

# Aristotle's Causal Pluralism

Nathanael Stein

Department of Philosophy, Western Michigan University,  
Moore Hall, Kalamazoo, MI 49008, USA, nstein0001@gmail.com

*Abstract:* Central to Aristotle's metaphysics and epistemology is the claim that '*aitia*' – 'cause' – is "said in many ways", i.e., multivocal. Though the importance of the four causes in Aristotle's system cannot be overstated, the nature of his pluralism about *aitiai* has not been addressed. It is not at all obvious how these modes of causation are related to one another, or why they all deserve a common term. Nor is it clear, in particular, whether the causes are related to one another as species under a single genus, such that there is a univocal definition of '*aitia*' which applies to all of them, or whether Aristotle means to assert that the four causes are homonyms. It is argued here that although there are strong reasons to group the four causes together, there are also powerful considerations on the side of homonymy. It is further argued that the four causes are more closely tied to the ontological theory of categories and predication than is often recognized. As a result, we can reconcile the competing demands of unity and plurality by taking one mode of causation, the formal cause, as basic, and accounting for the other modes with reference to it, in the manner of so-called *pros hen* homonyms.

## 1. Introduction: Aristotle's four causes

Aristotle is a causal pluralist: he believes there are four modes of causation, which have become known as formal, material, efficient, and final causation. These modes involve very distinct metaphysical relationships such as structuring, constitution, initiating motion, and being something's function or purpose, and Aristotle is certainly right to distinguish them from one another. Given their obvious differences, however, it is not clear why they should be grouped together, why they share a common term, or what could count against some other kind of metaphysical relation being deemed a causal one. Even if we think that 'cause' is a misleading English translation of Aristotle's term '*aitia*' or '*aition*', as some have argued,<sup>1</sup> we may

---

<sup>1</sup> I follow the traditional translation of '*aitia*' as 'cause' throughout, though whether or not the four causes are rightly called 'causes' – as opposed to 'because's, or 'explanations', or 'explanatory factors' – is sometimes disputed. Often the reservation has simply to do with scope: '*aitia*' appears to be used more widely than the English 'cause', at least in philosophical contexts, as will quickly become apparent to a reader of the *Physics* (see, for example, Charlton 1992, 98f.). Nothing in what follows turns on this difference. Some commentators, however, question whether Aristotle means to

wonder why all and only these four relationships are picked out as being of primary importance for Aristotle's metaphysics and epistemology. As with any pluralism, Aristotle's causal pluralism has a potential for both philosophical richness and sophistication, but only if we can answer these basic questions: How do the four causes relate to one another? What, if anything, unifies this plurality?

It is natural to understand these questions as asking what the four causes have in common. Even putting the matter in this way, however, presumes too much about the nature of Aristotle's pluralism. Aristotle may be claiming that there are four *kinds* of cause (and so the four causes would be related as species under a common genus, answering to a single definition of 'cause'), but he may also be making a stronger claim that 'cause' has four distinct *senses* or definitions, in which case there may be no common feature (besides the word) in virtue of which they are all causes.<sup>2</sup> We must therefore first establish whether the distinctions between the four causes are being drawn among senses or kinds.

In Aristotelian terms, we may begin by asking whether Aristotle considers the four causes to be *synonymous* or *homonymous* – that is, whether they are all called 'causes' in virtue of a single definition, or whether each type of cause involves a distinct definition of the term.<sup>3</sup> Answering that

---

draw metaphysically-grounded distinctions among different ways in which things determine each other, and have suggested that he be interpreted as making more epistemically grounded distinctions among ways in which *we* go about explaining them – ways we answer the question, 'Why?' (see, for example, Annas 1982, 319, and Vlastos 1969, 291–325). These alternatives suggest in turn the possibilities that the unity and inter-relation of the four causes might be grounded either in metaphysical or in epistemological features (these possibilities are discussed in section 4). The importance of the four causes for Aristotle's metaphysics and epistemology, his accounts of change, substance, nature, the soul, and biological phenomena, among others, is in any event undisputed, and I think the same questions of unity and interrelation arise on either understanding of the distinction.

<sup>2</sup> This particular question is raised by Gareth Matthews with great clarity as an instance of possible sense-kind confusion, at least on the part of commentators, between taking Aristotle to be saying that there are four kinds of cause and taking him to be saying that there are four senses of 'cause'. See Matthews 1972, 149–157. Sometimes we move back and forth between the locutions innocently, but where there is a genuine possibility of non-univocity we cannot do so without inconsistency, as Matthews shows. As he points out, most of Aristotle's commentators have not been sensitive to the distinction between senses and kinds, and use the two locutions as equivalents with regard to the four causes. Aristotle himself is deeply aware of the distinction, even if, as we shall see, he is not explicit as to how the four *aitiai* fare with respect to it.

<sup>3</sup> This distinction between synonymy and homonymy is not exhaustive, since a term might be primitive, i.e., indefinable. I do not think this is an attractive option for the causes, for reasons which will become clear. Thus, I think the key question with regard to '*aitia*' is one which is rightly assessed in terms of synonymy and homonymy, namely, whether the four causes bear their common term in virtue of one or several definitions. The relation between univocity and multivocity on the one hand, and syn-

question does not settle the matter, however: Aristotle himself frequently complains that his predecessors, especially Plato, have overlooked ambiguities in the senses of key philosophical terms such as 'being' or 'good', but it does not follow for him, or for us, that these different senses must be treated independently of one another. Some such terms may be very intimately related to one another (the paradigm illustration being the different senses of 'healthy', '*to hugieinon*', which Aristotle gives at *Topics* 106b35), while others share a name by mere linguistic accident (what he calls homonyms 'from chance', '*apo tuches*').<sup>4</sup>

Thus, in order to inquire as to the relations between the four causes, we must also establish what sort of relations we have reason to expect. If there is a univocal definition of 'cause' which applies to all four causal relations, we must say what that is, and what further properties differentiate the species of causation from one another. If there is no univocal definition of causation, however, then we must say whether and how these different senses belong together. There is pressure to give an account according to which they are not just interrelated but somehow unified as well, since Aristotle frames his account of scientific knowledge in terms of grasping causes – he appears to claim, that is, that causal relations *as such* are important for knowledge.<sup>5</sup> If 'cause' is non-univocal, then this account is threatened with incoherence unless we either provide some source of unity for its various senses or revise

---

onymy and homonymy on the other, is complex and not always clear. Some of the issues raised by the distinctions are discussed in what follows, especially section 2.

<sup>4</sup> *NE* 1096b26. Standard English examples are savings banks and river banks, or the geographical features and articles of clothing both called 'cape'. Aristotle's is '*kleis*', which refers both to keys and collarbones (*NE* 1129a30).

<sup>5</sup> *Phys.* 194b18; *Metaph.* 983a24; *Post. An.* 94a20; *Metaph.* 994b30, among others. This claim that we know when we know causes is in one sense not a fully general claim, since it presumes as background Aristotle's fourfold distinction at *Post. An.* 89b24 between knowledge of (1) the fact (*to hoti*), (2) the reason why (*to dioti*), (3) if it is (*ei esti*), and (4) what it is (*ti esti*), where the first and third come to the same, and ultimately, so do the second and fourth (in the sense that knowledge of what-something-is is at least one way of knowing the reason why). The relation between knowledge of the fact and of the reason Why is further discussed in *Post. An.* B, especially chapter 8, in which it is apparent that the search for *aitiai* yields knowledge of the reason Why with respect to a fact which is itself known in a different way, even though both kinds of knowledge may arise simultaneously (see also *Post. An.* 89b36f.). Aristotle's claims that knowledge is or involves grasping causal relations are thus most likely implicitly restricted to inquiry which seeks the reason Why, to the exclusion of knowledge whether something is the case. That scientific and philosophical knowledge seeks the *dia ti* may in turn be understood as the claim that grasping the reason Why is necessary and sufficient for such knowledge, or more modestly, only sufficient. He does not appear to consider whether it is also necessary, although he clearly demarcates classes of facts or entities which are important for scientific and philosophical knowledge but are not subject to causal inquiry, i.e., first principles and simples, which themselves have no causes, but must be understood in a different way (*Post. An.* 93a5, 93b21f.; *Metaph.* 1041b9).

our understanding of Aristotle's claim that knowledge consists in grasping causal relations.<sup>6</sup> That is, if the four causes correspond to distinct *senses* of '*aitia*' rather than kinds of cause, Aristotle needs a justification for taking them together as a theoretical unit upon which to base his conception of scientific knowledge. We might therefore distinguish two questions: (1) How do the four causes relate to one another? (2) How are they unified, if at all? The first could be given a detailed answer even if the second is given a negative or deflationary one; if I am correct, however, there is pressure on Aristotle to give an account of the relations among the four causes according to which they are unified in a non-trivial way. Unfortunately, although Aristotle frequently discusses and explicates the modes of causality he has distinguished, and occasionally discusses the four causes as a unit, he does not explicitly address the question of their interrelation, or whether the distinctions he has drawn are among senses or kinds.<sup>7</sup>

In what follows, I offer an account of the interrelation of the four causes on Aristotle's behalf. There are good textual and philosophical reasons, I argue, in favor of treating the four causes as homonymous, and so as not susceptible to univocal definition. On the other hand, even if they do not fall under a common genus, there is ample evidence that they are closely connected with one another. Aristotle's texts thus indicate both genuine diversity and deep connection between the causes, but they do not give an account of the nature of the distinction or the connection. Nonetheless, I think there is an attractive, and deeply Aristotelian, way of accommodating both of these features of the causes, and it may be discerned by looking carefully at Aristotle's discussions of causation and explanation. I offer an interpretation according to which the definitions of the four causes exhibit what is often called *pros hen* homonymy or focal connection: on this interpretation, one sense of '*aitia*', namely the formal cause, is taken as the fun-

<sup>6</sup> Similarly, if we were to define 'function' as a mapping which takes one value and yields another, since 'value' is ambiguous between something like a quantity and something like a moral principle (as when we talk, for example, of a society's values), we would have either to exclude one of the senses from the definition or make it clear why both belonged.

<sup>7</sup> The main texts in which Aristotle treats the four causes together are given and discussed below. Aristotle does occasionally remark that two or more causes may be, in some instances, "one" (usually the formal and final cause; see, e.g., *Phys.* 198a24, 199a30). It is important to note, however, that these remarks do not address the question I am raising. The claim that the form or essence of a human being is in some sense the same as its final cause is indeed a striking and important claim, but it is not fully general (not all formal and final causes are "one"), nor is it a claim about the nature of these causal relations as such. That is, Aristotle is not maintaining that the relationships of formal and final causation are the same thing, even in these instances, or that they always coincide the way they do in the case of the forms of substances like animals. Rather, these remarks are directed at the occupants of those causal roles and their interrelation (in certain special cases), and do not directly pertain to the natures of the roles themselves.

damental one, in terms of which the others are to be understood. Formal causation and its relations to the other causes are in turn to be explained, perhaps surprisingly, in terms of the nature of predication, as Aristotle conceives it. Thus, although Aristotle does not discuss it explicitly, there is a distinctively Aristotelian structure by which we can account for the diversity and the unity of the four causes, one which accords well with what he does say about them, and which does justice to some of his deepest metaphysical and epistemological commitments.

## 2. *The four causes: homonymy and synonymy*

A term's meaning may be in some sense *multiple* without its being non-univocal or homonymous: 'vehicle' refers to small passenger cars and to large trucks, but both classes fall under a single sense of the term. Likewise, the multiplicity of causes does not straightaway entail the homonymy of 'cause' or '*aitia*'. We must look, then, at the nature of the multiplicity involved as Aristotle conceives it, and examine whether there are strong enough reasons to view this multiplicity as an instance of homonymy.

Aristotle defines homonymy and synonymy at the start of the *Categories*:

Things are called homonymous of which the name alone is common, while the account (*logos*) of their being according to the name is different, the way both a man and a drawing are animals [...] But things are called synonymous of which the name is common and the account according to the name is the same, the way both man and ox are animals.<sup>8</sup> (1a1–8)

Roughly, then, if two things share a name and the name applies to them under the same definition (*logos*), they are synonyms; if the definitions in virtue of which they share the name are different, they are homonyms.<sup>9</sup>

Aristotle is unfortunately not explicit as to how the four causes fare with respect to this distinction. The *locus classicus* for the exposition of the four causal theory is in *Physics* II 3, where the theory is introduced as a natural continuation of the discussion in the last two chapters of the scope of natural science, and the importance to it of the notions of form and matter in particular. Here, however, the theory is not argued for or defended; it is simply presented as follows:

<sup>8</sup> All translations are my own.

<sup>9</sup> See Irwin 1981, 523–544, and Shields 1999, ch. 3. See below for more detailed discussion of the nature and kinds of homonymy. I follow Irwin and Shields in taking homonymy and synonymy to be exclusive and exhaustive for definable terms, which entails that things are homonymous whenever their definitions are not exactly the same – in other words, homonyms will include sets of terms whose definitions are partially identical, as well as those whose definitions have nothing in common.

Having made these distinctions, we must examine the causes: both what sort they are and how many in number [...] So then, one way of calling something a cause is that out of which something comes to be and which persists, such as the bronze of the statue and the silver of the cup, and the genera of these things; another is the form or paradigm – this is the account of the essence – and its genera (such the ratio two-to-one of the octave, and number in general), as well as the parts in the account. Further, there is the primary source of the change or rest: for example, the person who has deliberated is a cause, and the father of the son, and in general the maker of what is made and what changes something of what it changes.<sup>10</sup> Further, as the goal; this is that for the sake of which, for example health of someone going for a walk: ‘Why does he walk?’ ‘So that he may be healthy’, we say, and having spoken thus we suppose ourselves to have given the cause. (*Phys.* 194b23–35)

This passage does not state why or how we ought to think there are four causes; we are simply given a list, and the key claim that introduces the multiplicity of causal relations does no more than that. It is assumed from the start that investigating causes involves investigating their number, but we are not told whether the plurality is of kinds or senses.<sup>11</sup>

Nor is Aristotle more explicit elsewhere. Causes (*aitiai*) are not one of his examples of homonymy where he treats the latter directly in the *Categories* (1a1–6) or in the *Topics* (I 15–16). Nor does he say explicitly that the causes are homonymous, or that some things are only homonymously causes. Nevertheless, two features of Aristotle’s language strongly suggest that he thinks of the causes as homonymous:

(1) He expresses his claim that there are several modes of causation in many places as the claim that ‘*aitia*’ is “said in many ways” (“*pollachôs*

<sup>10</sup> It is difficult to capture the transitive sense of ‘*to metaballon*’, parallel with that of ‘*to poioun*’ just before it, without introducing explicitly causal-looking language, hence raising worries about circularity. Aristotle’s phrase here, however, is meant only to give the agentive side of the equation, as contrasted with the patient undergoing change. There may be a further worry as to whether Aristotle can or should offer a non-circular characterization of efficient causation, but that worry does not arise straightaway from Aristotle’s Greek.

<sup>11</sup> Several commentators have noted the lack of argumentation in Aristotle’s exposition of his theory, which has led some to claim that it must already be known to Aristotle’s intended audience (such as Ross 1936, 37), and others to conclude that Aristotle is here only surveying the uses of ‘*aitia*’ in ordinary Greek (see, e.g., Charlton 1992, 99). The discussion of ‘*aitia*’ in the so-called Philosophical Lexicon of *Metaphysics* Δ, elsewhere in which Aristotle gives precisely the kind of information we are seeking about ‘*aitia*’ for other terms such as ‘one’ (‘*hen*’), ‘actuality’ (‘*energeia*’), ‘necessary’ (‘*anagkê*’), and so on, is almost identical to that of the *Physics*. It appears to have been inserted either by Aristotle himself (which Ross offers as a possibility) or by the editors of the *Metaphysics* to replace missing text (as Asclepius 305.19 suggests; see Ross 1924, 292). Whatever the reason for Aristotle’s silence on the matter, if there are arguments on behalf of his pluralism about causes, or other clues about the type of pluralism involved, they are implicit, and require examination of Aristotle’s other discussions of causation and related issues.

*legeomenon*”), i.e., multivocal.<sup>12</sup> This is standard Aristotelian language for homonymy, and a common phrase Aristotle uses to indicate that a term is ambiguous in a philosophically significant way.<sup>13</sup>

The distinction between univocity and multivocity and that between synonyms and homonyms are not identical, however, and they may come apart, so the multivocity of ‘*aitia*’ does not straightaway entail homonymy. The former distinction pertains in the first instance to the number of senses a word has, whereas the latter pertains in the first instance, as we have seen, to things themselves and the definitions in virtue of which they share a common name.<sup>14</sup> Nevertheless, if the *logoi* in question are Aristotelian essences, then we might think that having multiple senses is both necessary and sufficient for applying to a variety of things under different definitions – i.e., that multivocity is necessary and sufficient for homonymy.<sup>15</sup>

Even if Aristotle does think there is multivocity without homonymy, however, it is unlikely that ‘*aitia*’ is an instance of it. Aristotle takes the trouble to discuss and explicate the various modes of causation, suggesting that there are non-obvious differences which must be pointed out and understood. Indeed, he views the failure to delineate these different notions as a crucial defect in the work of his philosophical predecessors, and his own work in that regard as a significant achievement.<sup>16</sup> This would be unnecessary, not to say overblown, if he felt he were simply uncovering unrecognized, vaguely stated, or surprising items to which the term applied in precisely the same sense. That is, the multivocity of ‘*aitia*’ very likely

<sup>12</sup> *Phys.* 195a4 and 195a29; *Metaph.* 983a26, 1013b4, 1052b4–8; *De An.* 415b9.

<sup>13</sup> E.g., *Top.* I 15; *Phys.* 227b3; *Metaph.* 1003b5. There are disputes about whether multivocity and homonymy are everywhere co-extensive for Aristotle (discussed below), but not about whether Aristotle often expresses claims about homonymy by saying that a given term is multivocal.

<sup>14</sup> This point is made clearly by Ackrill in his commentary on the *Categories* (Ackrill 1963, 71), and has been widely accepted since. This does not mean of course that homonymy has only to do with things, and Aristotle does sometimes speak of words (*onomata*) as homonyms (as with ‘contact’, ‘*haphês*’, at *Gen. Corr.* 322b29–32). It is nonetheless clear that, both canonically and in general, homonymy applies in the first instance to things rather than words.

<sup>15</sup> Shields (1999) and Owen (1960) argue, in that vein, that a term is homonymous if and only if it is *pollachôs legomenon*. Irwin (1981, 529), by contrast, entertains the possibility that Aristotle recognizes synonymous multivocals, i.e. terms which are *pollachôs legomena* but non-homonymous. This is not inconsistent with also maintaining that synonymy and homonymy are exhaustive and exclusive (see n. 9): Irwin’s claim, as I understand it, is that the synonymy/homonymy distinction and the univocal/multivocal distinctions do not track each other perfectly, so that a term may be “said in many ways” (i.e., multivocal) but nonetheless be applied in all cases under a single definition, hence synonymously. Shields and Owen, by contrast, maintain that a term and its referents are either univocal and synonymous, or multivocal and homonymous.

<sup>16</sup> *Metaph.* A, 984a16ff.



turns on the fact that the definition by which something counts as a formal cause is different from that by which something counts as an efficient cause. Hence, even if homonymy and multivocity come apart, there is good reason to suppose that this instance of multivocity is also a case of homonymy.<sup>17</sup>

(2) Aristotle refers to the different *aitiai* as ‘ways’ (*tropoi*).<sup>18</sup> This language is in line with his treatments of other homonyms, such as ‘nature’ (*phusis*) (*Metaph.* 1015b16), ‘one’ (*hen*) (1015b33, 1016a17), ‘substance’ (*ousia*) (1017b15), and ‘prior’ (*proteron*) (1018b30). Aristotle often, indeed, uses *tropon* to indicate that a word has or is being used in importantly different senses. He uses it, for example, to distinguish intrinsic from instrumental kinds of good (*NE* 1096b13); to distinguish the way pleasure completes an activity from the way in which the best object of perception completes the faculty of perception (1174b24); to distinguish the senses of ‘infinite’ (*to apeiron*, *Phys.* 204a3); the senses of ‘in’ (210a18); and different ways of dividing motion (234b21).<sup>19</sup> While not conclusive evidence by itself, the fact that Aristotle uses similar language for all of these cases, along with the qualificatory sense of *tropon* (which suggests that these ways of being a cause, or nature, or one, are less strongly similar than, say, kinds of car), adds strength to the view that Aristotle thought of the causes as homonymous.<sup>20</sup>

<sup>17</sup> For the purposes of this paper, I will therefore use ‘synonymous’ and ‘univocal’ interchangeably, and likewise ‘homonymous’ and ‘multivocal’, though I do not mean to discount these problematic cases. The point I wish to focus on is whether there are several definitions of ‘cause’, i.e., different properties or relations which all answer to the same term, but which are not related as species under a common genus; nothing turns on whether this also amounts to a difference in *meaning*, as has been disputed (see especially Irwin 1981, 533ff., and Shields 1999, ch. 3). In other words, my primary aim is to examine whether Aristotle’s fourfold distinction is a weak or strong one, and if strong, whether there is nonetheless a systematic connection or unity among the different causal roles or relations.

<sup>18</sup> *Phys.* 194b23, 195a27; *Metaph.* 996b5.

<sup>19</sup> With regard to the different senses of ‘being’ (*to on*), Aristotle also uses *tropon* with explicitly semantic import: ‘being’ is multivocal, he says, one sense being reserved for the different categories such as quantity, place, time, and anything else it might signify in this way (“*kai ei ti allo sēmainei ton tropon touton*”, 1026b1).

<sup>20</sup> There is a *prima facie* problem in discussing the possible ambiguity of ‘cause’, however, since we might think that ‘*aitia*’ is ambiguous or homonymous if and only if ‘Why’ (*dia ti*) or ‘The Why’ (*to dia ti*) is so as well. ‘*Dia ti*’ and ‘*aitia*’ form a complementary pairing, as we have seen: ‘*dia ti*’ as a question is answered by giving *aitiai*. When prefixed with a definite article, this pairing leads us to expect ‘*to dia ti*’ to be synonymous or at least interchangeable with ‘*aitia*’, and indeed Aristotle sometimes seems to use the phrase as a near synonym: at *Metaph.* 983a29, he states that the primary *dia ti* is a cause (*aition*) and principle (*archē*), and ‘*to dia ti*’ picks out the object of inquiry at *Metaph.* 1041a10. Further, Aristotle seems to imply an equivalence when he claims that the number of causes must be the one stated, since this is the number comprehended by (*perieilēphen*) *to dia ti* (*Phys.* 198a15). However, because all four *aitiai* form a complete answer to the question ‘Why’, we may therefore think that ‘*dia ti*’ cannot be homonymous (this point is made by Irwin 1981, 537). On the other hand, ‘*to dia ti*’ (unlike ‘*aitia*’ or ‘*aition*’) is usually accompanied by either a form of



There are further, philosophical reasons for Aristotle to take the four causes to be homonymous. Aristotle's claims that the goal of inquiry is to grasp *aitiai*, and that these four relations are the ones by reference to which we give the right type of answer to scientific questions, are not obvious on their face, nor are they trivial. Aristotle ought to say, if possible, what it is about these *aitiai* that makes them scientifically significant – he ought to be able to define the term.

There are, however, serious difficulties confronting any attempt to define '*aitia*' univocally. For one, the four kinds of relation picked out by Aristotle are simply so different that the terms of a proposed univocal definition risk losing all substantive content. This problem is well-illustrated by Moravcsik's attempt to treat the four causes as what he calls "generative factors"; he writes: "Using the term 'generate' in a very wide sense, premises generate conclusions, substances are generated from constituents, and animals can generate other animals".<sup>21</sup> It would be a very wide sense of 'generation' indeed which applied to the relation of the ratio 2:1 to the octave, or of health to after-dinner walks, which are Aristotle's own examples of the formal and the final cause, respectively (194b28 and 35). Such a sense divorces 'generation' from both its Aristotelian as well as its common usage, and in any event renders it no more effective than any other label, such as 'Aristotelian cause'.<sup>22</sup>

---

'*apodidōmi*' or '*anagō*' (see, e.g., *Metaph.* 983a28; *Phys.* 198a23, 198b5), and immediately after stating that the number of causes is four because this is the number of ways comprehended by the *dia ti*, Aristotle further glosses this as due to the fact that the *dia ti* is "referred to" ("*eis* [...] *anagetai*") the four kinds of thing named, i.e., essence, first mover, that for the sake of which, and matter (198b16–21), suggesting that the terms are not equivalent after all. Both '*dia ti*' and '*aitia*' seem, in fact, to have what we may call analyzed and unanalyzed uses. '*Dia ti*' in its most straightforward use may simply be a demand for explanation without presuming anything about the character of that explanation, and may in that sense be satisfied by an indefinite variety of answers, ranging from 'because something having property X came into contact with something having property Y' to 'because the Deity willed it thus'. That is, in a general or philosophically untutored way, there are indefinitely many ways of answering the question '*dia ti*', including false ones or ways a more sophisticated inquirer would reject as misguided; at 198b15, however, Aristotle clearly indicates that there are only four correct ways of giving the Why (and the use of the definite article may indicate that the use is the technical one). Similarly, '*aitia*' may well have a relatively non-technical sense, insofar as Aristotle wishes to say of his predecessors that they were inquiring into *aitiai* despite having no philosophical understanding of them as such (see, e.g., *Metaph.* 984b4, where Parmenides is characterized as having supposed there to be two causes). If *aitia* and *dia ti* have both technical and non-technical uses, then we need not suppose that the homonymy or synonymy of one entails or excludes homonymy of the other.

<sup>21</sup> Moravcsik 1975, 626.

<sup>22</sup> Other attempts to provide a univocal account of '*aitia*' tend to be more reductive in character. Robin (1910) attempts to show that all four causes are in fact manifestations of the formal cause or essence; Irwin (1988) argues that based on some of Ar-

There are, nonetheless, a variety of descriptions or characterizations which might correctly capture the extension of *'aitia'*, such as 'things cited as answers to Why-questions', or 'explanatory factors', and which pick out important features of the four causes. Our ability to give such a phrase or description, however, even if it is informative, does not entail univocity, unless we are in fact giving the definition in virtue of which something bears the relevant term.<sup>23</sup>

Moreover, these descriptions tend to appeal to terms or phrases which are in just as much need of clarification as *'aitia'* and 'cause', such as 'making it the case that', 'determining', 'is responsible for', 'explains', and so on. Such phrases are of little use, and simply postpone the question. In other words, descriptions such as 'being a genuine explanation of' or 'being responsible for' may give a false impression of univocity: whether or not they do would appear to depend, in fact, on whether the four relations Aristotle describes are themselves susceptible to univocal definition. Hence, a full evaluation of whether the four causes are homonymous will turn on whether it is possible to give such an account.<sup>24</sup>

It is thus well worth taking seriously the idea that Aristotle's claim that *aitia* is 'said in many ways' should be understood as a claim about homonymy, and that his characterizations of them in the passage cited from the *Physics* give an array of relations which may not be univocally defined. Aristotle's notion of homonymy is quite complex, however, and he is prepared to call a range of things homonymous in non-equivalent ways; we must therefore examine what is entailed by the homonymy of cause.<sup>25</sup>

The four *aitiai* cannot, for example, be chance homonyms: since they are all middle terms of demonstrative syllogisms (*Post. An.* 94a23) and answers to Why-questions, as we have seen; they share more than just the name

---

istotle's ways of speaking about causality and some general desiderata for a theory of causation, we may reduce the four causes on his behalf to types of efficient cause. The latter approach assumes, explicitly, that certain modern desiderata ought to be respected (e.g., that all causation should be analyzed as a relation between actual, discrete particulars such as events); whether it is the right approach therefore depends on the validity of those desiderata. While I consider Robin's reductive thesis to be too strong, I shall argue (see especially section 6) that his core suggestion regarding the primacy of the formal cause is, in fact, correct.

<sup>23</sup> Aristotle is clear on this point with regard to the nature of soul. He offers what appears to be a common account (*logos koinos*) under which all the kinds of soul fall, namely the characterization of soul in *De An.* II 2 as the first actuality of a natural body which has life in potentiality (412a27). Nevertheless, he denies that they are synonyms on the grounds that such an account fails to describe what is peculiar (*idion*) to each kind of soul (414b22–25).

<sup>24</sup> I am not arguing, then, that one *must* take the four causes to be homonymous, only that there is good reason to do so. I think such an account is in fact promising, and has further advantages over univocal accounts, which are given in the concluding section.

<sup>25</sup> For detailed treatments of the nature and varieties of homonymy, see Irwin 1981; Shields 1999; and Ward 2008.

'*aitia*'. Nor could any of the four causes be what are sometimes called spurious homonyms, the way the hands of cadavers are not, properly speaking, genuine hands, even though it is no accident that we tend to refer to them as such.<sup>26</sup> Aristotle distinguishes between genuine (*kath'hauto*) causes and non-genuine ones such as chance (*tuchê*) and luck (*to automaton*), and the four causes (along with their more precise permutations) all fall on the genuine side of the distinction.<sup>27</sup> Thus, none of the four causes is a cause *merely* homonymously.

The different senses of '*aitia*' would thus still be intimately connected with one another on the supposition that they fall short of synonymy.<sup>28</sup> What, then, is the nature of that connection?

There are a variety of ways in which related senses of a term might be connected with one another, as Aristotle recognizes, but one of the most philosophically intriguing, and one which provides a very strong degree of unity, is what is known as *pros hen* homonymy.<sup>29</sup> In cases like 'health', and more controversially, 'being', a single sense forms the primary or core sense, and the others are organized around and in reference to it. I shall argue that when we consider Aristotle's accounts of the causes, there is good textual and philosophical evidence to suggest that they are connected in this *pros hen* manner, with one sense – that of formal causation – functioning as the basic sense in terms of which the others are defined.<sup>30</sup>

<sup>26</sup> See, e.g., *De An.* 412b17–22; *Pol.* 1253a19–25; *Meteor.* 390a10–15.

<sup>27</sup> See especially *Phys.* 195a32f.

<sup>28</sup> In describing the different kinds of non-accidentally related homonyms, Irwin (1981) speaks of "connected" homonyms, while Shields (1999) speaks of "associated" homonyms.

<sup>29</sup> See especially *Metaph.* Γ 2. Owen (1960) uses the term 'focal meaning' for these cases, which is perhaps unfortunate for its linguistic connection; Irwin (1981) speaks of focal connection (531) and connected homonyms (526), which rightly avoids linguistic connotation. Shields (1999) speaks of associated and core-dependent homonyms, the latter of which captures the notion of asymmetry important to the most philosophically interesting cases of homonymy. Aristotle sometimes describes them as *pros hen*, as at *Metaph.* 1043a37. Not all philosophically interesting homonyms are related *pros hen*, it would seem. Aristotle suggests, for example, that the kinds of soul might turn out to have wholly discrete definitions which form an ordered series; he likens them to the basic kinds of geometrical figure, whose accounts (*logoi*), he claims, are distinct one from the other, and not interrelated, but rather form an ordered set (*De An.* 414b28–32).

<sup>30</sup> I am not of course claiming that this is Aristotle's considered view on the matter, since he does not address it directly. Nonetheless, these issues of definition, univocity, and interrelation are ones of which he is acutely sensitive and which he does much to develop, so we cannot accuse him of being unaware of the sorts of questions raised by his pluralism about *aitiai*. Thus, even though he does not discuss their unity and interrelation explicitly, it is plausible that Aristotle's treatment of the causes is amenable to a more rigorous formulation in terms of distinctions of which he is already well aware. My goal is to give an account which accords best with his philosophical commitments and methods, and to give it in as rigorous a manner as the problem allows.

### 3. The definitions of the individual causes

The question of the interrelation of the four causes thus turns on whether we can discern a structure among their respective accounts. Since Aristotle does not consider the matter explicitly, we ought naturally to look at the discussions of ‘*aitia*’ in general as well as the specific accounts of each of the causes, for which the main texts are *Physics* II 3 and *Posterior Analytics* B 11.<sup>31</sup> Whether these discussions should be treated as offering definitions is of course uncertain – that depends in part on whether the causes are homonymous. Nonetheless, these texts offer the most general discussions and the most explicit characterizations of the four causes, so they are the natural place to begin examining the nature of their difference and their interrelation.

Aristotle’s introduction of each of the causes in *Physics* II 3 is bipartite, consisting of a clause or phrase giving what appears to be a general characterization of the kind of cause, followed by salient examples:

(MC) “That out of which something comes to be and which persists” (“*to ex hou gignetai ti enuparchontos*”, 194b23)

(FC) “The form or paradigm; this is the account [*logos*] of the essence” (“*to eidos kai to paradeigma, touto d’estin ho logos ho tou ti ên einai*”, 194b26)

(EC) “The primary source of the change or rest” (“*hothen hê archê tês metabolês hê prôtê ê tês êremêseôs*”, 194b29)

(TC) “The goal; this is that for the sake of which” (“*to telos; touto d’estin to hou heneka*”, 194b32)<sup>32</sup>

One of the causes is given its characterization by means of a relative clause (“that out of which something comes to be and persists”); two, the formal and efficient causes, are characterized by means of nominal expressions (“the form or paradigm”). The description of the final cause uses both kinds of expression. These descriptions do not make explicit reference to one another, nor is one singled out – implicitly or explicitly – as primary.

We might hope that these accounts, while not explicitly referring to each other or overlapping in content, nonetheless share or imply a common background which could serve to clarify their connection. Thus, we might

<sup>31</sup> *Phys.* II 7 is an important chapter for understanding the causes as well, since Aristotle there sums up the importance of the four causes for the study of nature; however, he does not attempt there to clarify the nature of the four causal roles themselves any further than he already has in *Phys.* II 3, either through additional explication or examples.

<sup>32</sup> I follow Fine’s (1984) abbreviations for the kinds of cause, and use (TC) to indicate the teleological, or final cause.

argue that the causes are related to one another by being unified around a core case, or rather, a certain conception of the “effect” they all have.<sup>33</sup>

Unfortunately, the notion of a unity of effect for the causes is not borne out by Aristotle's examples, which are: for (MC), the copper of a statue and the silver of a cup; for (FC), the ratio two-to-one as the account of the octave; for (EC), the man who has deliberated and the father of the son; for (TC), health, for the sake of which someone walks. While Aristotle certainly thinks there are things such as artefacts and living beings which have all four causes,<sup>34</sup> and so are the result of their joint expression, his examples here illustrating the modes of causation are not coordinated in this way.

Nor does the relevant passage from *Posterior Analytics* B 11 (94a21–23) provide more of a clue. There, the descriptions of the causes are even more compressed, and they make no clear reference to effects:

(MC) “Those things which, being the case, necessarily this is” (“*to tinôn ontôn anagkê tout' einai*”)<sup>35</sup>

(FC) “The what it was to be” (“*to ti ên einai*”)<sup>36</sup>

(EC) “That which moved first” or “The primary mover” (“*hê ti prôton ekinêse*”)

(TC) “That for the sake of which” (“*to tinos heneka*”)

<sup>33</sup> Aristotle does not have a correlative term for ‘*aitia*’ which would correspond to ‘effect’; I use it here simply to indicate whatever it is that the causes in fact cause, i.e., an explanandum. The most highly developed recent version of such extrinsic unification around a single effect is that of Graham (1987), who views the causes as unified by a model based on the paradigm example of the statue – that is, the paradigm case of an effect these causes have. More recently, Hennig (2009) presents the four-causal schema as deriving from a pair of related questions about the sources and endpoints of natural changes and their subjects. The four causes would thus be unified (again extrinsically) insofar as they are all pertinent to questions raised by a certain conception of natural change.

<sup>34</sup> See, for example, the four causes of ‘man’ listed at *Metaph.* 1044a34f. Thus, the standard illustration by which Aristotle's four causes are often presented is, as Sprague (1968) observes, somewhat misleading, though not un-Aristotelian; despite the importance of substances and the artefacts as examples and illustrations, it is important for our purposes to note that, in the passages which constitute his direct presentation of the four causes, Aristotle does not choose to illustrate them by means of a single example.

<sup>35</sup> The phrase is unavoidably awkward and obscure, and notoriously problematic, as it is unclear how this description could accurately reflect the same notion of material causation discussed elsewhere in the corpus. It is certainly none of the other causes, however, and if this cause is truly different from the material cause described in the *Physics*, it is a kind of cause Aristotle abandons in his non-logical treatises. It seems more likely that his description here is, for whatever reason, an awkward expression of the material cause, perhaps meant to suit the geometrical example he has in mind, or to evoke hypothetical necessity, or perhaps recalling the sense in which the premises are the matter of the conclusion of a syllogism. See Ross 1945, 639, and Barnes 1994, 226.

<sup>36</sup> A literal rendering of the standard and technical Aristotelian phrase, usually and best translated as ‘essence’.

Aristotle's examples of effects here again remain largely disjoint, with one intriguing exception. For (MC), the effect is that the angle inscribed in the semicircle is a right angle (94a28); for (EC), the Persian war; for (TC), health again, as well as the existence of a house. Aristotle does not give a separate example of (FC) in this passage, however, but instead concludes the description of the geometrical example of (MC) with a puzzling remark: after showing why *half of two rights* is the middle term which is explanatory (*aition*) of *right* belonging to the angle of the semicircle, he says, "And this is the same as the essence, for this is what the account signifies. But then, the middle term has also been shown to be explanatory of the essence." (94a35f.)<sup>37</sup>

The passage is intriguing for current purposes insofar as it suggests a possibility for unifying the material and formal causes by way of their common effect. It is not clear from Aristotle's example whether he means his point about the formal cause to be general, especially since, as noted above (n. 35), we need to do some work on his behalf in order to align the cause described here as "Those things which, being the case, necessarily this is" with the material cause exemplified by metals in the *Physics*. At the same time, we ought to take seriously the possibility that the effects of the material and formal cause are identical in a principled way, especially given that, between the two most important passages in which Aristotle sets out to describe his theory of *aitiai*, he does not, in fact, give a clear example of formal causation for a straightforward case – the only explicit example being the ratio 2:1 as the essence of the octave.

Despite this possibility of unifying the causes by way of a single effect, however, there is no clear unity that extends to (EC) and (TC) in Aristotle's examples. The accounts Aristotle offers of the individual causes do not, therefore, display the kind of association or focal connection displayed by the descriptions of the various kinds of health or being, as they do not overlap in any apparent way, nor is one sense suggested to be more basic than the others. Nor, however, is another kind of structure suggested by the text. To the extent that Aristotle does have an account of their connection, therefore, we must rather look to those of his discussions which treat the notion of *aitia* in a general way.

<sup>37</sup> "Touto de tauton esti tōi ti ên einai, tōi touto sēmainein ton logon. Alla mēn kai tou ti ên einai aition dedeiktai to meson." I agree with Barnes's reading of this difficult sentence with the majority of the manuscripts against Ross's "to ti ên einai aition dedeiktai to meson (on)". The text and its meaning remain obscure nonetheless, and I have tried to render it neutrally. In particular, I have used 'explanatory' to translate '*aition*' here, since using 'cause' would seem to require adding more precision than the text in fact has (for example, whether the middle term is simply *a* cause or *the* cause).

#### 4. The general accounts of 'cause'

##### 4.1. The epistemological criterion

The four causes are meant to be answers to Why-questions, as laid out in the canonical introduction of the causes at *Physics* 194b18: "We don't suppose ourselves to know until we have grasped the 'Why' of each thing (and this is to grasp its primary cause) [...]" ("eidenai de ou proteron oiometha hekaston prin an labômen to dia ti peri hekaston (touto d'esti to labeîn tên prôtên aitian) [...]"). Aristotle does not further explain there just how causes are supposed to answer Why-questions, but as we have seen, this feature ought to be susceptible to further clarification. In the *Posterior Analytics*, by contrast, he is explicit that the four causes are all indicated by the middle terms of demonstrative syllogisms (94a23). By treating causes in terms of their relation to demonstrative syllogisms, Aristotle thus appears to be giving a systematic rendering of this core feature, showing precisely how causes are implicated in answers to the question 'Why?'.

The link between being an answer to a Why-question and being a middle term is detailed in *Posterior Analytics* B 11. By way of illustrating the causal role of the middle term, Aristotle answers several Why-questions by deriving the *explanandum* through a syllogism whose middle term is (he hopes) obviously the cause: Why did the Persian war befall the Athenians? (94a37) Because they were the first attackers. Why does he walk about? Why is there a house? (94b9)

Thus, a good answer to a Why-question allows us to derive the explanandum as the conclusion of a demonstrative syllogism. As such, causes are answers to Why-questions, not in the sense that causal statements have the logical form of Because-answers, or that they provide psychological satisfaction to inquirers, but in the sense that causes somehow provide the proper link between the premises of demonstrative syllogisms.

It is important, however, to be clear about the priority between causes and their role in demonstration – that is, about whether something is a cause in virtue of being a middle term, or is a middle term in virtue of being a cause. Aristotle of course does not mean to say that the linguistic terms composing syllogisms are causes, or that all middle terms pick out causes. Since 'term' ('*horos*') is ambiguous between linguistic entities and their referents, however, Aristotle may wish to assert a kind of conceptual connection between being a cause, on one hand, and being a property by which we may syllogistically derive that some A belongs to some C.

Consider, however, the non-explanatory syllogism given at *Posterior Analytics* 78a28f.: What does not twinkle is near; Planets do not twinkle; so the planets are near. The syllogism is non-explanatory since the cause of the planets' not twinkling is their nearness, and not vice



versa.<sup>38</sup> Moreover, that is the only reason the syllogism is non-explanatory – there are no formal grounds for ruling out ‘not-twinkling’ as an inappropriate middle term for demonstration. Yet even if the syllogism gets the causal priority wrong, one can still carry out valid deductive reasoning from the coextension of nearness and not-twinkling; indeed, it is Aristotle’s point in this passage that deductions to the effect that something is the case are not necessarily deductions of why something is the case, the difference being whether or not the middle term happens to be the cause.

So, on the assumption that all causes are appropriate middle terms for demonstrations, they are such in virtue of being causes, not vice versa. So, while their role as middle terms of demonstrative syllogisms is no doubt important to our understanding of the causes, that role is posterior to their being causes, and hence cannot serve as the ground of their interrelation.

#### 4.2. Causes as *archai*

The other important discussion in which Aristotle offers a feature common to all the causes is in *Metaphysics* Δ, not in the section on cause, but that on principle (*archê*).

In this section Aristotle writes: “Causes are spoken of in an equal number of ways [as principles are]; for all of the causes are principles.” (“*isachôs de kai ta aitia legetai; panta gar ta aitia archai*”, 1013a16) This claim is followed straightaway by a general claim about principles: “It is common to all the principles then to be the first point [or primary thing] from which something is or comes to be or is known.” (“*pasôn men oun koinon tôn archôn to prôton einai hothen ê estin ê gignetai ê gignôsketai*”, 1013a18) It is thus possible that in discussing ‘principle’, Aristotle is at the same time offering an account, either in whole or in part, of ‘cause’. If so, this passage is significant insofar as it indicates ontological or metaphysical features as well as the already familiar epistemological one – besides being principles by which things are known, causes are principles by which things exist or come to be.

What Aristotle means by these claims, however, is not apparent on the surface: on the one hand, ‘*archê*’ is one of the most difficult of Aristotelian terms, subject to a wide gamut of uses, and on the other, we cannot infer that there are as many kinds of cause as kinds of principle solely because the former are all examples of the latter. Indeed, the uses Aristotle has for ‘*archê*’ seem to outstrip his uses for ‘cause’. In particular, the first sense of ‘*archê*’, as that from which one would begin movement along some thing, e.g., the extremity of a line or road (1012a34), does not seem to correspond

<sup>38</sup> See Ross 1945, 552, and Barnes 1994, 156.

to any kind of cause.<sup>39</sup> Causes and principles are therefore probably not necessarily coextensive.<sup>40</sup>

Further, there is an interpretive issue regarding the range of phenomena which are meant to be metaphysical principles of existence and generation (*esti* and *gignetai*). The phrase “first point from which something is or comes to be” (“*to prôton [...] hōthen ê estin ê gignetai*”, 1013a18) may be interpreted narrowly to mean the existence or coming-into-being of something, most notably substance, or interpreted widely to mean, roughly, something's being the case and coming to pass. That is, under a narrow interpretation, causes are said to be principles of the existence and generation of things; under a wide one, they are principles of whatever occurs or is the case in the world, presumably including a wide range of ontological categories. From the passage itself we get no further direct evidence as to which interpretation is best.<sup>41</sup>

It is plausible, nonetheless, that Aristotle does mean to apply all three attributes to the causes: if so, they are principles of existence and generation or occurrence, as well as being principles of knowledge. Moreover, as we have seen, there appears to be an asymmetry between the ways in which causes may be principles. Even if being a cause of Y and being an answer to a Why-question pertaining to Y are coextensive, being an answer to a Why-question is posterior to bearing a causal relation as such (as we see in the case of demonstrative syllogisms). It is likely, then, that the ways in which causes are principles of existence and generation are prior to and entail their being principles of knowledge in some special way.

In sum, since the notion of being a principle outstrips the notion of cause, except insofar as there may be contexts in which Aristotle uses a restricted

<sup>39</sup> Compare also the use at *Metaph.* 1003b6: the different ways in which being is spoken of are all said to relate to a single principle (*pros mian archên*); this sense of ‘principle’ seems again to lie outside the range of ‘cause’.

<sup>40</sup> At least, not strictly speaking. The interplay between ‘*archê*’ and ‘*aitia*’ is a complex one. Often Aristotle appears to use ‘*archê*’ as a synonym for cause, as at *Metaph.* 984a26, and in *Metaph.* Γ he claims that, like ‘one’ and ‘being’ might be, they are “the same and one in nature” (“*tauton kai mia phusis*”), but differently defined (1003b24). Further, with the exception of the first meaning given in Δ 1, the senses of ‘*archê*’ Aristotle recognizes are all rather closely allied to notions which are importantly causal for him, such as the part where generation starts (1013a4), that whose choices move things (1013a10), and the starting point for knowledge (1013a14). Nonetheless, Aristotle has good reason for keeping ‘*archê*’ and ‘*aitia*’ distinct insofar he seems to wish to appeal to ‘*archê*’ as something more basic than ‘*aitia*’, in terms of which he can explicate the latter: the use of the word to define the efficient cause at *Phys.* 194b29, as well as frequent use of the phrase ‘*hê archê tês kinêseôs*’ (e.g., at *Metaph.* 984a27, 994a5) to indicate it seems to demand precisely this. It is likely, then, that ‘*archê*’ has both wide and narrow senses, where the latter is co-extensive with ‘*aitia*’. See also Ross 1924, 290.

<sup>41</sup> The issue is addressed indirectly in the next section: if we have an independent idea of what the effects of causes are, we can bring that information to bear on the question of what causes are principles of.

meaning of ‘*archê*’ which is synonymous with ‘*aitia*’, the discussion of causes as principles does not explain how the causes are related to one another, or why they should comprise a single, complete theory. It is helpful, however, to see that causes are principles in at least two distinct senses (*viz.* of existence or generation and knowledge) between which there is an important asymmetry. What we require, in fact, is an independent and more precise account of what it is for a cause to be a principle of existence and coming-into-being, and how being an ontological principle of this sort might entail being an epistemological principle as well.

### 5. *Causation and predication*

Thus far, none of the texts in which Aristotle discusses the causes provides a clear statement as to the manner in which they relate to one another. The general accounts of the causes provide two pieces of useful information, however: (a) causes are answers to Why-questions; (b) causes are principles of generation or existence, and their manner of being principles is such that they are thereby also epistemological principles. It is reasonable to suppose that it is as answers to Why-questions that causes are epistemological principles; thus, what is wanted is an account of how causes are ontological principles such that citing them answers Why-questions.

A way forward is then suggested, perhaps surprisingly, by a passage from *Metaphysics* Z 17. It may initially seem unlikely that we should find clarity concerning causation in one of the most difficult chapters of the discussion of substance in the *Metaphysics*. In this chapter, however, Aristotle has begun a fresh start in the inquiry, leading up to the identification of substance and form at 1041b8. This attempt to grapple with the nature of substance begins with a reminder that substance is a principle and cause (or, more likely, a principle in the sense of cause: “*hê ousia archê kai aitia tis estin*”, 1041a9). Because substance is a cause, it is an answer to Why-questions, and what follows this reminder is an elucidation of the nature of Why-questions themselves. More specifically, Aristotle here clarifies the nature of explananda:

The ‘Why’ is always sought in this way: why one thing belongs to some other. For to seek why the musical man is a musical man is either to inquire, as was said, why the man is musical, or something else. Now, the question in no way is inquiring why something is itself, for the fact and the existence of the thing must be clear – I mean for example that the moon is eclipsed – while the fact that something is itself is the one account and the single cause regarding all questions as to why the man is man or the musical musical [...] But one may inquire why man [or the man] is an animal of such a sort. This, then, is clear: we do not seek why what is a man is a man; one seeks therefore in respect of something why it belongs (that it does belong must be clear; for if it is not thus, nothing is sought). For example: Why does it thunder? Why is sound produced in the clouds? For thus one thing’s belonging to another is what is inquired into. (1041a10–26)

If this claim is meant to be fully general, it suggests that Why-questions have a single form: they all ask why one thing belongs to (*huparchei*) another – why some A belongs to some B. Less cryptically, we can take Aristotle as here referring to predication, or if we are understanding predication linguistically, to its metaphysical analogue.<sup>42</sup> That is, to ask why one thing belongs to another is, for Aristotle, to ask why something is predicated of (“belongs to”) a certain subject. The examples which follow appear to confirm that he thinks of this relation between Why-questions and predication in a general way. He would analyze each of the following as questions why A belongs to B: Why is the man musical? Why is man a certain kind of animal? Why does it thunder? Why are these bricks and stones a house?<sup>43</sup>

The examples given include not just instances of a substantial form making something the thing it is, but also of events such as thunder and of the inherence of accidents such as being musical. Thus, although *Metaphysics Z* is an inquiry into the nature of substance, Aristotle relates that inquiry to a point about causation and explanation which is itself fully general. Surprisingly, then, Aristotle's inquiry into substance in *Z* 17 may provide what is missing from the canonical discussion of the four causes in the *Physics*, namely, an understanding of why these four distinct relations belong together, and how they relate to one another.

Aristotle's suggestion, then, is that explananda all have the same metaphysical structure: what is explained when Why-questions are answered is the belonging (i.e., the inherence) of a predicable to a subject. The belonging of a predicable to a subject might thus be the effect of which causes are causes, and by the same token, that of which they are principles. In theory, of course, such a question as ‘Why does A belong to B?’ could be answered in a variety of ways according to different metaphysical frameworks. We might say that A belongs to B in virtue of B's participation in the Platonic form A, or in virtue of B's being a member of the class ‘A’ (if we are class nominalists, for example), or in virtue of the inherence of an *in re* universal A in B, or, to the extent it differs, the presence of a form A in suitable matter B. Given Aristotle's own metaphysical framework, the structure of explananda would be analyzed in terms of a predicable – either a substantial form or an item in a non-substance category – inhering in its proper subject.

<sup>42</sup> This reading of the passage thus takes ‘*huparchein*’ in its technical sense, used by Aristotle throughout the *Prior* and *Posterior Analytics*, for example, to indicate the relationship between a predicable and the subject in which it inheres. That this is the technical usage is especially suggested by the examples: see in particular *Post. An.* 93a39–b14, which also uses the examples of eclipses and thunder, and *Phys.* I 7, which uses the example of musicality to illustrate the theory of hylomorphic change being introduced.

<sup>43</sup> All cited between 1041a10–27.

That is, on Aristotle's view, to be an explanandum is to be a complex entity consisting of a predicable and the subject of which it is predicated.<sup>44</sup>

The passage from Z 17 and this interpretation of it are in conflict with the discussion in B 11 of the *Posterior Analytics* at first blush. The examples of Why-questions Aristotle gives there simply do not fit this model: "Why did the Persian War come upon the Athenians?" "Why does he walk about?" "Why is there a house?" (94a38–b9). Indeed, more generally, we can ask perfectly coherent and, it would seem, answerable Why-questions which do not have the form "Why does A belong to B?" – that is, which do not appear to ask why a certain predicate belongs to a subject. For example: Why are there houses/stars/zebras? Why is it raining? Why does being extended follow from being material?

However, once these questions are answered in Aristotle's way, the explanandum has been transformed from the way in which it appears in the Why-question to a conclusion of a syllogism of the form 'A belongs to B'. Thus, we should read the claim that the 'Why' is always asked with regard to why one thing belongs to another as in fact claiming:

(ES) Explananda all have a common subject-predicate structure.<sup>45</sup>

(ES) may be read in two ways, as related to Why-questions. We might interpret it as a semantic claim that:

(ES1) Why-questions have a hidden logical form, which, when analyzed, yields the form 'Why does A belong to B?'

Thus, a speaker asking 'Why does he walk about?' is properly understood as asking 'Why does walking belong to him?', much the way a Russellian analysis of definite descriptions yields their underlying logical form.

Alternatively, we may read (ES) as the claim that:

(ES2) An entity must have a certain structure in order to be an explanandum. An explanandum cannot be simple, but must rather be analyzable as having parts corresponding to subject and predicate, such that the latter belongs to the former.

<sup>44</sup> There are long-standing difficulties and scholarly controversies over whether and to what extent Aristotle's treatment of substance as form which is predicated of matter is consistent with his other remarks (e.g., in the *Categories*) about substance and predication (see, e.g., Lewis 1991 and Wedin 2000). The view proposed here is largely independent of those controversies, since it concerns substantial form only insofar as Aristotle considers it as a cause which is predicated, though of course the ultimate metaphysical details will depend on how we understand the relationship between form, matter, and substance, on one hand, and the nature of predication on the other.

<sup>45</sup> The link between Why-questions and the notion of predication is also found in the *Posterior Analytics*, especially in chapter B 2. There, however, inquiries are divided into those that seek an explanation for, on the one hand, why something is one of the things predicated of it, and on the other, for why something is *simpliciter* (i.e., why it exists; see especially 90a31–34). Significantly, in the passage at *Metaph. Z*, 1041a10 discussed above, the claim is not restricted.

The textual and philosophical evidence all weigh heavily for (ES2). There is no real semantic mystery regarding Why-questions analogous to the problem of truth-conditions for sentences involving definite descriptions which would justify reading (ES1). To ask 'Why ...?' is semantically unmysterious, even if the question may be asked about entities ranging from states of affairs to actions to desires, and answered by reference to a similarly wide array of entities.

Moreover, Aristotle's reasoning in the passage from Z 17 makes it clear that explananda have the structure they do in virtue of metaphysical, not semantic facts. For he argues that to ask why the musical man is a musical man is either (1) to ask why the man is musical, or (2) something else. But it cannot be, he continues, (2a) asking whether it is the case, since it must already be known to be the case if we are asking about it; nor can it be (2b) asking why something is itself, since everything is itself. So it must be asking (1), understood as the question why something is predicable of something (1041a10f.). (2a) is ruled out on pragmatic, not semantic grounds – it is not meaningless to ask whether something is the case when it is already known to be the case, just pointless. (2b) is ruled out on metaphysical and epistemological grounds: self-identity is true of everything, and so is not a subject of inquiry in regards to any one thing more than another. If there is an explanation of self-identity, it is the same one for everything.

Finally, in conclusion to his argument that form is substance, Aristotle asserts that no inquiry nor teaching of the sort he has just been referring to is possible with regard to simples (*Metaph.* 1041b9–11), since the nature of simples is such that they cannot be analyzed as something predicated of something else.<sup>46</sup> Thus, again, the scope and content of Why-questions is determined by the metaphysical facts which make them pertinent or answerable, not the semantic facts of what they might mean.

We may say then that rather than conflicting with the *Posterior Analytics*, this account best explains the relation Aristotle sees between causes and the middle terms of syllogisms, and how these are explanations: for Why-questions must be analyzed as asking why some A belongs to some B in order to be metaphysically tractable, and the scientific demonstrative syllogism is the one that makes it apparent how it is that this A does indeed belong to this B.

If, therefore, we read Aristotle as claiming that explananda exhibit a predication structure, we may interpret the four causes as associated with the instantiation of that structure. That is, to be a cause is to be metaphysically related in one of four distinct ways to a certain kind of entity whose structure is, in turn, comprised of an 'A' term which inheres in some subject 'B'. If so, we may suggest the following definitions for the causes by revising the accounts given in the *Physics* along these lines:

<sup>46</sup> See also *Phys.* 190b11, where Aristotle concludes that everything that comes to be is composite ("to gignomenon hapan aei suntheton esti").

(1) Something is a material cause of A's belonging to B =<sub>def</sub> it is that out of which the structured entity [A's belonging to B] comes to be and which persists.

(2) Something is a formal cause of A's belonging to B =<sub>def</sub> it is the form or paradigm, which is the account [*logos*] of what it is for A to belong to some B.

(3) Something is an efficient cause of A's belonging to B =<sub>def</sub> it is the source of the motion or change in which A comes or continues<sup>47</sup> to belong to B.

Finally, (4) Something is a final cause of A's belonging to B =<sub>def</sub> it is that for the sake of which A belongs to B.<sup>48</sup>

The definition of the formal cause may be more succinctly put, once the predicational structure of the explanandum is made clear: formal causes simply are the predicables, both substantial forms and the entities in the non-substance categories such as quality and quantity. There need not, and presumably would not, be a distinct predicable in this sense for every linguistic predicate or property in the loose sense of the term. Rather, these predicables are restricted to the genuine entities which populate the categories, and which, as the metaphysical components of basic facts such as Socrates' *existing* or his *being pale*, make true predicative statements possible. Formal causes are thus "sparser" than predicates generally.

To be an instance of a formal cause, then, is simply to be some A which in fact belongs to its proper subject, B. In the case of the generation of statues and living beings, A and B can be identified with the Aristotelian substantial form and matter. This identification does not generalize, however, since such forms are not the only kind of predicable, nor are all subjects the matter of a hylomorphic compound.<sup>49</sup>

<sup>47</sup> The definition disjoins coming to be and continuing to be in keeping with Aristotle's original definition of efficient cause as the principle of change or rest. The disjunction seems necessary, since there are efficient causal relations which do not involve the gain of a predicate: if an object is held aloft by a current of air, and when this has passed, by a magnetic field, the field seems to be the efficient cause of the object's being suspended even though it did not cause it to become suspended. One might object to the disjunctive nature of the definition, or that the relata of causation must be events. These are fair questions to raise, perhaps, but this is not the place to consider them. One may simply note that, if there are good arguments for restricting the ontology of causal relata, Aristotle's definition may perhaps be accordingly revised without being thereby vitiated; alternatively, the disjunctive nature of the definition may turn out to be justified by further analysis of efficient causation.

<sup>48</sup> These definitions may be further refined, since as written they preserve some of the roughness of Aristotle's initial characterizations in the *Physics*. However, those refinements would presumably require a more precise understanding of the metaphysics of these relations themselves, which is beyond the scope of this paper.

<sup>49</sup> Indeed, if the explanandum is wisdom's belonging to Socrates, the subject is a substance, which would seem to be a form. Likewise, the material cause of an event such as the Persians attacking the Athenians might be, perhaps, the Athenians, or Athens, depending on how we construe the property of being attacked. I take it to be a virtue of this interpretation that it allows the four-causal framework to be extended to phe-



To be a formal cause is thus to be an instantiated predicable. The notion of an instantiated predicable is in turn contained in each of the other definitions: material causes are the proper subjects in which predicables inhere, efficient causes are agents whose activity yields the instantiation of predicables (for example, the replacement of one predicable, such as *pale*, by another one of the same genus, such as *tan*), and final causes are the purposes or goods for the sake of which predicables inhere.

The revised definitions above therefore exhibit a *pros hen* structure: the formal cause is defined as some A which belongs to some B, while each of the other three definitions advert to the existence or coming-into-existence of such a compound structure. The formal cause, that is, appears to be basic insofar as its account is simply the account of what it is for some entity in one of the categories of being to exist, whether that amounts to the existence of a hylomorphic compound such as man (if the form is a substantial form), or the instantiation of an attribute (for non-substantial forms). Thus, the efficient, material, and final causes are each defined with reference to the instantiation of a formal cause, which is in turn defined independently.<sup>50</sup>

This set of definitions thus provides a unity to the causes without reducing them to a single kind of causation, and connects them directly to one another rather than by way of features extrinsic to them. That is, the four causes are related to one another not because they all fall under a second-order non-essence-specifying account, nor because they share the feature of being answers to Why-questions, but because their definitions overlap. They are all relationships which bear in mutually irreducible ways on the possibility of fact-like entities whose structure is distinguishable into two parts, one of which belongs, in some sense which must be explained, to the other.

---

nomena which are not straightforward hylomorphic compounds or substrates which lose and gain contraries, even if these constitute its core of applicability.

<sup>50</sup> A worry may then arise that this strategy will be too generous: if causes are relations defined with reference to the formal cause, what prevents us from defining an indefinite number of increasingly absurd "causes", such as "*that which does not prevent A from belonging to B*", or "*that which desires that A belongs to B*", or "*that in front of which A belongs to B*"? These are all defined in terms of the notion of inherence, but they hopefully do not count as causes. On one hand, this worry is a general one which arises on any attempt to understand a given concept as unified *pros hen*: the device itself requires some way of restricting the items which merit the relevant term. The four causes are no exception, but since the worry expressed is a general one for the device of focal connection it may to some extent be set aside here. No doubt part of the solution as regards causes, however, consists in the fact that these relations, unlike the negative or gerrymandered ones, are required for an understanding of how these basic fact-like entities come to exist at all. If there are arguments in favor of recognizing other such relations, then they might merit the term 'cause' as well, but not otherwise.

### 6. *Consequences of this interpretation*

The preceding interpretation shows that there is a satisfying account in terms of *pros hen* homonymy, according to which we may understand both the diversity of causal relations as well as the reason Aristotle has for grouping them together. If we accept, first, that the causes are connected or associated homonyms for Aristotle, and second, that the definitions of the different senses of ‘cause’ overlap in that they all make reference to predication, we can see how the causes are both metaphysical and epistemological principles such that their metaphysical characteristics entail their epistemological ones. For if the causes are necessary or necessary and sufficient conditions for predication, they constitute the metaphysical basis for the possibility of (true) assertoric speech and, by parallelism, thought – and hence knowledge. Without the causes, there could be no metaphysical ‘combination’, and hence all thoughts and assertions of combination would be false.

This feature is a significant advantage in favor of taking ‘*aitia*’ to be homonymous, since it gives us a powerful answer to the question of why these relations are grouped together, given that they are so clearly different when considered just as roles or relations. Still, one might argue for a univocal account: one might claim that ‘*aitia*’ is primitive and hence indefinable, but nonetheless univocal, or perhaps that there is something like an analytic connection between ‘*aitia*’ and the notion of being a correct explanation, which we can keep separate from the actual metaphysical relations by which something falls under the description. ‘*Aitia*’ might thus be univocal regardless of the differences in the natures of the metaphysical relations Aristotle describes. One drawback of such an approach, however, is that we must then take the notion of *being explanatory* (in the robustly epistemic sense) as basic if we are to keep ‘*aitia*’ univocal, and hence lose the ability to say precisely what being explanatory consists in. On my reading however, we can say quite a bit about why these relations are explanatory, even though we give up univocity. Nor, however, is it necessary to claim that there is *no* usage of ‘*aitia*’ which is roughly univocal and means roughly ‘correct explanation’ in order for this strategy to work. Starting with a pre-theoretical notion of being a correct explanation, we might perceive that there are four very different metaphysical relationships which are especially important for the way things are (and this is Aristotle’s point); there is no single, general explanatory relationship under which these relationships fall as species. Accordingly, when properly analyzed, ‘*aitia*’ turns out to be non-univocal. We might make a stronger claim that ‘*aitia*’ has a univocal ‘nominal definition’ but a non-univocal ‘real definition’. Something along those lines might be accurate (see n. 20), but Aristotle’s treatments do not display such a sharp distinction.

A further consequence of this interpretation is that the formal cause becomes primary, since the definition of the formal cause, while not trivial

(since it appeals to the notion of essence), is clearly the most basic of the set: the account of the formal cause states what it is for a certain predication complex to exist.<sup>51</sup> If that is right, then it begins to look as though Aristotle has not, after all, strayed very far from Plato in criticizing the latter for overlooking efficient causes:<sup>52</sup> formal causation is in fact primary for both, but whereas Socrates in the *Phaedo* claims that participation in a Form is the *only* cause of something's being, e.g., beautiful,<sup>53</sup> Aristotle insists (besides protesting against Plato's separate Forms themselves) that more is required to understand how something comes to be beautiful.

It is important, however, not to confuse the primacy of formal causation with the primacy of form, in the sense of essence – essences may be primary entities in some sense for Aristotle, but it is not for this reason that formal causation is primary relative to the other causes. Formal causation is indeed basic, but according to Aristotle, it covers both *per se* and *per accidens* predications, as is made clear by the role of the formal cause in both qualitative change and the nature of substance – essences are not the only kind of predicable. The priority of formal causation derives from its being definitionally primary as a kind of relation, not from the metaphysical importance of form. In other words, formal causation is basic because of its general features as a causal role, not because the occupants of that role are themselves ontologically fundamental.

Further, we can see why the generation of a substance or artefact is rightly thought to be a core case for causation.<sup>54</sup> Substances and artefacts are ultimately the best or most important examples of form-in-matter, since they are, respectively, the most primary kind of being and the model by which Aristotle thinks the nature of substance is best illustrated. Thus, they constitute the most important test for the theory of causes.<sup>55</sup> The preceding

<sup>51</sup> I thus agree with the core of Robin's (1910) proposal to take the formal cause as primary. Unlike Robin, however, this account makes no attempt to reduce the other modes of cause to the formal one, nor is there any pressure to do so.

<sup>52</sup> *Metaph.* 991b3f.

<sup>53</sup> "For it seems to me that if there is something beautiful besides the beautiful itself, its being beautiful is because of nothing other than its participating in the beautiful; and I say so for everything." (*Phaedo* 100c4–6) This is only the so-called "safest [*asphalestaton*] answer" to the question of how things come to be what they are. How this answer relates to the "cleverer [*kompsoteran*] answer" (105c), and whether Aristotle unfairly claims that Plato ignored efficient causes are complex matters – see Fine (1984) for discussion.

<sup>54</sup> As argued by Graham (1987) and as implicitly understood in many general presentations of the four causes, about which see Sprague (1968).

<sup>55</sup> This would also explain why Aristotle thinks that causation is applicable among entities in non-substance categories though not literally so: "Again, if the causes of substances are causes of everything, still different things have different causes and elements, as was said; the causes of things that are not in the same class, e.g., of colors, sounds, substances, and quantities, are different except in an analogical sense" (1071a25–27). Since substance is to be analyzed in terms of form and matter, and the

interpretation thus confirms this thought, but does not have the undesirable consequence of making other phenomena such as qualitative alteration and locomotion turn out to involve non-paradigmatic notions of causation.<sup>56</sup>

A possible disadvantage of this approach is that the notions of efficient and final causation appear to be downgraded: there cannot be inherence without matter and form, but one would think it possible for there to be inherence without movers or goals, at least in the sense that the inherence relation itself involves only two parties. That is, what many theorists would argue is the *only* part of Aristotle's theory that corresponds to a causal notion is in fact, on this view, not even the basic mode of causation.<sup>57</sup>

Whether or not this is a genuine disadvantage of the theory depends on whether we accept Aristotle's general approach to causation as metaphysically explanatory relations, and perhaps as well on how the theory performs in giving an account of the details of the different modes of causation.<sup>58</sup>

Aristotle has at his disposal a simple but powerful way of accounting for both the plurality and the unity of his four causes, one which coheres both with his underlying metaphysical framework and with the theoretical work the causes are meant to perform. Despite being manifestly different metaphysical relationships, and so plausibly homonymous, they are ultimately structured around the relationship between predicables and their subjects.

---

entities of the other categories ontologically depend on substance, the notion of properties belonging to non-substances is arguably different in a metaphysically important way from relations between substances and entities in the other categories.

<sup>56</sup> This defect affects accounts such as Robin's (1910) and Graham's (1987), which make form in the sense of substance or an artefact the unifying factor for the causes, without distinguishing it from formal causation considered more generally. Similarly, there is no need on this proposal to follow Hennig (2009) in making natural change essential to all four causes as such, nor in attempting to reduce the causal/explanatory roles to two, the formal/final and material/efficient roles, such that the members of each pair differ only in the nature of their explananda (*viz.*, subjects of change and the changes themselves).

<sup>57</sup> Indeed, if each of the definitions given above corresponds to a different sense of '*aitia*', we may think that there is all the more reason to doubt that the theory of the four *aitiai* is a theory of causes or causation. This objection would be too quick, however, since it presumes that 'cause' is univocal, which we have reason to doubt. Even if we assume 'cause' to be univocal, however, it would nonetheless be the case that Aristotle's cluster of *aitiai* embeds an understanding of a variety of metaphysical determination relations, including that picked out by the allegedly univocal 'cause', together with the assertion that these relations should be treated together. That fact should not be obscured by the semantics of 'cause' in contemporary English.

<sup>58</sup> Thus construed, Aristotle's theory also offers an intriguing possibility for the problem of offering a definition of causation: on his account, the task of defining causation is transformed into the task of defining a special group of concepts, no one of which is required to bear the full load of answering to everything we would wish to count as causal. The task of defining, say, *efficient* causation would then be more tractable, if only because that of explaining the relation of efficient causation to the primary notion is a more clearly delimited one.

Indeed, it makes sense for Aristotle to wish to relate his theory of causation to his theory of predication: his account of scientific explanation and understanding becomes far more compelling if the metaphysical relations it privileges have a clear and direct connection to his views about the basic constituents of the natural world.

Thus, we can account for the interrelation of the four causes by means which are characteristic of Aristotle's way of handling philosophically important concepts: a form of pluralism which both draws important distinctions, but, by treating the distinct modes together, allows for a subtler understanding of the issue at hand.<sup>59</sup>

- Ackrill, J. L. (ed.) 1963. Aristotle. *Categories and De Interpretatione*. Oxford.
- Annas, J. 1982. "Aristotle on Inefficient Causes". *The Philosophical Quarterly* 32, 311–326.
- Barnes, J. (ed.) 1994. Aristotle. *Posterior Analytics*. Oxford.
- Charlton, W. (ed.) 1992. Aristotle. *Physics. Books I and II*. Oxford.
- Fine, G. 1984. "Forms as Causes: Plato and Aristotle". In *Mathematik und Metaphysik bei Aristoteles. Akten des zehnten Symposium Aristotelicum*. Ed. A. Graeser. Bern, 69–112.
- Graham, D. W. 1987. *Aristotle's Two Systems*. Oxford.
- Hennig, B. 2009. "The Four Causes". *The Journal of Philosophy* 106, 137–160.
- Irwin, T. H. 1981. "Homonymy in Aristotle". *The Review of Metaphysics* 34, 523–544.
- . 1988. *Aristotle's First Principles*. Oxford.
- Lewis, F. A. 1991. *Substance and Predication in Aristotle*. Cambridge.
- Matthews, B. 1972. "Senses and Kinds". *The Journal of Philosophy* 69, 149–157.
- Moravcsik, J. M. 1975. "'Aitia' as Generative Factor in Aristotle's Philosophy". *Dialogue* 14, 622–638.
- Owen, G.E.L. 1960. "Logic and Metaphysics in Some Earlier Works of Aristotle" *Journal of Hellenic Studies* 77, 103–111. Reprinted in *Logic, Science and Dialectic*. Ed. M. C. Nussbaum. Ithaca 1986.
- Robin, L. 1910. "Sur la conception aristotélicienne de la causalité". *Archiv für Geschichte der Philosophie* 23, 1–28; 184–210.
- Ross, W. D. (ed.) 1924. *Aristotle's Metaphysics*. Oxford.
- . (ed.) 1936. *Aristotle's Physics*. Oxford.
- . (ed.) 1945. *Aristotle's Prior and Posterior Analytics*. Oxford.
- Shields, C. 1999. *Order in Multiplicity. Homonymy in the Philosophy of Aristotle*. Oxford.
- Sprague, R.K. 1968. "The Four Causes: Aristotle's Exposition and Ours". *The Monist*, 298–300.
- Vlastos, G. 1969. "Reasons and Causes in the *Phaedo*". *Philosophical Review* 78, 291–325.
- Ward, J. K. 2008. *Aristotle on Homonymy: Dialectic and Science*. Cambridge.
- Wedin, M.V. 2000. *Aristotle's Theory of Substance: The Categories and Metaphysics Zeta*. Oxford.

<sup>59</sup> For comments and criticism of earlier drafts of this paper I am especially grateful to Christopher Shields, Terence Irwin, Vasilis Politis, Gail Fine, and audiences at the University of Oxford and the Pacific Division of the American Philosophical Association. I would also like to thank the anonymous referees for this journal for their helpful comments and suggestions.