of motion and rest), and complex natural processes typically take a complex course. The derived notions of final causes only apply to complex processes and the things that undergo them. It is only in this case that one may say that a process is typically involved in or typically contributes to another process. This makes it easy to see why the topic of final causality has so often given rise to the question of reducibility—the question of whether complex teleological processes may be reduced to the simple steps they involve. It is often assumed that when one does so, all appearance of final causality disappears. Now since all causal processes must admit of a complex description (as argued in Chapter 8), it is doubtful whether one can split up any of them into parts that are still causal processes but not any longer complex. But even if this should be possible, it would

not make all final causes go away. By splitting up a complex natural process into its elementary parts, one may sever the connections in terms of which the derived notions of a final cause have been introduced. Only parts of complex processes can have remote final causes, and if one considers these parts in isolation, one considers them insofar as they have no remote final causes. This, however, does not mean that they do not have proximate final causes. The reason why it seems that there are no final causes left is only that the notion of a proximate final cause is less familiar than the notions of a function or a purpose. Still, there is such a thing as a proximate internal final cause of a process. Its proximate final cause is its essence or typical course, and all causal processes must have one. A cause must be a cause of something specific, and therefore all causal processes must take some specific course. Unless the elementary parts into which a complex process is divided take an entirely unspecific and unpredictable course, they still have final causes (cf. Geach 1975, p. 93). If their course were entirely unspecific and unpredictable, however, there would be no reason to call them causal.

Hence, it is not the case that processes that have final causes must be more complex than processes that "only" have efficient causes (pace Johnson 2005, p. 288). There is no trade off between explanations of natural processes in terms of final versus efficient causes (cf. Geach 1975, p. 85). Since all causal processes are instances of types of processes, all causal processes must have an essence. This essence, however, is their proximate final cause, so that all causal processes are teleological. They all have a final cause. Teleological processes may be divided into parts that are causal processes, and although this leads to an abstraction from some of their final causes, the parts would no longer be causal processes if one would abstract from their proximate final causes. In this sense, final causes are irreducible.

It is very common in this context to encounter a misconceived criticism of Wieland's account of the Aristotelian final cause (Wieland 1992, §16; transl. in 1975). For instance, Johnson writes that according to Wieland, "teleology is for Aristotle ... simply 'a concept of reflection'—a matter of 'as if' (als ob) reasoning—a reflective

not constitutive concept of natural things" (Johnson 2005, p. 183).[10] Against this alleged position of Wieland it is then argued that for Aristotle, teleology is real and irreducible, whereas Wieland supposedly claims that it is not real but reducible to something else that is real. However, Wieland only uses the phrase "as if" in a passage where he argues that <u>chance</u> is a kind of "as if" teleology (1992, p. 258-9; cf. Physics II 8, 199b21: ὥσπερ). He argues that for Aristotle, chance cannot be an independent force, because "something is fortuitous only if it could come about also for the sake of an end" (p. 259). This is virtually the opposite of the claim that Johnson and others attribute to him: that teleology is not a real force in nature because it is only a matter of "as if" reasoning. Where Wieland writes that "with chance, an apparent, 'as if' teleology is involved" (p. 259), his critics imagine him saying that all teleology is merely apparent and "as if."

Further, Wieland claims that <u>all four</u> causes (and other Aristotelian principles) are concepts of reflection (p. 202). Therefore, that teleological concepts are concepts of reflection cannot be taken to entail that they are less constitutive for natural things than any other kind of explanatory concept. To be sure, the application of concepts of reflection does not yield a representation of a further object.[11] This means that when we relate two representations to each other as representations of the matter and form of one thing, or of the efficient and final cause of one process, this comparison does not yield a further representation, which would presumably be a representation of a causal and teleological tie between the objects of the compared representations. I have argued in previous chapters that there are no such further objects. That there is no such further object, however, does not mean that causality and teleology are not real. Both are at least as real as the fact that natural things are material substances. In particular, final causes are at least as real as efficient ones.

Moreover, although Wieland claims in one passage that all teleological explanation

---

[10] Cf. Charlton (1970, p. 120), Balme (1987, p. 280-1), Gotthelf (1987, p. 228), Furley (1996, p. 68), Hankinson (1998, p. 143), and Lennox (2001b, p. 251).

[11] Cf. Chapter 2, section 6.

must be backed up by explanation in terms of efficient causes, he might as well have said the reverse. For he writes that "the causes are ordered with respect to each other; but this ordering is reciprocal" (p. 263). Just as complex teleological processes may be reduced to less complex causal processes, complex causal processes may be reduced to less complex teleological processes. Teleological processes may be "reduced" to causal ones because causal processes <u>are</u> teleological processes.

Teleological Reasoning

Intentional actions may be classified into types on the basis of the intention with which they are done. Natural processes, in contrast, fall under types because they are governed by principles of motion and rest inherent in the natural things that are involved in these processes. The main aim of this chapter is to develop an account of teleological reasoning that does for natural processes what practical reasoning does for intentional actions. Charles claims that intentional agency and goal-directedness in nature are two fundamentally different modes of "teleological causation" (1991). I show how one of them, natural teleology, can be derived from the other, intentional agency. It may be derived by a process of generalization, so that as a consequence intentional agency turns out to be a special case. I begin with an account of practical reasoning, and then on this basis develop accounts of two kinds of reasoning that specifically apply to complex natural processes, which I call teleological and functional reasoning.[1] Once it is clear how these kinds of reasoning work and how they differ from practical reasoning, it will also be clear to what kinds of process they apply, and how these processes differ from pseudo-processes and intentional actions.

It is often been noted that practical reasoning differs from speculative (ordinary, theoretical) reasoning, but it has not always been clear wherein the difference lies.[2] Two differences have been pointed out. First, it is sometimes suggested that practical reasoning differs from speculative reasoning in that it leads to the performance of an action rather than to the truth of a proposition. Second, practical reasoning has been

---

[1] According to Mayr (1976), I should probably speak of "teleonomic reasoning." I agree with Mayr that there is a way of referring to ends that does not imply some cosmic purpose. However, I do not think that it is necessary to introduce a new term in order to make this point. The old term did not necessarily have this meaning, and if it would be restricted to this meaning, it would be of no use for us.

[2] See Corcilius (2008a) for an overview.

said to follow its own inference rules. As they stand, both assumptions are in need of qualification.

1. The Action as Conclusion

In De Motu Animalium 7 Aristotle says that in the case of practical reasoning, "the conclusion which results from the two premises is the action" (701a11–13, tr. Nussbaum). However, as every reader of Nicomachean Ethics VII 7 knows, we often fail to act according to our plans and deliberations (1150b19-21). In such cases, practical reasoning does not result in an action. Now since there is no concluding without a conclusion, if the action were the conclusion, there could be no practical reasoning without an action. This, however, only means that the conclusion that follows from a practical syllogism cannot be a particular action performance. When Aristotle says that the conclusion of a practical syllogism is an action, he may still mean an action type rather than a token performance (Clark 2001). If this is true, practical reasoning is about action types rather than particular instances of them. It relates action types to action types. This also means that practical reasoning does not differ much from speculative reasoning on this record. The conclusion of a piece of speculative reasoning is not that the thinker actually entertains some particular belief; it follows whether it happens to be drawn or not. The conclusion of a speculative syllogism is a belief type: something that should be entertained by anyone who accepts the premises. Likewise, the conclusion of a practical syllogism is not that an agent actually performs a certain action, but it is something that should, as a matter of practical logic, be done by anyone who underwrites the premises. In both cases, the conclusion is not a particular occurrence but a type of occurrence that may be instantiated by an actual affirmation or action performance. Particular actions are thus analogous to particular acts of affirming a conclusion (Nicomachean Ethics VII 3, 1147a26-29).

Corcilius (2008b) argues that practical deliberation differs categorically from the practical syllogism that Aristotle describes in De Motu Animalium. According to

Corcilius, deliberation is propositional, it is not about particulars, and it does not result in a particular action but in a proposition; whereas the De Motu practical syllogism is not propositional, is about particulars, and leads to a particular action. What I have said implies that the difference between the two is not as big as Corcilius thinks. Practical reasoning in general leads from action types to action types, and such types are best represented by propositions. In this sense, all practical reasoning is propositional. This does not mean that in order to relate actions to further actions, or proximate to remote final causes, one needs to conceptualize their relations. One may just do one thing in order to do another thing. In De Motu Animalium, Aristotle says about speculative reasoning that putting together the premises is thinking the conclusion (701a10-11). Connecting A, B and C is connecting A and C, whether one is aware of this or not. This is not only true when A, B, and C are speculative propositions but also when they are actions or action types. Doing A in order that B and B in order that C is doing A in order that C.

That practical reasoning relates action types to action types leads to a corollary concerning the initial premise of the practical syllogism, which is often taken to be a statement to the effect that something is obligatory, good, or desired. Like the conclusion, this premise should also be thought of as a characterization of an action type—something to be done—rather than the statement that it is to be done (cf. Anscombe 1957, §35). Practical reasoning does not lead from imperatives or desires to actions but from things to be done to further things to be done (Kenny 1979, ch. 11). That something is good, desirable, or obligatory is not something to be done.

In Intention, Anscombe discusses two apparently different ways of explaining actions. First, in order to explain an action, one may offer a description of the same action in terms of a more distant aim (its remote final cause). Asked why a man is pumping water, one may answer that he is replenishing the water supply, thus giving another description of what the man is doing. Second, one may explain an action by practical reasoning. She concludes that both are ultimately the same:

The interest of the account [that Aristotle gives of practical reasoning] is that it describes an order which is there whenever actions are done with intentions; the same order as I arrived at in discussing what 'the intentional action' was, when the man was pumping water. (1957, §42)

This is another way of saying that in practical reasoning, we relate action types to further action types, or what is the same, we relate aims and ends of an agent to further aims and ends of the same agent. That an action is of a type means that it has an essence. Since the essence of an action, like the essence of a natural process, is its final cause, relating action types to action types is the same as relating final causes to final causes. Practical reasoning is thus concerned with the mereology of action types: Action types are shown to contribute to, or to involve, other action types.

In any case, practical reasoning is not distinctive in that its conclusion is a particular action performance. The conclusion may be taken to be a particular action in the same sense in which the conclusion of a speculative syllogism may be taken to be a particular act of affirming a proposition. More properly, however, the conclusion of a practical syllogism is taken to be an action type.

2. Inference Rules

Now for the second allegedly distinctive feature of practical reasoning: that it follows its own inference rules. In order to see whether this is true, we should consider a particular example, such as the following practical syllogism from Posterior Analytics II 11:

Why does he walk about after dinner? Because walking prevents the foodstuffs from remaining on the surface at the mouth of the stomach and this is healthy. (Cf. 94b11-26)

This example is taken from a text that is generally considered difficult. Aristotle begins II 11 with the statement that all four kinds of causes, and hence also the final cause, can serve as a middle term in a syllogism. In general, the middle term is the term that appears in both premises of a simple, categorical syllogism and links them. Now an attempt to make the final cause (health) the middle term of a syllogism would lead to either of the following:

(a) Digestion contributes to health, walking leads to health, hence walking contributes to digestion.

(b) Digestion contributes to health, health requires walking, hence digestion involves walking.

(c) Health involves digestion, walking leads to health, hence walking leads to digestion.

However, neither (a) nor (b) are valid practical inferences. For instance, (a) breaking an egg may contribute to making an omelet, and heating a pan may lead to making an omelet, but heating the pan does not therefore contribute to breaking the egg. Likewise, (b) breaking an egg may contribute to making an omelet, and making an omelet may require heating a frying pan, but breaking the egg does not therefore involve heating the pan. Instance (c) is a valid form of practical reasoning, but it does not seem to be what Aristotle wants. It leads to the wrong conclusion. The syllogism that he presents is not supposed to establish the fact that walking leads to digestion but rather that because walking is already known to facilitate digestion, and digestion contributes to health, walking contributes to health. It seems that contrary to Aristotle's initial announcement, the middle term can only be digestion and not health. If this is so, the example has the following general form:

(*)   Doing A leads to B

B contributes to G

Therefore, doing A contributes to G

Syllogism (*), however, is a perfectly valid speculative one. It does not follow its own special inference rules (Natali 2001, p. 83; Müller 1982, p. 74). On the other hand, (*) still does not answer the question with which Aristotle started. His question was: "Why does this man walk about?" The answer to this question is actually provided by the following complex syllogism of which (*) constitutes only a part:

(1)   For a goal G (of this man),

(*)   Doing A leads to B

B contributes to some G

Therefore, doing A contributes to some G

To infer A as a further goal (of this man)

This syllogism starts with the specification of a goal G of an agent, e.g., health. On the basis of a speculative syllogism from which it follows that A, walking, contributes to this goal, we infer A as a possible derived goal of the same agent. The so-called practical syllogism thus turns out to be an ordinary instance of speculative reasoning, framed by specifications of things to be done. The embedded speculative reasoning (*) is often referred to as "the" minor premise, but as Kenny points out, there are often several premises that lead from the specification of an end to something that can be brought about (1979, p. 139-42). For instance, as Aristotle writes in Metaphysics Z 7, a doctor may reason that a certain treatment causes warmth, that warmth causes some further state that contributes to equability, and that equability belongs to health. In any case, the real calculating is done exclusively by the speculative part of a practical syllogism, in accordance with the rules of speculative reasoning. There are no special practical inference rules; there is only a special use of ordinary inference

rules (Allan 1955, p. 328; Kenny 1975, p. 87).

In Posterior Analytics II 11, Aristotle writes that it becomes clearer how the final cause can be a middle term once one "exchanges the accounts" (δεῖ δὲ μεταλαμβάνειν τοὺς λόγους, 94b21-22). This remark has been notoriously unclear to the commentators. The account of practical reasoning proposed here might yield a reading of the passage that has not been considered so far: Aristotle could be suggesting that we regard the entire minor premise of a practical syllogism as an embedded piece of hypothetical reasoning. "Account" (λόγος) may well mean 'reasoning' (cf. Posterior Analytics ed. Ross, p. 644), and in Prior Analytics I 23, μεταλαμβάνειν also refers to the temporary substitution of a hypothesis for the claim that was to be established in the first place (41a39, cf. ed. Ross, p. 371). When Aristotle speaks of "exchanging the accounts," he may thus ask us to substitute bits of reasoning for a premise. In practical reasoning, we would be supposed to substitute a chain of reasoning, i.e. (*), for the minor premise. In this sense, practical reasoning would proceed κατὰ μετάληψιν (Prior Analytics I 29, 45b17). It would proceed by replacing the minor premise by a speculative syllogism that shows how its terms are linked. We replace the minor premise, that walking contributes to health, by a speculative syllogism that shows how walking and health are linked. This would also make it easier to understand how the final cause can be a middle term. For although health is not a middle term in (*), it does appear twice in (1): Given health as a goal, doing A contributes to health, from which one may infer A as a derived goal. And in general, a middle term is simply a term that appears twice in the premises of a syllogism but not in the conclusion. Just as in speculative reasoning, the middle term (health) appears in both premises but not in the conclusion.

I conclude that practical reasoning is an application of speculative reasoning in a practical context. I leave open the particular nature of the embedded speculative reasoning and the rules that apply to it. So far, any kind of reasoning will do, as long as it shows that A causes, is involved in, leads to, or contributes to G. In most cases, the reasoning will not be an instance of logical deduction, since usually such a

deduction does not show how to achieve a practical end. One might therefore demand that more specifically there must be a causal chain leading from A to G (cf. Mackie 1974, p. 273; Müller 1982, p. 107). But it is also probable that the notion of a causal chain is best explicated on the basis of practical (or teleological) reasoning, such that an efficient cause would be defined, roughly, as whatever may be brought about or manipulated in order to bring about some effect in certain circumstances (Gasking 1996, ch. 7; Woodward 2003, ch. 2). As long as an account of causality in terms of possible manipulative strategies seems promising, it would be imprudent to already presuppose a specific account of causation in an account of practical (and teleological) reasoning.

3. Discussion

Aristotle also writes that the starting point of practical reasoning is the endpoint of the reasoned action, and the starting point of the action is the conclusion of the reasoning (Z 7, 1032b15-26; cf. Eudemian Ethics 1248a20-29). When there is a suitable chain of speculative reasoning that leads from A to B, a chain of practical reasoning may lead from B to A. In (*), we infer G from A, and in (1) we infer A from G. In this sense, practical reasoning may be said to be the <u>mirror image</u> of speculative reasoning (Kenny 1966, p. 73).

This claim, that practical reasoning is a mirror image of speculative reasoning, might be misunderstood. One might think that in order to obtain a valid practical syllogism, one may simply invert a given chain of speculative reasoning. However, consider the following example of a practical syllogism given by Anscombe (1957, §33):

(2)   Vitamin X is good for all men over 60

Pig's tripes are full of vitamin X

I'm a man over 60

Here's some pig's tripes

Me eating what's here (action type)

If this is a mirrored speculative syllogism, one should be able to mirror it back and obtain the original speculative syllogism. This, however, seems to yield the following rather useless sequence of sentences (cf. Geach 1975, p. 91).

(2')   Me eating what's here (action type)

Here's some pig's tripes

I'm a man over 60

Pig's tripes are full of vitamin X

Vitamin X is good for all men over 60

What is wrong here is that the entire practical syllogism is taken to mirror a speculative syllogism. However, the trick is not to invert the complete sequence of sentences in a speculative syllogism but rather to insert the relevant piece of speculative reasoning into the pattern given in (1) above:

(2'')   For the goal of getting vitamin X,

　　　　(*)   Pig's tripes contain vitamin X

　　　　　　　Here's some pig's tripes

　　　　　　　Therefore, eating what's here gives one vitamin X

　　　　To infer eating what's here as a derived goal

Practical inference mirrors speculative reasoning only on a large scale. The incorporated reasoning is not mirrored in its details, but its starting point is taken as the conclusion of the overall practical deliberation and vice versa.

Von Wright gives an account of practical reasoning that differs considerably from the one presented here (e.g. 1983, ch. 1 & 2). His basic rule for practical inferences is that whoever wants an end must also want the necessary means. Supposing that A is a necessary means for attaining G, we may have the following situation:

(3)   I need to attain G

　　　　G necessitates A

　　　　Therefore, I need to accept that A

This instance of allegedly practical reasoning obeys the rules of ordinary logic and already works in the forward gear. It is not a mirror image of a speculative syllogism, not even on a large scale. However, I do not think that von Wright's favorite example is a case of practical deliberation at all. That G necessitates A does not constitute an instrumental or teleological connection between G and A; that is, it does not imply that one may choose A in order that G or vice versa. It is entirely open what steps shall be taken in order to achieve G, and this is not what reasoning (3) is about. Reasoning (3) is about what is necessary, and practical deliberation is not about what is necessary. It is about what is suitable. It does not relate ends to their necessary

consequences, but means to ends. This is not to criticize von Wright, who does point out that he is not concerned with proper Aristotelian practical inference (1983, p. 1).[3]

A further worry should be addressed before moving on. The following is true of speculative reasoning: If A implies G, then A & B also implies G. A syllogism that shows how A and G are linked may thus be replaced by a syllogism that shows how A & B and G are linked. Now as I have argued, given a task G, and given that A implies G, one may infer A as a derived task. However, that one may do A in order to achieve G does not imply that doing A & B would also be an appropriate means for achieving G. In cases where a syllogism that shows how A and G are linked is an appropriate "minor premise" of a practical syllogism, the other syllogism that shows how A & B and G are linked may be inappropriate. For instance, that one may break an egg in order to make an omelet does not imply that one may break an egg and the kitchen window in order to make an omelet. Note, however, that this kind of inference would not even be admissible for some kinds of speculative reasoning, especially not for causal reasoning (cf. Geach 1975, p. 93). If turning on the heater causes the room to warm up, turning on the heater and opening the window need not. Hence, turning on the heater and opening the window is not an appropriate means of achieving certain goals that may be achieved by only turning on the heater. In such cases, A & B does not imply what A implies because there is a conflict between A and B. On the other hand, where A and B do not conflict, it remains true that doing A & B is also a way of doing A. In these cases, doing A & B where A alone suffices may still be inappropriate because doing B is unnecessary. For instance, I may burn down my house in order to roast a pig (Anscombe 1995, p. 20). However, the reason why this seems inappropriate is only that burning down my house is likely to flout other objectives, such as having a shelter or saving money. Therefore, we have to assume that practical reasoning is evaluated against a backdrop of all relevant goals, goods, and ends of an agent, including efficiency. Indeed, if there is in fact nothing that

_____

[3] It might be to criticize Nussbaum, who relies on von Wright in her explanation of Aristotelian practical reasoning (1985, p. 176).

speaks against burning down my house, it might be quite efficient to roast a pig that way (cf. Kenny 1975, p. 91).

This is to admit that often there are several alternative ways of achieving a practical aim, and one may have to choose among them according to some rules and principles. The most general rule is probably that the means should be reasonably efficient as compared to the available and known alternatives (cf. Nicomachean Ethics III 3, 1112b16-17). However, I mention the necessity of introducing some such efficiency clause in practical reasoning only to later put it aside. I do not care about the exact formulation of it here, since it is not be needed in teleological (and functional) reasoning. The efficiency clause is necessary because agents may often deliberately choose among several alternative means for achieving an end. Nature has no such deliberate choice.

Let me propose the following canonical form of practical reasoning:

(4)   For a set G of tasks of an agent,

           Given that as a result of speculative reasoning,

                achieving A contributes to some B ∈ G

           modulo some efficiency condition

           To include A in the set of tasks G

On the basis of a valid speculative inference, plus an efficiency condition, we may add further tasks of an agent to a set of already known tasks of this agent (cf. Geach 1975, p. 89). By tasks I mean things to be done, which are action types. When I speak of a set of tasks, I do not mean an extensionally defined set that could not gain or lose members. Rather, G should be taken to be something like an agenda, a collection of objectives, or a story about what an agent wants in her life, which may be continued, expanded, and revised. Often, an agent may choose a means A for accomplishing one of her tasks without being directly able to achieve A. In such cases, A will simply be

added to the set of tasks G and the search for appropriate further means will go on. By subsequently adding further actions to be done to the set G, one may establish a rather complex agenda. Once an item in this agenda has been achieved, it may be deleted from the list of things to be done. For instance, I may set out to open a door and infer that turning a key would be a way of doing so, only in order to proceed to the question of how to get hold of a key that would do the job. Once I have the key, I can open the door and thus delete all steps involved in doing so from my agenda. The further tasks that we may infer from the tasks of an agent may conflict with other tasks in her agenda or with tasks that may be inferred from them. Hence, that a task B may be accomplished by doing A need not mean that anyone who wants to attain B also should want to attain A. But whether this is so or not depends on what other tasks the agent already has, and what other tasks may be inferred from them by way of practical reasoning. That is, the means for discovering such conflicts is again practical reasoning.

According to the present account of practical inference, the specifications of goals in G may always be taken to be descriptions of actions to be performed. Not all possible goals of an agent are of this kind. For instance, the goal of what a doctor does may be the health of a patient. Applying a massage may be a means of attaining this goal and hence, in given circumstances, applying a massage is the same as curing the patient. But neither of them, massaging and curing, is the same as the health of the patient. Health is not a possible action of the doctor at all but only the aim of her actions. The account of practical reasoning that I give here is limited in that it restricts goals to actions. In the present context, this is preferable for the following reason. As long as a goal state is not actually achieved, it does not obtain. The health of the patient does not exist until the doctor has cured the patient. For this reason, it seems that when we explain what the doctor is doing by relating it to the health of the patient, we must refer to the contents of her mind, or of the mind of an observer. In contrast, the action of curing the patient does occur before the doctor has completed it. There is something outside the mind of the agent or observer that we can point at and say, "This is the curing of the patient." Since my aim is to generalize practical

reasoning in order to apply it to animals and plants that cannot express any intentions and feelings, we had better confine ourselves to entities that we can point at (cf. Kenny 1963, p. 122-3). We can point at a beaver and say, "Look, it is felling a tree," even before the tree felling is complete. We do not need to refer to a future state of affairs or to a conception the beaver is assumed to have of this state. The felling of the tree is presently going on, and still it may be the goal of what the beaver is doing.

4. Natural Teleology

After his brief discussion of practical reasoning in Metaphysics Z 7, 1032b6–10, Aristotle remarks that "things which are formed by nature are in a similar situation" (Z 9, 1034a33). By this he does not mean that we describe nature in anthropomorphic terms but conversely, that insofar as we act for reasons, we imitate nature (Physics II 8, 199a8–b32). That actions are similar to natural processes means that we should think of them in similar ways. In Parts of Animals I 1, Aristotle writes that in general, the mode of demonstration in the natural sciences differs from the one in the theoretical sciences. The theoretical sciences explain something by way of speculative reasoning, and their starting point is something that is. The natural sciences, in contrast, start from something that is going to be and infer what needs to be there for this to happen (640a3-4). This kind of reasoning is very similar to practical reasoning: Given a certain end, suitable means are inferred.

What I have achieved so far is to give an account of practical reasoning that does not rely on an explicit reference to the beliefs and desires of an agent. In practical reasoning as described above, we do not explain actions by citing beliefs and desires but by relating them to further actions and tasks of the same agent. This account has the considerable advantage that it may be fairly directly transferred to other kinds of teleological processes. I have in mind explanations of the following kind:

> For animals, in order to maintain their proper working state, an appropriate means is to sleep. Therefore, animals sleep. (Cf. De Somno 2, 455b13-26)

Animals do in fact sleep, and that they do so contributes efficiently to their maintaining their proper working state. Since the latter is one of the goals that animals may be assumed to have, we may infer that they sleep in order to maintain their proper working state.

Such instances of teleological reasoning do not involve the assumption that particular animals ever decide to sleep. In this respect, teleological reasoning differs from practical reasoning. Practical reasoning is typically reflected in the actions it is about. Human agents do not just act in order to achieve something; they also conceive of their actions as directed at this aim. Therefore, we do not only explain the behavior of humans by relating it to further human behavior, but we also naturally assume that humans are capable of describing their own actions in ways that reflect their purposiveness. To act intentionally is to bring one's own actions under a certain concept such that they may be accounted for by practical reasoning, and only beings that are at least in principle capable of practical deliberation may be said to act intentionally.[4] Even where agents act without deliberating, they have a certain understanding of what they are doing and how it may be described in terms of its aims. Without this understanding, they could not perform intentional actions. In this sense, practical knowledge is "the cause of what it understands" (Anscombe 1957, §48). The practical knowledge of the agent is the basis on which we may account for her actions by way of practical reasoning.

In contrast, the behavior of non-human animals and plants admits of teleological explanation not because the animals and plants in question would in principle be able to give such an explanation. Teleological explanation is not deliberation (Müller 1982, p. 120). Teleological reasoning may serve to explain what animals and plants are doing, but animals and plants do not choose what to do on the basis of teleological reasoning (cf. Charles 1991, p. 107-8). Their behavior is not caused by anything comparable to practical knowledge. Accordingly, the explanation of why animals

---

[4] Michael Thompson, in conversation.

sleep is not intended to represent anything that is actually involved in making it so. The conclusion of a teleological syllogism is a fact that we already know by observation. Animals do in fact sleep, the reasoning goes, and this is why they do it. (Or, equivalently: Animals do in fact maintain their proper working state, and this is how they do it.) In teleological reasoning, we do not infer things to be done from other things to be done, but we relate observed behavior to further observed behavior. We point out how some observed trait of a living being fits into a general story about this type of living being.

For this reason, no efficiency clause needs to be supplied in teleological reasoning. We need not assume that nature should always find the most efficient way of doing things. As things are, nature does find quite efficient ways of achieving aims, but there is no need for a rule that prevents inefficiency. The task of teleological reasoning is not to produce teleological behavior but only to account for it. Even Aristotle's famous principle that nature does nothing in vain only serves to explain why things are as they are. As Aristotle says, "if there is a better way, this is in accordance with nature" (Progression of Animals, 704b17-18). Since teleological reasoning does not cause what follows from it, the principle that nature does nothing in vain is not needed to prevent arbitrary possible scenarios from being realized.

The most important difference between practical and teleological reasoning is probably that the former applies to particular agents and their specific agenda, whereas the latter primarily applies to types of living beings (Fitzpatrick 2000, p. 154-8; Thompson 2008, p. 78). To be sure, practical reasoning is also about types of agents and types of situations. As Kenny remarks, the initial premise in a practical syllogism states "that a certain <u>kind</u> of action is good for certain <u>kind</u> of agent" (1979, p. 139). Practical reasoning starts from "a general knowledge of the facts of life" (Natali 2001, p. 75). Teleological reasoning, however, is about types in a stronger sense. The reason for this is that beings that do not deliberate cannot express their preferences other than by acting on them. Deliberating agents can form and express their intentions without already acting on them, and they may thus pursue a

highly individual and idiosyncratic agenda. It is in many cases not easy, and sometimes impossible, to directly see what an agent intends by doing a certain thing. In such cases, one may often ask the agent what she aims at. In contrast, beings that do not deliberate have only one way of expressing their intentions, if they have any: namely acting on them (cf. Schofield 1991, p. 38-9). Therefore, there is also only one way of determining what they are after: to look what they are actually aiming at. All one can do in order to find out what animals or plants aim at is to describe what they are doing in more general terms. The very fact that animals and plants do not deliberate what to do means that they cannot choose to put an item or an action to a non-standard use. If their actions are goal-directed at all, then there must be some general fact about their life form that shows this. This is the reason why teleological accounts of animal and plant behavior are usually in the habitual present tense (Woodfield 1976, p. 124).[5]

Despite all these differences between practical and teleological reasoning, there is one striking formal similarity. In both kinds of reasoning, the conclusion may be added to the set of items with which the reasoning starts. In practical reasoning, we infer particular tasks from other particular tasks. In teleological reasoning, we infer what things of a type typically do from other things that they typically do.

Teleological reasoning starts with some knowledge about what typical instances of a life form T are known to do. By going through an embedded chain of speculative reasoning, we may situate further observed facts within this more general story. If there is a reasonable account according to which an observed trait contributes to something that typical Ts are known to do, the observed trait may be included among the things that Ts typically do. (As everywhere else in this book, "typical" has a normative sense here. Typical is what is in accordance with standards of goodness.) Since all that is at stake is the relation of observed behavior to already assumed or

---

[5] Allen (unpublished) notes that "for the sake of" refers to typical aims and purposes when followed by the generic present tense, and to intended aims or purposes when followed by the progressive.

known facts, the efficiency clause may be dropped. Accordingly, the canonical teleological syllogism looks like this:

(5)   For a set G of things that a typical T does,

           Given that some particular T is known to do A, and that

           as a result of speculative reasoning,

                    a typical T's doing A contributes to some $B \in G$,

           ----

           To include in G that a typical T does A

For instance, suppose that we see a cat purr and that we know that the frequency of the sound it makes stimulates bone growth and healing. Bone growth and healing certainly contribute to something else cats may be said to typically do, such as hunting, eating, and copulating. On this basis, we may reason that for a cat to purr has a point, and that the purring is typical and not abnormal. As in the case of practical reasoning, this is only a reasonable hypothesis to start with. The next step is to look at how the claim that cats typically purr relates to other claims about what cats typically do (just as we must ask how the task of an agent relates to the rest of her agenda). If we should find out that most healthy and normal cats do not purr or that purring also has important detrimental effects, we may revise the claim that purring is typical for cats. On the other hand, if we realize that purring also is a way of communicating, we may confirm it.

Consider a second example. We know that the leaves of maple trees turn red before they fall, and we may ask whether there is a teleological explanation for this. As it turns out, maple trees use certain chemicals in order to prevent other plants from growing in their immediate vicinity, and these chemicals are brought to the ground by the falling leaves. Since these chemicals cause the leaves to be red, we are now in possession of both an explanation for their color and of a more accurate description of what is happening. It is not exactly typical for maple trees to have red leaves. What is

typical for them, because it fits into a general account of what maple trees aim at, is to produce the chemicals that make them red. A tree without red leaves would not be atypical because of the color but only because of the lack of these chemicals. The color is merely a by-product of what is typical for maple tree. As a mere by-product, it does not have a final cause (cf. Physics II 8, 199a3-5).

One might object that such reasoning may easily justify the inclusion of items in the natural history of a plant or animal that do not seem to belong there. For instance, we know that cats typically die after a certain time. Now suppose that we encounter a particular cat with a lethal disease. Can we conclude that since this disease contributes to what cats typically do, cats typically have this type of disease? Again, it is important to ask whether the disease in question has detrimental effects that other causes of death do not have. For instance, the disease might be contagious such that it also afflicts younger animals and causes premature death. This would count against including the disease in the set G after all. In contrast, a syndrome that does not cause anything other than death at an appropriate age would probably not even count as a disease but rather as a natural cause of death.

5. Functions

It is only a small step from teleological to functional reasoning. Here is, again, an example provided by Aristotle.

> Birds that eat flesh must be able to fly. Large wings contribute to this ability, and therefore the wings of birds that eat flesh are large. (Parts of Animals IV 12, 693b26–694a3)

We start from the plausible assumption that predatory birds typically are good at flying. The embedded speculative reasoning, however, is not concerned with birds but only with one of their functional parts. Since large wings contribute to the ability to fly, the reasoning goes, they also contribute to what predatory birds typically do. This

is a syllogism of the following general form.

(6)   For a set G of things that a typical T does,

Given that some particular F is known to do A, and that

as a result of speculative reasoning,

a typical F's doing A contributes to some B ∈ G,

To include in G that a typical F does A

Again, we start with some general knowledge about Ts and the observation of a particular trait, behavior, or property of an F. The conclusion may then be that this particular trait is not merely particular but that it has a more general point. The only difference between teleological and functional reasoning is that function executions do not contribute to the tasks of the function bearer F itself but to the tasks of some T ≠ F. A function of a thing is something it typically does, which contributes to something that another thing typically does. Syllogism (6) is thus not about fitting observed traits of an F into a story about Fs but about fitting traits of an F into the story about Ts. Put in terms introduced in Chapter 9, section 3, functional reasoning assigns external final causes to things.

In a number of cases, the function bearer F is a part of the living being in question, as with our bodily organs and tissues. Sometimes, however, F is not properly called a part of T even when it is located inside the organism, like certain enzymes and vitamins, or the stones that certain animals swallow in order to facilitate digestion. It is sometimes claimed that items have functions only with regard to a containing system (Cummins 1975, p. 763), but it is difficult to give a non-circular account of the relevant sense of "system" and "containing." The stones in the stomach of a duck are not proper parts of its organism but at best functional parts of it, and the functional parts of a system are best defined as items that execute a function within that system. The account on offer avoids this circularity by specifying a more general relation

between a function bearer and the life form in which it figures. Prima facie, if there is a valid instance of functional reasoning according to which what an item typically does contributes to what a type of living being typically does, then this item has a function. The function bearer need not be part of any containing system other than the system of process types tied together by functional and teleological reasoning. An item that has a function in the life of a certain type of living being may even be located strictly outside its organism. This is the case with the breasts of mammals that have a function in the lives of their offspring, but also with spider webs, beaver dams, and the empty seashells that serve as a shelter for hermit crabs.

Thus, an item has a biological function insofar as what it typically does may be accounted for by functional reasoning. This account of what it is for an item to have a biological function has several advantages compared to other traditional and contemporary ones. First, since it is phrased in terms of what an item typically does, it does not amount to the claim that every beneficial effect of an item would be its function. The primary aim of Wright's etiological account of functions is to distinguish between the function of an item and the accidental beneficial effects of its presence (1973, p. 141-2). Some things are beneficial without having a function; some have a function without being beneficial. Wright argues that in order to see the difference, one must acknowledge that functions are explanatory, in the sense that they explain why a thing is present or performs a certain task. This leads him to the following analysis: "The function of X is Z means: ... (a) X is there because it does Z & (b) Z is a consequence (or result) of X's being there" (p. 161). In cases where an item is designed to do Z, this is more or less obvious. Some items may be designed to do Z but still never do Z, but in most cases, the fact that they do Z explains why they are there. For items that were not designed to do anything, the explanation is most likely an evolutionary one. That items of a certain type did Z in the past explains why they are still around.

However, accounts such as Wright's imply that an item can only have a function if it has the appropriate kind of history. It must have been designed to do Z, must have

done Z in the past, or must be a descendent of things that have done Z in the past. This leads to the claim that if there were instant organisms ("hopeful monsters"), which would not have the right kind of history, their parts could not have functions (cf. McLaughlin 2001, p. 88). Some think that this is counterintuitive, and the account that I propose does not entail it. In particular, it does not involve the assumption that evolution has taken place.

There is a difference between an item's having a certain function and the reason why it came to have this function. One may explain why the parts of an organism have their function by referring to the evolution of this organism. This very possibility, that one can explain one by referring to the other, is threatened when one is defined in terms of the other. If one defines functions as effects of an item's presence that may be explained by reference to evolutionary theory (as in Neander 1991), the claim that evolutionary theory can explain why the parts of organisms have a function becomes vacuous. This is not to deny that evolutionary explanations for the presence of biological functions are correct and useful. Also, the notions of a function and a selected effect may well be extensionally equivalent such that for certain purposes one can be used as well as the other. But they should not be identified by definition. As Woodfield writes, "to think that evolutionary theory can explicate the meaning of 'function' is to get things back to front" (1976, p. 118).

Moreover, as Grim (1974) points out, Wright's account still does not reliably distinguish between function and accidental benefit. The general reason for this is that it is still formulated in terms of particular items and circumstances. A particular item may end up somewhere as a consequence of its doing Z, such that it keeps doing Z because of this, where Z is still not its function. Whether an item has a function or not can often not be seen by looking at particular instances and circumstances. An item has a function by being an instance of a type of items, instances of which generally have this function. Therefore, the notion of a function should be explained on the level of types of items and types of circumstances. The account suggested in this

chapter does this; most etiological accounts do not.[6]

The present account is also preferable to other alternatives to etiological accounts. Since it is formulated in terms of what is typical, it does not define the function of an item in terms of its actual disposition, tendency, or propensity to bring about a certain effect (Bigelow and Pargetter 1987). This would be wrong because a broken item may not in any sense tend to or be disposed to do Z, although doing Z may still be its function. It may have this function without being the least able to execute it. For although it is not at all disposed to do Z, it is still of a type whose instances typically do Z. Also, the account presented here does not relativize functions to possibly arbitrary human goals, such as the otherwise similar account given by Cummins (1975). Rather, functions are explained by relating them to objective accounts of what certain living beings, human or non-human, typically do in their lives. It explains how items may have functions in the absence of conscious beings who impose these functions on them.

6. Conclusion

I have provided accounts of what it means for a natural process to be teleological, and for a natural thing to have a function. A process is teleological insofar as it may figure as the conclusion of a teleological syllogism, and in teleological reasoning we relate (types of) natural processes to (types of) other natural processes. Two special forms of teleological reasoning have been discussed: practical and functional reasoning. An event is an intentional action insofar as it may be properly thought to result from practical reasoning. By analogy, a process is teleological insofar as it may be related, by teleological reasoning, to other processes typically undergone by instances of the same type. Finally, an item has a function insofar as what it typically

---

[6] Etiological accounts take two steps at once. They acknowledge that functions should be explained on the level of types but immediately go on to reduce this level to something particular, e.g., by referring to the particular history of a particular item. This second step is not necessary, and it causes more problems than it solves.

does stands in a certain relation to what instances of a certain life form typically do. Actions, function executions, and other teleological processes are thus teleological by bearing certain relations to other processes, and the nature of these relations is rendered explicit in the respective kinds of reasoning.

Teleological reasoning relates processes to other processes to which they contribute, and this means, in the terminology introduced in the previous chapter, that it relates the proximate and internal final cause of a process to its remote and possibly external final causes. In general, teleological reasoning traces the relations between processes or things and their remote and external final causes. If there are remote and external efficient causes, there should also be such a thing as (efficiently) causal reasoning. It would be the mirror image of teleological reasoning. In a similar way, one should be able to distinguish between proximate and remote <u>formal</u> causes and thus be able to relate the paradigmatic form of a thing to the paradigmatic forms of its parts. For instance, lungs are typically parts of organisms, and therefore one might refer to the paradigmatic form of a living being as a remote formal cause of its lungs. If this can be done, there should also be something like formal reasoning, and its complement, material reasoning. These four kinds of reasoning would correspond to the four causes. Teleological reasoning would mirror causal reasoning, and formal reasoning would presumably mirror material reasoning in a similar way. Further, teleological and formal reasoning would be closely linked, at least for living beings. Seen in this light, functional reasoning turns out to be a hybrid form of reasoning. It relates the final cause of what an organism does to the final cause of what one of its functional parts does, and thereby also the form of the organism to the form of its part. It traces two links at once: the links between the form of an item to the form of its parts, and the links of what the item does to what its parts do. I have here only given a detailed account of how teleological, practical, and functional reasoning work. There might be several ways in which the other kinds of reasoning differ. To describe them would be the topic of more chapters, if not another book, which I have not written.

Conclusion

What follows is a rough summary of some of the more basic claims made in this book. It goes without saying that they cannot all be adequately expressed and defended within a few pages; they have already been explained and argued for in the rest of this book. Also, I have sometimes distinguished between what Aristotle literally says and what I take to be a possible and helpful position that he does not explicitly take, which nonetheless matches and explains other things he says. In this conclusion, such distinctions are blurred. It should therefore not be taken to represent Aristotle's stated views.

Throughout this book, I have been concerned with natural things and processes. Natural things are things with an inherent principle of motion and rest, and natural processes are processes that are governed by such principles. Principles of motion and rest may be thought of as standards of typicality that apply to the processes that are governed by these principles. That they are inherent to a thing means that such standards of typicality follow from a proper account of the thing's nature. Both natural processes and things are instances of types insofar as they are subject to standards of typicality, not insofar as they meet them. Natural things are such that one may assess on the basis of their own nature whether the processes in which they are involved are typical and natural for them. A natural thing is a thing for which there are intrinsic standards of typicality. The standards of typicality that apply to artifacts, in contrast, do not arise from an account of their nature but from an account of the nature of the living beings that produce and use them.

Aristotle's four causes are primarily causes of natural things and processes. They correspond to four questions that one must raise about them in order to treat them as the natural things and processes they are. Natural things are capable of change. In order to study them, one must therefore study the processes in which they are typically involved. Further, the ways in which natural things change and develop are

governed by standards of success and typicality. Natural things may fail to change and develop according to these standards, and the natural processes they undergo may take atypical courses or remain incomplete. For each natural thing, one must therefore distinguish between that which potentially meets the respective standards and what it would be if it satisfied them. The first is something that potentially is a thing of a certain kind, out of which such a thing may come to be. The second is what the first potentially is. The same distinction between a potential and its possible actualization must be observed concerning natural processes. All in all, four questions must be raised, none of which is a Why-question:

	out of what?	what?
concerning natural things:	material cause	formal cause
concerning natural processes:	efficient cause	final cause

The Material Cause

   The material cause of a natural thing is something that potentially is this thing. This does not mean that everything that potentially is a thing is its material cause. One must in any case begin with a given material substance, a τόδε τι, and may then ask which thing it is, in this case, whose potential is being partly or fully realized by the substance. To describe something as the material cause of a natural thing is to describe it as potentially this thing. In order to describe a potential as such, one must describe how it is actualized. Therefore, the material cause of a thing is not conceptually separable from this thing. The form of a potential is the form of its actualization. A potential and its actualization are neither two different things, nor are they two parts, ingredients, or constituents of one thing, nor is one an attribute of the other. They are the same in the sense that one of them is potentially the very same as the other actually is. A potential beaver is potentially the very same as an actual

beaver (namely a beaver). On the other hand, a potential and its realization are not identical because the potential may fail to be realized, so that it is only potentially what the actualization is. Actualities cannot, as such, remain merely potential. Potentials and actualities have different modal properties.

Essences

The matter of a natural thing is something that potentially is a typical instance of the thing's type. The paradigmatic form of a natural thing is what its matter potentially is. "Potentially" has a generic and normative sense here: What a thing potentially is, is what instances of its kind typically are. The paradigmatic form of a thing is not a feature or property of this thing and it is not a further thing that would be related to it, but it is also not identical to this thing. Paradigmatic forms are essences, and the essence of a thing is what its definition defines. Essences are general because definitions are general. The paradigmatic form of a thing is what this thing is, and what this thing is may be what other things are as well. This, however, does not imply that essences are properties. The definition of a compound thing defines a compound thing, and if the essence of the thing is what this definition defines, it must also be a compound thing.

Before we can attribute properties or features to anything, we must identify a subject to attribute them to. Essences are that in terms of which we may identify natural things as instances of types. The difference between these two acts, identifying a thing and attributing properties to it, can be explained by using Austin's distinction between two directions of fit. Austin distinguishes between several ways in which a sentence such as "#123 is a rhombus" may be used. For instance, one may utter it as an answer to the question, about an item (#123), whether it fits a given type ("rhombus"). This is what Austin calls casting: An item is referred to in its capacity of fitting a description. In casting, the direction of fit is item to type; item #123 is chosen in virtue of fitting a description. Another possible use of "#123 is a rhombus" is to give it as an answer to the question, about #123, what type it instantiates. Here, #123

is described as a rhombus, and the direction of fit is description to item. The description "rhombus" is presented as fitting the item. Austin calls this use <u>stating</u>.

On the basis of Austin's original distinction between casting and stating, I distinguish two parts of what I call a Sellarsian sentence ("This such is so and so"). In a Sellarsian sentence, "this such" casts an item as falling under a certain description or fitting a type, and "is so and so" states that the item thus cast has certain properties. Sellarsian sentences are somewhat artificial constructs because it happens only very rarely that an item is freshly cast and described in one single sentence, and there are many ways of casting things other than uttering (parts of) sentences. The division of labor between the parts of a Sellarsian sentence needs to be projected onto our use of language and thought as a whole. If this succeeds, it may help us understand what essences are. Essences may then be taken to be that in terms of which we cast a thing, in order to state something of it.

Casting something is like calling a name. One may call a name without knowing whether anyone with that name is present. If someone responds, one has good prima facie reasons for assuming that this person has that name. When one successfully casts an item as an instance of a type, one also has good prima facie reasons for assuming that it actually is an instance of this type, even though in some cases it may fail to be one. Whether it is possible to miscast an item as something else depends on how general the casting term is. "Empirical object" is one of the most general casting terms, and if Kant is right, certain substantive claims must hold true of all objects that we can possibly cast as empirical objects. By reflecting on the way in which we must cast empirical objects in order to get hold of them, we can find out certain things that must be true of them. Since we can do this prior to actually casting and investigating any actual item, this reflection may result in a priori knowledge about all empirical objects. Further, it does not seem to be possible to miscast an empirical object for something else. Everything that can be cast by using an empirical casting term must be an empirical object. It may be that there is no empirical object at all to be cast, but if one succeeds in casting anything as an empirical object, it must be one. As such,

this object must have the features that all empirical objects must have.

"Natural thing" is a less general but still quite general casting term. It casts a subclass of empirical objects, just as "empirical object" casts a subclass of "object." By reflecting on the way in which we must cast natural things and processes, we may get to know certain things that must hold true of all natural things. I take it that when Aristotle tells us what questions we must ask in order to see natural things and processes as the natural things and processes they are, he engages in this kind of reflection. He tells us something that we can know a priori about all natural things and processes.

The Formal Cause

Formal causes are essences of natural things, and essences are general. To describe the essence of a thing is to describe the general type that the thing instantiates. I argue that the safest way to define a type of natural things is to pick a focal instance and specify a relation that all other instances must have to this instance. Further, I show that in cases where this relation involves reproduction or copying, its description involves standards of success. That something is a copy of another thing does not simply mean that it does in fact resemble the first; it only means that it is <u>supposed</u> to resemble the first in certain respects. There may be failed and atypical copies. To call something a copy or replica of another thing is thus not to say what it is like. Rather, it is to say what it should be like, and thus what standards of typicality apply to it. Further, that one living being is an offspring of another one does not even imply that it is supposed to resemble this other living being. The offspring may be more typical than its parent. All it must do in order to qualify as an offspring is to meet certain standards of health, so that it is capable of living a life of a certain kind. In order to qualify as a good offspring, it must meet some more specific standards of typicality, which are tied to its nature.

Natural things have inherent principles of motion and rest, and this means that there are objective reasons why certain standards of typicality apply to them. They

can only be cast, as the natural things they are, by applying these standards. Natural things may fail to act and develop according to their own principles of motion and rest. When this happens, they are atypical by standards that lie in their own nature. By casting a natural thing as an instance of a type, one sets up a certain standard, as it were, and waits for something that subscribes to it (and may or may not satisfy it). The standards that we set up when we cast a natural thing as such may remain unsatisfied in two ways. First, nothing might show up that is subject to them. Second, something might show up and be subject to them but still not satisfy them.

The Efficient Cause

The material cause of a natural thing is something that potentially is that thing. Accordingly, the efficient cause of a natural process might be taken to be something that potentially is this process (so that not everything that potentially is a natural process is its efficient cause). This, however, is not exactly what Aristotle says. Aristotle speaks of agents as efficient causes of what they do, and an agent is not potentially an action. On the other hand, agents are efficient causes only insofar as they act, and what they do is a natural process only insofar as it is governed by a principle of motion and rest inherent in the agent. Therefore, one cannot, in this context, separate agents from their actions, or actions from their agents. When an agent realizes its potential, a potentially acting thing becomes an actually acting thing; or what is the same: a potential action of a thing comes to be an actual action of this thing. Aristotle usually refers to the things involved in a process as efficient causes, and he often refers to examples where one cause is an efficient cause of another process. These cases are more familiar, but less basic; they can be described in terms of the more basic notion of an efficient cause introduced here.

The efficient cause is the one of the four causes that has an effect; there are no effects corresponding to material, formal, or final causes. There is also no such thing as material, formal, or final causation. All causation is efficient causation. I argue that even though one may always distinguish between an efficient cause and its effect,

causation is not a relation between distinct items. It is also not a special kind of process that connects two distinct processes. Rather, the effect of an efficient cause is the realization of its potential. Once a potential is realized, it does not differ from its realization. Therefore, once and insofar as an efficient cause actually leads to its effect, it is not distinct from this effect. To describe a process as the efficient cause of another process is to describe both of them as the beginning and end of the same process; or more generally, as different stages of the same process. Since a potential is not conceptually separable from its realization, one cannot first perfectly separate cause and effect and then find out what their causal relation consists in.

That a process is causal means that it admits of a complex description, according to which the beginning of it turns into a process of a generally specifiable type. All causal processes must therefore be instances of generally specifiable types. Further, all processes are such that as long as they are going on, they are not yet complete, and some part of them has not yet happened. This part must be something specific, e.g., a movement with a certain direction, for if nothing specific is missing, nothing is missing. Thus all ongoing natural processes must proceed toward an end that they might fail to meet. (If they had reached this end, they would be over and would not any longer go on.) This end, which is their completion, is their final cause.

The Final Cause

Final causes, in a basic sense, are for natural processes what formal causes are for natural things. They are the essences (or "limits") according to which natural processes proceed. Just as the matter of a natural thing is inseparable from its paradigmatic form, there can be no causal processes without final causes. The proper account of the nature of a natural thing implies certain standards of typicality. To proceed for the sake of a final cause is to proceed according to such a standard.

The essence of an item is also that in terms of which it may be cast as a subject of possible predication. Likewise, the final cause is that in terms of which we may cast an ongoing process as an instance of a type. The essence of a process makes us expect

a certain further course that the process should take, which is the course that instances of its type typically take. The final cause of a natural process is the course that is typical for instances of its type. Since the typical course of a process is not a further process that this process undergoes, the final cause of a natural process is a paradigmatic instance of its kind. This paradigmatic instance is "the best" in the sense that it constitutes the best-case scenario.

We usually refer to things in order to attribute properties to them. We often refer to processes in order to say that they are parts of, contribute to, or are involved in further processes. This explains why the notion of a final cause, as introduced here, is less familiar than the notion of a remote aim or purpose. Since in most cases, the final cause of a process is that in terms of which we cast it in the first place, it would be redundant to explicitly attribute this final cause to the process. For instance, when we have already referred to a process as an instance of pushing, saying that its final cause is pushing does not add anything of interest. More often, we associate the basic final cause of a complex process with the final causes of its parts. We say, for instance, that the final cause of pushing is opening the door. Opening the door is a more complex process, of which pushing is a part. The proximate final cause of opening the door is opening the door, and this proximate final cause is here referred to as a remote final cause of its part, pushing. Complex processes have complex final causes. These complex final causes may be reduced to their simple parts, but this is not to reduce final causes to efficient causes. Complex processes also have complex efficient causes, and these complex efficient causes may be reduced to their less complex parts. Still, these parts have final causes, since they are what they are only by taking some specific course.

I elaborate on this by discussing teleological reasoning. Teleological reasoning is about the mereology of causal processes and their final causes. Practical reasoning is an instance of teleological reasoning. In practical reasoning, we relate actions that are means to actions that are ends. Practical reasoning mirrors speculative reasoning in the following sense. From a task A, one may practically infer B as a derived task, if

one can show, by speculative reasoning, that B leads to A. However, practical reasoning mirrors speculative reasoning only on a large scale. The speculative syllogism is not mirrored in its details. There are no special inference rules that apply exclusively to teleological reasoning; rather, it is an application of ordinary inference rules in a special context. If there were such things as formal and material reasoning, formal reasoning would presumably mirror material reasoning in the same way.

Bibliography

Albritton, Rogers (1957). "II. Forms of Particular Substances in Aristotle's Metaphysics." The Journal of Philosophy 54 (22), p. 699-708.

Allan, D. J. (1955). "The Practical Syllogism." In L. de Raeymaker (ed.), Autour d'Aristote, Offert à A. Mansion. Louvain: Publications Universitaires.

Allen, James (unpublished). "Aristotle on Chance."

Allison, Henry E. (2004). Kant's Transcendental Idealism. New Haven: Yale University Press.

Ameriks, Karl (1992). "Kantian Idealism Today." History of Philosophy Quarterly 9 (3), p. 329-342.

Annas, Julia (1982). "Aristotle on Inefficient Causes." The Philosophical Quarterly 32 (129), p. 311-326.

Anscombe, Elizabeth (1953). "The Principle of Individuation." Proceedings of the Aristotelian Society Supplement (27), p. 83-96.

——— (1957). Intention. London: Basil Blackwell.

——— (1981). "Causality and Determination." In Collected Philosophical Papers, vol. 2. Oxford: Basil Blackwell.

——— (1995). "Practical Inference." In Rosalind Hursthouse, Gavin Lawrence and Warren Quinn (eds.), Virtues and Reasons. Oxford: Clarendon Press.

Anscombe, Elizabeth, and Geach, Peter T. (1961). Three Philosophers. Oxford: Basil Blackwell.

Aquinas, Thomas (1954). In Octo Libros Physicorum Aristotelis Expositio. Taurin: Marietti.

Aristotle (1831-70). Aristotelis Opera, ed. Bekker. Berin: Georg Reimer.

——— (1957). Physics, Books I-IV, Loeb edition. London: Harvard University Press.

Aronson, Jerrold L. (1971). "On the Grammar of 'Cause'." <u>Synthese</u> 22, p. 414-430.

Austin, J. L. (1962). <u>Sense and Sensibilia</u>. Oxford: Clarendon Press.

——— (1979). <u>Philosophical Papers</u>. Oxford: Clarendon Press.

Balme, David M. (1987). "Teleology and Necessity." In Allan Gotthelf and James G. Lennox (eds.), <u>Philosophical Issues in Aristotle's Biology</u>. New York: Cambridge University Press.

Barcan Marcus, Ruth (1971). "Essential Attribution." <u>The Journal of Philosophy</u> 68 (7), p. 187-211.

Beckner, Morton (1968). <u>The Biological Way of Thought</u>. Berkeley: University of California Press.

Beebee, Helen (2006). <u>Hume on Causation</u>. London: Routledge.

Beere, Jonathan (2006). "Potentiality and the Matter of Composite Substance." <u>Phronesis</u> 51 (4), p. 303-329.

Belnap, Nuel D. (1970). "Conditional Assertion and Restricted Quantification." <u>Noûs</u> 4 (1), p. 1-12.

Ben-Yami, Hanoch (2001). "The Semantics of Kind Terms." <u>Philosophical Studies</u> 102 (2), p. 155-184.

Bigelow, John, and Robert Pargetter (1987). "Functions." <u>The Journal of Philosophy</u> 84, p. 181-196.

Bobzien, Susanne (1999). "Chrysippus' Theory of Causes." In Katerina Ierodiakonou (ed.), <u>Topics in Stoic Philosophy</u>. Oxford: Clarendon Press, p. 196-242.

Bonitz, Hermann (1870). <u>Index Aristotelicus</u>. Berlin: Reimer.

Bostock, David (1995). <u>Aristotle, Metaphysics Book Z and H</u>. Oxford: Clarendon Press.

Brand, Myles (1975). "On Philosophical Definitions." <u>Philosophy and Phenomenological Research</u> 36 (2), p. 151-172.

Brandis, Christian August (1853). <u>Handbuch der Geschichte der Griechisch-Römischen Philosophie</u>, vol II, 2, 1: Aristoteles, seine akademischen

Zeitgenossen und nächsten Nachfolger. Berlin: Reimer.

Brody, Baruch A. (1973). "Why Settle for Anything Less Than Good Old-Fashioned Aristotelian Essentialism." Noûs 7 (4), p. 351-365.

Brunschwig, J. (1979). "La forme, prédicat de la matière." In Pierre Aubenque (ed.), Études sur la Métaphysique d'Aristote. Paris: Vrin.

Buck, Roger C., and David L. Hull (1966). "The Logical Structure of the Linnaean Hierarchy." Systematic Zoology 15 (2), p. 97-111.

Carlson, Gregory N., and Francis Jeffrey Pelletier (1995). The Generic Book. Chicago: University Press.

Cartwright, Nancy (2007). Hunting Causes and Using Them. Oxford: Cambridge University Press.

Chakravartty, Anjan (2007). A Metaphysics for Scientific Realism. Cambridge: University Press.

Chappell, Vere (1973). "Matter." The Journal of Philosophy 70 (19), p. 679-696.

Charles, David (1991). "Teleological Causation in the Physics." In Judson, Aristotle's Physics.

——— (1994). "Matter and Form: Unity, Persistence, and Identity." In Scaltsas et al., Unity, Identity, and Explanation in Aristotle's Metaphysics.

——— (2002). Aristotle on Meaning and Essence. Oxford: Clarendon Press.

Charlton, William (1970). Aristotle's Physics I,II. Oxford: Clarendon Press.

——— (1972). "Aristotle and the Principle of Individuation." Phronesis 17, p. 239-249.

——— (1983). "Prime Matter: A Rejoinder." Phronesis 28, p. 197-210.

Cherniss, Harold F. (1954). "A Much Misread Passage of the Timaeus (Timaeus 49c7-50b5)." American Journal of Philology 75, p. 113-130.

Chisholm, Roderick M. (1969). "Austin's Philosophical Papers." in K. T. Fann (ed.), Symposium on J. L. Austin. London: Routledge.

——— (1976). "The Agent as Cause." In Myles Brand and Douglas Walton (eds.), Action Theory, Dordrecht: Reidel.

Clark, Philip (2001). "The Action as Conclusion." Canadian Journal of Philosophy 31 (4), p. 481-506.

Cocchiarella, Nino (1976). "On the Logic of Natural Kinds." Philosophy of Science 43 (2), p. 202-222.

Collingwood, Robin G. (1940). An Essay on Metaphysics. Oxford: Clarendon Press.

Cooper, John M. (1973). "Chappell and Aristotle on Matter." The Journal of Philosophy 70 (19), p. 696-698.

Corcilius, Klaus (2007). Streben und Bewegen. Berlin: De Gruyter.

——— (2008a). "Aristoteles' praktische Syllogismen in der zweiten Hälfte des 20. Jahrhunderts." Logical Analysis and History of Philosophy 11, p. 101-32.

——— (2008b). "Two Jobs for Aristote's Practical Syllogism?" Logical Analysis and History of Philosophy 11, p. 163-84.

Cornford, Francis MacDonald (1937). Plato's Cosmology: The Timaeus of Plato. London: Routledge.

Coughlin, Glen (2005). Aristotle, Physics or Natural Hearing. South Bend: St. Augustine's Press.

Cousin, D. R. (1955). "How Not to Talk." Analysis 15 (4), p. 73-81.

Cox, J. W. Roxbee (1955). "Fitting and Matching: A Note on Professor Austin's 'How to Talk'." Analysis 16 (1), p. 6-11.

CSEL = Corpus Scriptorum Ecclesiasticorum Latinorum. Wien: Verlag der Akademie der Wissenschaften.

Cummins, Robert (1975). "Functional Analysis." The Journal of Philosophy 72, p. 741-765.

Dancy, Russell (1975). "On Some of Aristotle's First Thoughts About Substance." The Philosophical Review 84 (3), p. 338-373.

——— (1978). "On Some of Aristotle's Second Thoughts About Substance: Matter."

The Philosophical Review 87, p. 372-413.

Dardis, Anthony (1993). "Sunburn: Independence Conditions on Causal Relevance."
Philosophy and Phenomenological Research 53 (3), p. 577-598.

Davidson, Donald (1980). Essays on Actions and Events. Oxford: Clarendon Press.

Dobzhansky, Theodosius (1935). "A Critique of the Species Concept in Biology."
Philosophy of Science 2 (3), p. 344-355.

Donnellan, Keith S. (1966). "Reference and Definite Description." The Philosophical
Review 75, p. 281-304.

——— (1968). "Putting Humpty Dumpty Together Again." The Philosophical
Review 77 (2), p. 203-215.

——— (1977). "The Contingent a Priori and Rigid Designators." Midwest Studies in
Philosophy 2, p. 45-60.

Dowe, Phil (1995). "Causality and Conserved Quantities: A Reply to Salmon."
Philosophy of Science 62 (2), p. 321-333.

Ducasse, C. J. (1932). "Of the Nature and Efficacy of Causes." The Philosophical
Review 41 (4), p. 395-399.

——— (1957). "On the Analysis of Causality." The Journal of Philosophy 54 (13), p.
422-426.

Ehring, Douglas (1997). Causation and Persistence. New York: Oxford University
Press.

Engstrom, Stephen (2002). "Kant's Distinction between Theoretical and Practical
Knowledge." Harvard Review of Philosophy 10, p. 49-63.

Fair, David (1979). "Causation and the Flow of Energy." Erkenntnis 14 (3), p.
219-250.

Falcon, Andrea (2008). "Aristotle on Causality." Stanford Encyclopedia of
Philosophy, http://plato.stanford.edu/archives/fall2008/entries/aristotle-
causality/.

Fales, Evan (1979). "Relative Essentialism." The British Journal for the Philosophy of

Science 30 (4), p. 349-370.

——— (1990). Causation and Universals. London: Routledge.

Farber, Paul Lawrence (1976). "The Type-Concept in Zoology During the First Half of the Nineteenth Century." Journal of the History of Biology 9 (1), p. 93-119.

Feldman, Fred (1973). "Sortal Predicates." Noûs 7 (3), p. 268-282.

Fine, Gail (1987). "Forms as Causes: Plato and Aristotle." In Andreas Graeser (ed.), Mathematics and Metaphysics in Aristotle. Bern: Haupt.

Fine, Kit (1992). "Aristotle on Matter." Mind 101 (401), p. 35-57.

Fitzpatrick, William J. (2000). Teleology and the Norms of Nature. New York: Garland Publishing.

Foot, Philippa (2001). Natural Goodness. Oxford: Clarendon Press.

Fraassen, Bas C. van (1980a). "A Re-examination of Aristotle's Philosophy of Science." Dialogue [Canadian Philosophical Review] 19, p. 20-45.

——— (1980b). The Scientific Image. Oxford: Clarendon Press.

Frankl, Wilhelm M. (1907). "Illusive Reihen." Annalen der Naturphilosophie 6, p. 16-19.

Franklin, James (1986). "Aristotle on Species Variation." Philosophy 61 (236), p. 245-252.

Franklin-Hall, Laura (2008). From a Microbiological Point of View. PhD thesis. New York: Columbia University.

Frede, Michael (1987). Essays in Ancient Philosophy. Minneapolis: University of Minnesota Press.

——— (1992). "On Aristotle's Conception of the Soul." In Martha C. Nussbaum and Amélie Oksenberg Rorty (eds.), Essays on Aristotle's De Anima. Oxford: Clarendon Press.

——— (1994). "Aristotle's Notion of Potentiality in Metaphysics Θ." In Scaltsas et al. Unity, Identity, and Explanation in Aristotle's Metaphysics.

Frede, Michael, and Günther Patzig (1988). Aristotleles' Metaphysik Z. München: C. H. Beck.

Freeland, Cynthia (1991). "Accidental Causes and Real Explanations." In Judson, Aristotle's Physics.

Furley, David (1985). "The Rainfall Example in Physics II 8." In Allan Gotthelf, ed., Aristotle on Nature and Living Things. Pittsburgh: Mathesis.

——— (1996). "What Kind of Cause is Aristotle's Final Cause?" In Michael Frede and Gisela Striker (eds.), Rationality in Greek Thought. Oxford: Clarendon Press.

Garrett, Don (1997). Cognition and Commitment in Hume's Philosophy. New York: Oxford University Press.

Gasking, Douglas (1996). Language, Logic and Causation. Carlton South: Melbourne University Press.

Geach, Peter Thomas (1962). Reference and Generality. Ithaca: Cornell University Press.

——— (1975). "Teleological Explanation." In Stephan Körner (ed.), Explanation. Oxford: Blackwell.

Ghiselin, Michael (1981). "Categories, Life, and Thinking." Behavioral and Brain Sciences 4, p. 269-313.

Gill, Mary Louise (1989). Aristotle on Substance. Princeton: University Press.

——— (1991). "Aristotle on Self-Motion." In Judson, Aristotle's Physics.

——— (1994). "Individuals and Individuation in Aristotle." In Scaltsas et al., Unity, Identity, and Explanation in Aristotle's Metaphysics.

Glennan, Stuart S. (1996). "Mechanisms and the Nature of Causation." Erkenntnis 44 (1), p. 49-71.

Godfrey-Smith, Peter (forthcoming). "Causal Pluralism." In Helen Beebee (ed.), The Oxford Handbook to Causation.

Goodman, Nelson (1983). Fact, Fiction, and Forecast. Cambridge: Harvard University

Press.

Gotthelf, Allan (1987). "Aristotle's Conception of Final Causality." In Allan Gotthelf and James G. Lennox (eds.), Philosophical Issues in Aristotle's Biology. New York: Cambridge University Press.

——— (1988). "The Place of the Good in Aristotle's Natural Teleology." Proceedings of the Boston Area Colloquium in Ancient Philosophy 4, p. 113-139.

Granger, Herbert (1985). "The Scala Naturae and the Continuity of Kinds." Phronesis 30 (2), p. 181-200.

——— (1987a). "Aristotle and the Finitude of Natural Kinds." Philosophy 62 (242), p. 523-526.

——— (1987b). "Deformed Kinds and the Fixity of Species." The Classical Quarterly 37 (1), p. 110-116.

——— (2000). "Metaphysics Z.11.1036b28: Αἰσθητόν or Αἰσθητικόν?" The Classical Quarterly 50 (2), p. 415-423.

Griffiths, Paul E. (1996). "Darwinism, Process Structuralism, and Natural Kinds." Philosophy of Science 63, p. S1-S9.

Grim, Patrick (1974). "Wright on Functions." Analysis 35, p. 62-64.

Gulley, Norman (1960). "The Interpretation of Plato, Timaeus 49 D-E." The American Journal of Philology 81 (1), p. 53-64.

Gupta, Anil (1980). The Logic of Common Nouns. New Haven.

Haag, Johannes (2007). Erfahrung und Gegenstand. Frankfurt: Klostermann.

Hacking, Ian (1991). "A Tradition of Natural Kinds." Philosophical Studies 61, p. 109-126.

Hankinson, R. J. (1998). Cause and Explanation in Ancient Greek Thought. Oxford: Clarendon Press.

Hanson, Norwood Russell (1958). Patterns of Discovery. Cambridge: University Press.

Hantz, Harold Donovan (1939). The Biological Motivation in Aristotle. PhD thesis.

New York: Columbia University.

Hart, H. L. A., and Tony Honoré (1985). Causation in the Law, 2nd ed. Oxford: Clarendon Press.

Harte, Verity (1996). "Aristotle Metaphysics H6: A Dialectic with Platonism." Phronesis 41 (3), p. 276-304.

Harvey, William (1976). An Anatomical Disputation Concerning the Movement of the Heart, tr. Whitteridge. Oxford: Blackwell.

Haslanger, Sally (1994). "Parts, Compounds, and Substantial Unity." In Scaltsas et al., Unity, Identity, and Explanation in Aristotle's Metaphysics.

Hennig, Boris (2008). "Matter in Z3." Foundations of Science 13, p. 199-215.

——— (2009). "The Four Causes." The Journal of Philosophy 106(3), p. 137-160.

Heuer, Peter (2008). Art, Gattung, System. Freiburg: Alber.

Heymans, Gerard (1902). "Über Erklärungshypothesen und Erklären überhaupt." Annalen der Naturphilosophie 1, p. 473-485.

Hocutt, Max (1974). "Aristotle's Four Becauses." Philosophy 49 (190), p. 385-399.

Hübner, Johannes (2001). "Ursache / Wirkung I." In Historisches Wörterbuch der Philosophie, vol. 11, Darmstadt: Wissenschaftliche Buchgesellschaft.

Hugly, Philip, and Charles Sayward (1981). "Expressions and Tokens." Analysis 41, p. 181-187.

Hull, David L. (1976). "Are Species Really Individuals?" Systematic Zoology 25 (2), p. 174-191.

Humberstone, I. Lloyd (1992). "Direction of Fit." Mind 101 (401), p. 59-83.

Hume, David (1975). Enquiries Concerning Human Understanding and Concerning the Principles of Morals. Ed. Selby-Bigge. Oxford: Clarendon Press.

——— (2000). A Treatise of Human Nature. Ed. David Fate Norton. Oxford: University Press.

Irwin, Terence H. (1988). Aristotle's First Principles. Oxford: Clarendon Press.

Johnson, Monte R. (2005). Aristotle on Teleology. Oxford: Clarendon Press.

Johnston, Mark (1997). "Manifest Kinds." The Journal of Philosophy 94 (11), p. 564-583.

Johnston, Timothy D. (1981). "'Species-typicality': Can Individuals Have Typical Parts?" Behavioral and Brain Sciences 4, p. 291-292.

Judson, Lindsay, ed. (1991). Aristotle's Physics: A Collection of Essays. Oxford: Clarendon Press.

Kant, Immanuel (1781/1787). Kritik der reinen Vernunft. First edition, Riga: Hartknoch 1781 = A. Second edition, 1787 = B.

——— (1902-1983). Gesammelte Schriften (Akademie-Ausgabe). Berlin: Reimer.

Kenny, Anthony J. (1963). Action, Emotion, and Will. London: Routledge.

——— (1966). "Practical Inference." Analysis 26 (3), p. 65-75.

——— (1975). Will, Freedom, and Power. London: Basil Blackwell.

——— (1979). Aristotle's Theory of the Will. London: Duckworth.

Kirwan, Christopher (1993). Aristotle, Metaphysics Books Γ, Δ, and E. Oxford: Clarendon Press.

Kitcher, Philip (1984). "Species." Philosophy of Science 51 (2), p. 308-333.

Kripke, Saul A. (1980). Naming and Necessity. Oxford: Blackwell.

Kung, Joan (1977). "Aristotle on Essence and Explanation." Philosophical Studies 31 (6), p. 361-383.

——— (1981). "Aristotle on Thises, Suches, and the Third Man Argument." Phronesis 26 (3), p. 207-247.

LaPorte, Joe (1996). "Chemical Kind Term Reference and the Discovery of Essence." Noûs 30 (1), p. 112-132.

Le Pore, Ernest, and Barry Loewer (1987). "Mind Matters." The Journal of Philosophy 84 (11), p. 630-642.

Lee, Edward N. (1967). "On Plato's Timaeus, 49D4-E7." The American Journal of

Philology 88 (1), p. 1-28.

Lennox, James G. (2001a). Aristotle, On the Parts of Animals I-IV. Oxford: Clarendon Press.

——— (2001b). Aristotle's Philosophy of Biology. New York: Cambridge University Press.

——— (unpublished). "Bios and Explanatory Unity in Aristotle's Biology."

Lesher, James H. (1971). "Aristotle on Form, Substance, and Universals: A Dilemma." Phronesis 16 (2), p. 169-178.

Lewis, David K. (1986). "Causation." In Philosophical Papers, vol. II. New York: Oxford University Press.

——— (2004). "Void and Object." In John D. Collins, Edward J. Hall and L. A. Paul (eds.), Causation and Counterfactuals. Cambridge: MIT Press.

Lewis, Frank A. (1982). "Accidental Sameness in Aristotle." Philosophical Studies 42 (1), p. 1-36.

——— (1994). "Aristotle on the Relation between a Thing and its Matter." In Scaltsas et al., Unity, Identity, and Explanation in Aristotle's Metaphysics.

Linnaeus, Carl (1758). Systema Naturae, 10th ed. Holmiae: L. Salvius.

Longuenesse, Béatrice (1998). Kant and the Capacity to Judge. Princeton: University Press.

Longworth, Francis (2006). Causation, Counterfactual Dependence, and Pluralism. PhD thesis. University of Pittsburgh.

Loux, Michael J. (1979). "Form, Species and Predication in Metaphysics Z, H, and Θ." Mind 88 (349), p. 1-23.

——— (1991), Primary Ousia. Ithaca: Cornell University Press.

Machamer, Peter, Lindley Darden, et al. (2000). "Thinking About Mechanisms." Philosophy of Science 67 (1), p. 1-25.

Mackie, John L. (1974). The Cement of the Universe. Oxford: Clarendon Press.

Malter, Rudolf (1981). "Reflexionsbegriffe." Philosophia Naturalis 19(2), p. 124-150.

Martinich, A. P. (1977). "The Attributive Use of Proper Names." Analysis 37 (4), p. 159-163.

Matthen, Mohan (forthcoming). "On Aristotle's Teleology of Living Things." In Georgios Anagnostopoulus (ed.), Blackwell Companion to Aristotle.

Matthews, Gareth B. (1990). "Aristotelian Essentialism." Philosophy and Phenomenological Research 50, p. 251-262.

Mayr, Ernst (1976). "Teleological and Teleonomic: A New Analysis." In Evolution and the Diversity of Life. Cambridge: Harvard University Press.

——— (1996). "What Is a Species, and What Is Not?" Philosophy of Science 63 (2), p. 262-277.

McDowell, John (1998). "Lecture II: The Logical Form of an Intuition." The Journal of Philosophy 95 (9), p. 451-470.

McLaughlin, Peter (2001). What Functions Explain. Cambridge: University Press.

Mellor, D. H. (1995). The Facts of Causation. London: Routledge.

Meyerson, Emile (1926). Identité et réalité. Paris: F. Alcan.

Micraelius, Johannes (1653). Lexicon Philosophicum. Jena: Jeremia Mamphrasius.

Millikan, Ruth Garrett (2000). On Clear and Confused Ideas. New York: Cambridge University Press.

Modrak, Deborah K. (1979). "Forms, Types, and Tokens in Aristotle's Metaphysics." Journal of the History of Philosophy 17, p. 371-381.

Moravcsik, Julius M. E. (1974). "Aristotle on Adequate Explanation." Synthese 28, p. 3-17.

——— (1991). "What Makes Reality Intelligible?" In Judson, Aristotle's Physics.

——— (1994). "Essences, Powers, and Generic Propositions." In Scaltsas et al., Unity, Identity, and Explanation in Aristotle's Metaphysics.

Müller, Anselm W. (1982). Praktisches Folgern und Selbstgestaltung nach Aristoteles.

Freiburg: Alber.

Mure, G. R. G. (1975). "Cause and Because in Aristotle." Philosophy 50 (193), p. 356-357.

Natali, Carlo (2001). The Wisdom of Aristotle. Albany: SUNY Press.

Neander, Karen (1991). "Functions as Selected Effects: The Conceptual Analyst's Defense," Philosophy of Science 58 (2), p. 168-84.

Norton, John (2003). "Causation as Folk Science." Philosopher's Imprint 3 (4), p. 1-22.

Nussbaum, Martha C. (1985). Aristotle's De Motu Animalium. Princeton: University Press.

Ostwald, Wolfgang (1903). "Über Erklärungshypothesen und Erklären überhaupt." Annalen der Naturphilosophie 2, p. 506-526.

Owens, Joseph (1963). The Doctrine of Being in the Aristotelian Metaphysics. Toronto: Pontifical Institute of Medieval Studies.

Pearl, Judea (2000). Causality. New York: Cambridge University Press.

Plato (1578). Platonis Opera. Geneva: Henricus Stephanus.

Platts, Mark (1983). "Explanatory Kinds." The British Journal for the Philosophy of Science 34 (2), p. 133-148.

Barnes, Jonathan (2006). Porphyry's Introduction. Oxford: University Press.

Price, A. W. (1979). "Doing Things Explicitly with Words." Philosophical Studies 36 (4), p. 345-357.

Putnam, Hilary (1975). Philosophical Papers. New York: Cambridge University Press.

Rapp, Christof (1995a). "Allgemeines konkret – Ein Beitrag zum Verständnis der Aristotelischen Substanzlehre." Philosophisces Jahrbuch 102(2): 83-100.

——— (1995b). Identität, Persistenz und Substantialität. Freiburg: Alber.

——— (1996). " 'Kein Allgemeines ist Substanz' (Z 13, 14-16)." In: Christof Rapp, ed., Aristoteles, Metaphysik, Die Substanzbücher (Z, H, Θ). Berlin:

Akademie Verlag.

——— (2007). <u>Aristoteles zur Einführung</u>. Hamburg: Junius Verlag.

Rödl, Sebastian (2003). "Norm und Natur." <u>Deutsche Zeitschrift für Philosophie</u> 51(1): 99-114.

——— (2005). <u>Kategorien des Zeitlichen</u>. Frankfurt: Suhrkamp.

Rorty, Richard (1973). "Genus as Matter: A Reading of Metaphysics Z-H." In: Alexander Mourelatos and Richard Rorty (eds.), <u>Exegesis and Argument</u>. Assen: Van Gorcum.

Ross, W. D. (1936). <u>Aristotle's Physics</u>. Oxford: Clarendon Press.

——— (1949). <u>Aristotle's Prior and Posterior Analytics</u>. Oxford: Clarendon Press.

Ruse, Michael (1987). "Biological Species: Natural Kinds, Individuals, or What?" <u>The British Journal for the Philosophy of Science</u> 38 (2), p. 225-242.

Russell, Bertrand (1912). "On the Notion of Cause." <u>Proceedings of the Aristotelian Society</u> 13, p. 1-26.

——— (1948). <u>Human Knowledge, Its Scope and Limits</u>. London: Allen and Unwin.

Salmon, Wesley C. (1998). <u>Causality and Explanation</u>. New York: Oxford University Press.

Scaltsas, Theodore, David Charles, and Mary Louise Gill, eds. (1994). <u>Unity, Identity, and Explanation in Aristotle's Metaphysics</u>. Oxford: Clarendon Press.

Schlick, Moritz (1979). <u>Philosophical Papers</u>. Dordrecht Reidel.

Schofield, Malcolm (1991). "Explanatory Projects in Physics, 2. 3 and 7." In Henry Blumenthal and Howard Robinson (eds.), <u>Aristotle and the Later Tradition</u>. Oxford Studies in Ancient Philosophy, Supplementary volume.

Schrödinger, Erwin (1945). <u>What is Life?</u> Cambridge: University Press.

Sellars, Wilfrid (1957). "I. Substance and Form in Aristotle." <u>The Journal of Philosophy</u> 54 (22), p. 688-699.

——— (1968). <u>Science and Metaphysics</u>. London: Routledge & Kegan Paul.

——— (1976). "Kant's Transcendental Idealism." <u>Collections of Philosophy</u> 6, p. 165-181.

——— (1977). "Aristotle's Metaphysics: An Interpretation." In <u>Philosophical Perspectives, History of Philosophy</u>. Atascadero: Ridgview Publishing Company.

Silverman, Allan (1992). "Timaean Particulars." <u>Classical Quarterly</u> 42, p. 87-113.

Skyrms, Brian (1984). "EPR: Lessons for Metaphysics." <u>Midwest Studies in Philosophy</u> IX, p. 245-255.

Smith, J. A. (1921). "τόδε τι in Aristotle." <u>Classical Review</u> 35, p. 19.

Sobel, David, and David Copp (2001). "Against Direction of Fit Accounts of Belief and Desire." <u>Analysis</u> 61 (269), p. 44-53.

Sosa, Ernest, and Michael Tooley, eds. (1993). <u>Causation</u>. New York: Oxford University Press.

Spirtes, Peter, Clark N. Glymour, et al. (2000). <u>Causation, Prediction, and Search</u>, 2nd ed. Cambridge: MIT Press.

Sprague, Rosamond K. (1968). "The Four Causes: Aristotle's Exposition and Ours." <u>Monist</u> 52 (2), p. 298-300.

Sprigge, Timothy L. S. (1971). "Final Causes." <u>Aristotelian Society Supplement</u> 45, p. 149-170.

Stahl, Donald E. (1981). "Stripped Away: Some Contemporary Obscurities Surrounding Metaphysics Z.3." <u>Phronesis</u> (29), p. 177-180.

Stekeler-Weithofer, Pirmin (unpublished). "Das Einzelne und das Allgemeine. Generisches Wissen als materialbegriffliche Sinnbedingung empirischer Aussagen."

Strawson, Peter F. (1959). <u>Individuals</u>. London: Routledge.

——— (1961). "Singular Terms and Predication." <u>The Journal of Philosophy</u> 58 (15), p. 393-412.

——— (1985). "Causation and Explanation." In Bruce Vermazen and Merril B.

Hintikka (eds.), Essays on Davidson. Oxford: Clarendon Press.

Strobel, Benedikt (2007). "Dieses" und "So etwas." Göttingen: Vandenhoek & Ruprecht.

Taylor, Richard (1966). Action and Purpose. Englewood Cliffs: Prentice-Hall.

Thompson, Michael (2008). Life and Action. Cambridge: Harvard University Press.

Thomson, Judith Jarvis (1998). "The Statue and the Clay." Noûs 32 (2), p. 149-173.

Tugendhat, Ernst (2003). ΤΙ ΚΑΤΑ ΤΙΝΟΣ. Freiburg: Alber.

Tversky, Amos (1977). "Features of Similarity." Psychological Review 84 (4), p. 327-352.

Vlastos, Gregory (1969). "Reasons and Causes in the Phaedo." The Philosophical Review 78 (3), p. 291-325.

Wallace, J. (1965). "Sortal Predicates and Quantification." The Journal of Philosophy 62, p. 8-13.

Warnock, G. J. (1989). J. L. Austin. London: Routledge.

Watkins, Eric (2005). Kant and the Metaphysics of Causality. New York: Cambridge University Press.

Wedin, Michael V. (1993). "Nonsubstantial Individuals." Phronesis 38 (2), p. 137-165.

Whewell, William (1847). The Philosophy of the Inductive Sciences. London: John W. Parker.

Whiting, Jennifer E. (1986). "Form and Individuation in Aristotle." History of Philosophy Quarterly 3 (4), p. 359-377.

——— (1991). "Metasubstance: Critical Notice of Frede-Patzig and Furth." The Philosophical Review 100 (4), p. 607-639.

Wieland, Wolfgang (1975). "The Problem of Teleology." In Jonathan Barnes, Malcolm Schofield and Richard Sorabji (eds.), Articles on Aristotle. London: Duckworth.

——— (1992). <u>Die Aristotelische Physik</u>. Göttingen: Vandenhoek & Ruprecht.

Wiggins, David (2001). <u>Sameness and Substance Renewed</u>. New York: Cambridge University Press.

Wilkerson, T. E. (1993). "Species, Essences and the Names of Natural Kinds." <u>The Philosophical Quarterly</u> 43 (170), p. 1-19.

Winston, Judith E. (1995). <u>Describing Species</u>. New York: Columbia University Press.

Witt, Charlotte (1989). <u>Substance and Essence in Aristotle</u>. Ithaca: Cornell University Press.

Wolterstorff, Nicholas (1975). "Toward an Ontology of Art Works." <u>Noûs</u> 9 (2), p. 115-142.

Woodfield, Andrew (1976). <u>Teleology</u>. Cambridge: University Press.

Woods, Michael (1967). "Problems in Metaphysics Z, Chapter 13." In Julius Moravcsik (ed.), <u>Aristotle: A Collection of Critical Essays</u>, New York: Doubleday.

Woodward, James (2003). <u>Making Things Happen</u>. New York: Oxford University Press.

Wright, Georg Henrik von (1983). <u>Practical Reason</u>, Philosophical Papers vol. 1. London: Basil Blackwell.

Wright, Larry (1973). "Functions." <u>Philosophical Review</u> 82, p. 139-168.

Zeyl, Donald J. (1975). "Plato and Talk of a World in Flux: Timaeus 49A6-50B5." <u>Harvard Studies in Classical Philology</u> 79, p. 125-148.

——— (2000). <u>Plato. Timaeus</u>. Cambridge: Hackett Publishing Company.