SOCRATES’ NEW AITIA: CAUSAL AND METAPHYSICAL EXPLANATIONS IN PLATO’S PHAEDO

RAVI SHARMA

The section of the Phaedo that recounts Socrates’ intellectual ‘autobiography’ and culminates in his turn to the theory of Forms (95 e–102 b) has long been considered a key passage for understanding Plato’s mature philosophy. Besides offering explicit commentary on various themes of Presocratic thought, it is one of the few Platonic texts purporting to explain some of the considerations that motivate the theory of Forms.

Yet as to what precisely is going on in the passage, there has been considerable disagreement. And although scholarly dispute is often a source of interpretative riches, the lack of agreement here extends, quite remarkably, even to the broadest outlines of an adequate interpretation. The central difficulty concerns the nature of the new aitia—the new mode of explanation—that Socrates advances starting at 100 b of the dialogue.¹ Some interpreters consider Socrates

¹ No single English expression adequately captures the sense of the word αἰτία. “Cause” is far too narrow to cover its range of uses and, I shall argue, is inappropriate for the present passage. “Explanation” and “mode of explanation” are less misleading, and I shall use them wherever I venture to translate the word. Some have argued that “explanation” wrongly encourages one to think of an αἰτία as a linguistic or propositional item, but I agree with Barnes in thinking that incorrect about English usage (J. Barnes, Aristotle: Posterior Analytics, 2nd edn. (Oxford, 1993), 89–90). On one influential reading of the passage, a Platonic αἰτία just is a propositional item: see M. Frede, ‘The Original Notion of Cause’, in M. Schofield, M. Burnyeat, and J. Barnes (eds.), Doubt and Dogmatism (Oxford, 1986), 217–49, repr. in id., Essays in Ancient Philosophy (Minneapolis, 1987), 125–50 at 129. Note, however, the rebuttals in I. Mueller, ‘Platonism and the Study of Nature (Phaedo 95 ff.)’ ['Platonism'], in J. Gentzler (ed.), Method in Ancient Philosophy [Method]
to be promoting a new vision of causal explanation. Others have insisted that he is pursuing explanations of another sort—‘logical’ or perhaps ‘metaphysical’ ones.

There is no way to specify a concept of causation that would be uncontroversial for interpreters of classical philosophy, let alone for professional philosophers generally. In characterizing one sort of interpretation as ‘causal’, I mean simply that it takes Socrates to be concerned with explaining the process of change, however precisely the character of such an explanation may be conceived. Causal readings of the *Phaedo* are typically set out in Aristotelian language—by way of a judgement that Socrates is pursuing one or more aspects of Aristotle’s ‘four-cause’ doctrine. Thus, it is variously held that Socrates’ new *aitia* is either ‘efficient’ or ‘final’ or ‘formal’ in nature.¹

The ‘logical’ or ‘metaphysical’ readings also vary considerably. What they have in common is that, in one way or another, they all understand Socrates to be concerned with analysing the constitu-

Socrates' New Aitia

section of a state of affairs, though not with explaining the process by which that state is produced. To put the point suggestively, though vaguely, they all consider Socrates to be interested in being rather than becoming.

As will emerge, I agree with those who regard Socrates as pursuing explanations of a metaphysical sort. In recent years several versions of such a reading have been defended in the literature. Yet they have not gone far enough. Specifically, they have tended to neglect a problem posed by certain details of the passage for any interpretation whereby Socrates renounces causal concerns entirely. Confronting that problem head-on will necessitate a new account of the 'autobiography' as a whole.

I shall begin by discussing the problem. It turns on Socrates' assertion at several points that he is ultimately concerned with the explanation of coming-to-be (and passing-away) no less than with that of being. As I shall show, the available versions of the metaphysical reading have failed to do justice to this assertion, which implies that Socrates' preferred mode of explanation is not completely heterogeneous with those he has left behind.

Having identified the problem, I shall go on to argue that some type of metaphysical reading is none the less correct: the competing, causal readings do not adequately account for the way in which Socrates characterizes his new aitia at 100 B–101 C of the dialogue. I shall subsequently underscore that judgement by examining Aristotle's testimony as to what concerns Socrates here. Aristotle has sometimes been taken to support an interpretation whereby Socrates' interests are securely causal in nature. But he is in fact much too intent on pressing his own theories to provide us with an impartial account of what is at stake in the dialogue.

In order ultimately to resolve the tension between the causal and the metaphysical readings, I shall need to undertake a broad reinterpretation of the 'autobiography'. Beginning with Socrates' criticisms of materialist explanations (96 C–97 B), I shall suggest that they all turn on an epistemic constraint governing Socrates' understanding of what constitutes an adequate aitia. Grasping the point will allow us to see why Socrates initially turns to teleological speculation as an alternative to materialism (97 B–99 C) and why he later finds himself driven towards a project of metaphysical analysis (100 B–102 D). One virtue I shall claim for the interpretation to be
developed is that it shows the entire passage to be much more carefully structured than readers have tended to think.

I

Let me begin with the challenge that must be faced by any reading of Socrates’ new aitia as non-causal in character. It rests on several passages of the narrative recounted from 95ε to 100α.¹ In each passage Socrates describes his larger concerns in ways that would suggest he expects his aitia to make for improved explanations of a causal sort. Indeed, it is difficult to see how what he says might plausibly be interpreted otherwise.

When he embarks on his autobiographical excursion, Socrates announces that he is about to undertake a general enquiry into ‘the mode of explanation concerned with coming-to-be and passing-away [διόλου γάρ δεῖ περὶ γενέσεως καὶ φθορᾶς τὴν αἰτίαν διαπραγματεύσασθαι]’ (95ε 10–96α 1).² Just a few lines later, at 96α 8–10, he begins the tale of his youthful pursuits by adding:

ὑπερήφανος γάρ μοι ἐδόκει εἶναι, εἰδέναι τὰς αἰτίας ἑκάστου, διὰ τί γίγνεται ἕκαστον καὶ διὰ τί ἀπόλλυται καὶ διὰ τί ἐστι.

I thought it grand to know the explanation of each thing, why each comes to be and why it perishes and why it is.

Were those the only remarks of that kind, they could perhaps be

¹ Here and throughout, I am speaking only of the Platonic Socrates. I shall not be concerned with the potential relevance of anything in the Phaedo to the views of the historical Socrates.

² Compare also 97c 6–7. The term γένεσις and its cognates might be used either for a thing’s coming-to-be simpliciter or else for the coming-to-be of some further determination of the thing—that is, for so-called qualitative change. The noun φθορά and the verb ἀπόλλυσθαι (cf. 96α 9; 97b 5) are poorly suited to serve as counterparts of the latter notion. But if in fact Socrates neglects to distinguish clearly between the two senses of γένεσις/γίγνεσθαι, he might well be content to use φθορά/ἀπόλλυσθαι as a generic alternative. For further comment, see D. Gallop (trans. and comm.), Plato: Phaedo [Phaedo] (Oxford, 1975), 170–1. A. Nehamas argues that when Socrates speaks of coming-to-be and passing-away, he is centrally concerned not so much with change as with the co-presence of opposites: ‘What falls under “coming-to-be and perishing” . . . include[s] all cases where one can say that something has a certain property at time t₁ and that it does not have that property at time t₂, where k and i may or may not be identical. In fact, Socrates seems most concerned with cases of the former sort’ (‘Predication and Forms of Opposites in the Phaedo’, Review of Metaphysics, 26 (1972), 401–91 at 403–4). This is an unlikely interpretation of γίγνεσθαι/ἀπόλλυσθαι, and it is not well supported by Socrates’ remarks about his youthful preoccupations.
In speaking as he does, Socrates might be referring solely to the passages that immediately follow, the ones detailing his early interest in various ‘Presocratic’ theories. In other words, he might be saying simply that he needs to relate the story of his initial fascination with those theories as a way of establishing a contrast by which to clarify his new sort of explanation—the one to which he now subscribes. However, that interpretation would prove incompatible with several later passages of the ‘autobiography’, which clearly show that the remarks in question are meant to cover Socrates’ mature concerns as well as those he has left behind.

At 97 b 3–7 Socrates offers the following comment regarding his turn away from the approach to explanation pursued by his materialist predecessors:

I no longer persuade myself that I know how it is that one thing [or a unit: ἕν] comes to be, nor in sum how anything else comes to be or perishes or is—at least, by this approach to the investigation [κατὰ τὸν τρόπον τῆς μεθόδου]. Instead, I muddle about in my own way by another approach [τὸν ἄλλον τρόπον αὐτὸς εἰκότως φάω], and in no way do I accept the former one.

The other approach to which Socrates alludes is precisely the mode of explanation that he will elaborate shortly, starting at 100 b. The way he introduces it suggests that he thinks of it as yet another means of pursuing the overall investigation (methodos), the one into coming-to-be, perishing, and being.’ That reading is confirmed by 99 d 1, where Socrates describes his new intellectual orientation, his ‘second voyage’, as involving a search for ‘the explanation’. Since

---

1 In a note on 97 b 6, J. Burnet takes the phrase τρόπος τῆς μεθόδου to mean simply “method of enquiry”, as μέθοδος alone later means. (See Plato's Phaedo [Phaedo] (Oxford, 1911), 103.) In that case, the words τῆς μεθόδου would not refer specifically to the investigation (into coming-to-be etc.), as I am presuming. One might then be tempted to interpret the mention of a second τρόπος (τῆς μεθόδου) as a reference to a method that is directed at an altogether different goal. However, any such proposal would be ill-founded. Socrates says nothing in the immediate context to specify what the supposed new goal might be; nor does he so much as hint that his initial goal has been abandoned. No matter how one understands the phrase τρόπος τῆς μεθόδου, his remark is most straightforwardly understood as implying that the new τρόπος serves the same end as the former one—the explanation why each thing ‘comes to be or perishes or exists’. That consideration undermines R. M. Bolton’s judgement that the present passage ‘seems clearly to indicate that [Socrates] rejects peri phuseos historia altogether, and replaces it with a different methodos’ (‘Plato’s Discovery of Metaphysics: The New Methodos of the Phaedo’ [‘Discovery’], in Gentzler (ed.), Method, 91–111 at 101).

6 τὸν δεύτερον πλοῦν ἐπὶ τὴν τῆς αἰτίας ἐλήμφησα δε σεπραγμάτευμα βαλόμενο σοι... ἐπίδειξιν ποιήσομαι, ὦ Κέβης.
Socrates has just confessed to lacking any satisfying teleological explanations of the sort he had once sought from Anaxagoras, he can be referring only to the overall investigation, the one mentioned at 97 B.⁷

On that point, a word of clarification is in order. It has been common to read 99 D 1–2 as if Socrates is still concerned with teleological explanation in his 'second voyage'. But whatever may be the appeal of such a reading when considered in the light of broad speculations regarding Plato's moral and scientific preoccupations, it is straightforwardly invalidated by the preceding lines, 99 C 6–9, where Socrates laments his inability to produce explanations of a teleological sort. The conflict between those lines and D 1–2 was noted by Vlastos, who pointedly remarked: 'Socrates makes it abundantly clear that he is still, at the time of speaking, “deprived” of the teleological aitia he had been looking for. But it is no less clear that . . . the “second journey” . . . is a method of inquiry on which he has already been engaged.'⁸ In recent years, several interpreters have sought to revive the teleological reading by charging that Vlastos misreported the text of 99 C 8. As the argument runs, Socrates there uses the aorist ἐστερήθην rather than the perfect ἐστέρημαι: he was denied the teleological aitia, but it is not at all clear what we should take his present state to be.⁹ So far as it goes, that observation is quite correct. However, the larger objection is none the less ill-founded. In the lines that immediately precede, Socrates clearly confesses his lack of teleological explanations when he declares:

ἐγὼ μὲν οὖν τῆς τοιαύτης αἰτίας δὴν ποτὲ ἐχει μαθητής ὅτου ἡδίσταν ἔδωκαν; ἂν γενοίμην.

For my part, I would most gladly learn from anyone at all just how this [teleological] mode of explanation works.

---

⁷ The expression “second voyage” is proverbial, although its precise meaning is uncertain. It may originally have been used to indicate a turn to the oars when the wind has subsided and the use of a sail is impossible, or else it may have signified the making of a second, safer journey when the first has failed. (See Burnet, Phaedo, ad loc., and note the extended discussion in Y. Kanayama, ‘The Methodology of the Second Voyage and the Proof of the Soul’s Indestructibility in Plato’s Phaedo’ ['Methodology'], Oxford Studies in Ancient Philosophy, 18 (2000), 41–100 at 87 ff.) There is some support for both views, though the first is perhaps better attested. Whichever way one chooses to interpret the image, it is clear that Socrates does not have in mind the pursuit of a new ‘destination’, or goal: the purpose of the voyage is still ἡ τῆς αἰτίας ζήτησις.


Here, the aorist γενοίμην has a purely aspectual significance and is not to be understood as a past tense. In view of that, there is no warrant for construing Socrates’ second voyage as if it were aimed at more adequate teleological explanations; and thus, as mentioned above, the only reasonable way of understanding the phrase τῆς αἰτίας at 99 d 1 is as a reference to the explanation of coming-to-be, perishing, and being.

Scholars who take Socrates to be introducing a notion of metaphysical explanation at 100 b ff. therefore need to show how their interpretation may be squared with 97 b and 99 d. But they have not done so. To cite only the most prominent example: despite an otherwise exacting discussion, Vlastos simply omits to mention 97 b. As for 99 d, he contents himself with rendering the expression ἡ τῆς αἰτίας ζήτησις vaguely, as “the search for aitia”. Given his broader reading of Socrates’ ‘autobiography’, it is not at all clear what Vlastos might take the latter phrase to mean. He would presumably want to understand “the search for aitia” in some general way, one that is neutral with respect to metaphysics and the various forms of causal explanation that Socrates ostensibly abandons. Yet Vlastos insists not only that there is what he calls a ‘radical’, or ‘cate-

10 In Attic prose, the aorist optative is rarely, if ever, used to express past possibility. Had Socrates wanted to express such a notion, he would most likely have employed ἄν with an imperfect or aorist indicative, the choice of tense depending on the degree to which he wanted to emphasize that his desire went unfulfilled. Gallop’s translation is therefore incorrect: ‘Now I should most gladly have become anyone’s pupil, to learn the truth about a reason of that sort’ (emphasis added).

11 I might add that there is no basis here for Gallop’s speculation (Phaedo, 176–7) that it is Plato (rather than Socrates) who thinks of the second voyage as a route to some sort of teleological explanation. Whatever may be the merits of such a reading when it comes to accounting for other, later details of Socrates’ arguments, it is at this stage unsupported by anything in the text. I emphasize the point because a roughly similar reading has recently been advocated, albeit tentatively, by R. M. Dancy (Plato’s Introduction of Forms [Introduction] (Cambridge, 2004), 293 with n. 8), who maintains that even if Socrates is not presently pursuing teleological explanations, he still has them in mind as his ultimate, as yet unrealized, goal. As Dancy argues, if one refuses to concede that point, then the whole discussion of teleology from 97 b to 99 c ‘sounds like a digression’. I shall eventually explain what I take the relevance of that material to be.

12 ‘Reasons’, 82 n. 15. In that note Vlastos also cites with approval N. R. Murphy’s account of the passage, but Murphy simply ignores the mention of a ζήτησις τῆς αἰτίας. In arguing that Socrates abandons any concern for scientific explanation, Murphy insists on understanding the ‘second voyage’ as one directed towards a new destination. Yet he neglects to explain why that destination is still described as one that will yield ‘(the) explanation’. (See The Interpretation of Plato’s Republic [Interpretation] (Oxford, 1951), 145–56.) The passage is not discussed by Bolton in the course of his metaphysical reading of Socrates’ new αἰτία.
gorial’, difference between metaphysical (or ‘logico-metaphysical’) and ‘physical’ explanation, but also that Socrates is perfectly aware of that difference. Indeed, he argues that Socrates’ ‘autobiography’ is properly read as the drama of a gradual awakening to the relevant distinctions. For at 96c–97b Socrates discusses a number of paradoxes that supposedly perplexed him in his early pursuit of natural science (περὶ φύσεως ἱστορίαν, 96a7) and ultimately contributed to his turn away from such enquiries. As Vlastos would interpret the passage, all of the paradoxes spring from a conflation of ‘mechanical’ and ‘logical’ sorts of explanation. Socrates is thus chronicling his early inability to disambiguate explanations of those two sorts. None the less, Vlastos argues, Socrates eventually comes to grasp the relevant distinctions: in subsequently elaborating his new mode of explanation and showing it to be immune to the paradoxes (100b–101c), he effectively uses the idea of a properly ‘logico-metaphysical’ enquiry as a means of diagnosing his prior confusion. That interpretation is certainly an elegant one, and it is elaborated with an exemplary rigour and attention to detail. However, it can hardly explain the phrase ἡ τῆς αἰτίας ζήτησις at 99d. After all, there would be little sense in supposing that, just as Socrates is preparing to expose his youthful errors to the light of day, he would still presume there is some fundamental connection between ‘logic’ and natural science and would still speak as if there is a general notion of explanation that covers both sets of concerns.

The foregoing difficulty in Vlastos’s reading is emblematic of a more general problem in the literature when it comes to reconciling a metaphysical interpretation of the new αἰτία with Socrates’ contention that he is concerned with the explanation of how each thing ‘comes to be or perishes or is’. When faced with that problem, one might well be tempted to suppose that the metaphysical interpretation is but an idealization of certain confused remarks at 100b·., or perhaps even a falsification of a theory whose primary


14 See ibid. 95–102.

15 Towards the end of his paper (‘Reasons’, 102 ff.) Vlastos suggests that Socrates’ new αἰτία ultimately has some causal significance to the extent that it testifies to what Vlastos considers the Platonic dream of establishing an a priori science of nature. But Vlastos nowhere proposes that Socrates actually presents his metaphysical αἰτία with an eye to accomplishing that purpose. Indeed, he thinks that Plato is doubtful as to whether such a science can in fact be worked out (108–9). Hence it would not be open to him to maintain that what Socrates has in mind at 99d1 is in fact some overarching project of causal explanation.
significance is intended to be causal. Any such conclusion would be precipitate, however. As we shall now see, when Socrates presents his new *aitia*, he says nothing to justify the idea that he considers the Forms to be components of a new causal theory. In fact, his description of the new *aitia* is best taken to support some version of a metaphysical reading.

II

To be sure, Socrates at one point characterizes the Form the Beautiful as the reason objects ‘become beautiful’, and many interpreters have understood his remark as evidence that he supposes the Forms to be causes of qualitative change. However, Socrates does not in fact ascribe any causal power to the Forms. He has a different conception of the way in which a Form explains the beauty of other things. The passage crucial for appreciating his point is 101 c 2–9, where he clarifies the sense in which he takes the Forms to supply an account of ‘becoming’. He instructs Cebes to respond as follows to requests for explanations of how things come to be:

You would loudly cry out that you don’t know how each thing comes to be [γιγνόμενον], except in so far as it comes to participate in the particular reality of whatever it is in which it comes to participate [μετάσχον τῆς ἰδίας οὐσίας ἐκάστου ὁ ὁ μετάσχει]. In the cases just mentioned [at b 10–c 1], you don’t know of any explanation for having come to be two other than coming to participate in the dyad [τῆς δύο ἰδιότητος μετάσχει], and whatever is to be two must come to participate [μετάσχει] in this; and likewise, whatever is to be one, in the monad. As for those splittings and attachings and the

---

16 τῷ καλῷ τὰ καλὰ γίγνεται καλά, 100 ε 3. I here follow the editors of the recent OCT edition in preferring the reading of MS T, which they deem the sole representative of the second family of the manuscript tradition. (The same reading is also recorded by the corrector of MS B; see E. A. Duke, W. F. Hicken, W. S. M. Nicoll, D. B. Robinson, and J. C. G. Strachan (eds.), *Platonis opera*, vol. i (Oxford, 1995), ad loc.) The verb γίγνεται is omitted by members of the first family (MSS B and D) and is placed after the second καλά by the members of the third. Several lines earlier, at 100 δ 8, the text is likewise uncertain. The reading favoured in the recent OCT, namely τῷ καλῷ σάντα τὰ καλὰ καλά, is that of T and several members of the third family. The members of the first family, which alone omit the γίγνεται at ε 3, uniformly include it at δ 8 before the final καλά. Given the conflicting evidence, at least one occurrence of γίγνεται may well be an interpolation, but it is not at all clear that both are (as Burnet suggested on the basis of an inaccurate report of the manuscript readings). Other editors, e.g. Robin and Rowe, have chosen to omit γίγνεται at ε 3 but to print it instead at δ 8.
rest of such subtleties, you would dismiss them, leaving it to those more sophisticated than yourself to make use of them in answering.

Socrates’ language is highly emphatic, the point of emphasis being that his new aitia offers no explanation of the process of change. That is brought home by his insistent use of the aorist μετασχέω, along with the cognate noun μετάσχεσις—which occurs in this passage only and is probably Socrates’ own coinage. Here, the aorist has no temporal significance at all, as is nicely underscored by the way in which it is used in conjunction with verbs of several tenses—present and future as well as aorist (γιγνόμενον c 3, γενέσθαι c 5, ἔσεσθαι c 6–7). Instead, Socrates uses μετασχέω and μετάσχεσις with a purely ‘ingressive’ force. That is to say, he is effectively indicating that all talk of coming-to-be is simply to be replaced by, or translated into, talk of coming-to-participate. And thus, the only sense in which Forms may be said to explain becoming is that they figure centrally in the account or analysis of what obtains when several objects come to be two, say.

The passage therefore provides no basis for maintaining that the Forms serve as ‘efficient’ causes. Nor can it readily be taken to support a thesis that although the Forms are not themselves causal agents, they are none the less part of a broad schema for a theory of efficient causation. The aitia that Socrates insists on at 101 c 4 is said to consist solely in the proposal that a thing (or group of things) comes to participate in the dyad, and Socrates never so much as hints that his way of speaking may be understood as an

---

17 The proposal that the aorist has ingressive force was made at least as early as F. M. Cornford’s Plato and Parmenides [Parmenides] (London, 1939), 79, and has since been endorsed in several prominent discussions of the passage. See e.g. Murphy, Interpretation, 147, and Vlastos, ‘Reasons’, 87 n. 32. Regrettably, the proposal has been passed over by most subsequent translators and commentators, who have not ventured any alternative explanation of the aorist. For instance, there is no hint that Socrates uses the aorist rather than the present in Gallop’s otherwise meticulous translation and commentary, or in the thoughtful paper by Annas, who cites Socrates’ remarks here as clear evidence that he is still concerned with developing a theory of causation (‘Inefficient’, 325). For several noteworthy exceptions to the general tendency, see Fine, ‘Forms’, 376 n. 41; C. J. Rowe (ed.), Plato: Phaedo (Cambridge, 1993), 245; and Mueller, ‘Platonism’, 8a.

18 Socrates here glosses over the difficulty of analysing the concept two, the logic of which cannot be assimilated to that of a concept like beautiful (his earlier example). As I am inclined to think, he never confronts the difficulty directly. Certainly, nothing he later says about number (at 104 x ff.) evinces a developed theory.

19 See especially Fine, ‘Forms’, 376, where Socrates’ use of the ingressive aorist is cited in support of the idea that the Forms are ‘causally relevant’ by being constituents of events that are causes.
opening onto some broader account of change. Indeed, at 100d 4–8, he pointedly refuses to venture any analysis of the Form–object relation, and he does so while continuing to maintain that the Forms are the crucial components of his own mode of explanation. Were Socrates to believe that such an incomplete analysis could form the basis of a theory of efficient causation—one, moreover, that justifies his abrupt dismissal of the explanations offered by his opponents—his confidence would be utterly without foundation. His way of proceeding would certainly be no more compelling than that of a materialist who insisted that twoness involves an association of units but who neglected to explain the exact nature of the association—whether it consists in mere juxtaposition or in some other form of relation—and who compounded that fault by haughtily refusing to specify the process through which the association could be said to come about. (Cf. 96 e 6–97 b 3, to be discussed below.)

It might be suggested here that Socrates’ talk of participation in a Form actually represents the discovery of another kind of cause altogether, say some type of ‘formal’ one. But given Socrates’ rejection of ‘splittings and attachings’ and like ways of speaking,
any such interpretation would be most unlikely. If in fact Socrates was interested in some proto-Aristotelian notion of a formal cause, he would have to be thinking of it as but one component of a broader theory. Taken in isolation, after all, the notion of formal causation could provide no account of why one thing comes to be beautiful (for instance) at a particular time and in contrast to other things that do not thus change. However, it would make little sense to describe Socrates as attempting to recommend one component (a formal one) of some larger causal theory by disclaiming any interest in another (an efficient one). It is precisely Socrates’ frustration with explanations involving ‘splittings and attachings’ that ostensibly motivates him to develop a more adequate mode of explanation. Once he does so, he never entertains the possibility that the two modes can exist side by side or, more generally, that there are any intelligible sorts of explanation besides the one he has chosen to pursue. Thus, at 100c5 he insists that an object is beautiful ‘for no other reason’ than that it participates in the Beautiful, while several lines later, at δ 4–5, he stresses that ‘nothing else makes [an object] beautiful’—that is, nothing other than its relation to the Beautiful.23 It would scarcely be plausible to maintain that although Socrates recognizes himself as an innovator with regard to the notion of a formal cause, he is none the less so confused about what he is proposing as to entertain that

23 A new defence of the idea that Socrates is committed to ‘formal’ causes has been offered by Sedley, who argues that formal causes constitute the ‘essence’ or ‘kernel’ of a causal statement, while other, efficient considerations are merely ‘secondary’ (‘Causes’, 116). Sedley proposes to read the comments about participation at 100c–d precisely as a relegation of efficient considerations to a secondary status; but it is not clear what it would mean for Socrates to promote a causal theory that completely refuses to specify how efficient causes may be said to function. In elaborating what he has in mind, Sedley relies heavily on a legal analogy: ‘When we have accounted for the murderous act by pointing to the murderous person, we have already said all that there is to say about where the actual responsibility lies. That is why . . . Plato does not include in the irreducible kernel of a causal statement the process by which a cause acts. How the murderousness was transmitted is no more important to a causal account than it was at Phaedo 100d 3–6 3 to establish whether it is by sharing, presence or whatever that the Beautiful comes to make things beautiful’ (124; cf. 116). Here, the ‘secondary’ causal considerations seem to drop out of play entirely. But that makes for an awkward causal theory, since it would not seem possible to give an account of responsibility without in fact seeking to explain the process of causation. To take Sedley’s example, what would one say about a situation in which several people with ‘murderous impulses’ are present in the vicinity of a ‘murderous act’? Should all of those people be deemed equally responsible for the relevant motivation with respect to the unfortunate victim and, therefore, for the act itself?
formal explanations may be considered replacements for efficient ones.\(^4\)

One ought therefore to grant that Socrates is abandoning causal considerations entirely at 100\(b\)–102\(a\). And once one does so, there is little to suggest that he intends to return to causal speculation at any point in the ensuing discussion, where he uses the theory of Forms as the basis for his final proof of the soul’s immortality (103\(c\)–107\(a\)). In the context of that proof, Socrates famously introduces a ‘more subtle’ or ‘more refined’ (κομψότεραν, 105\(c\) 2) way of answering questions, but he is not thereby sketching a programme for causal explanation. A full treatment of what he has in mind would take us well into the details of the final proof and would demand a separate study, especially in view of the divergent interpretations given to a number of difficult passages at 104\(b\)–105\(c\). But for the purposes of the present discussion, a few points will suffice.

In developing the sort of refinement with which he is concerned, Socrates introduces several examples that effectively acknowledge and exploit certain ‘necessary’ relations among entities, whether they be Forms or else what have been called “immanent characters”. The example he develops at greatest length involves the judgement “three is odd”. As Socrates puts the point, ‘whatever the character three takes hold of must be not merely three but also odd’ οἶσθα γὰρ δήπου ὅτι ἃ ἂν τῶν τριῶν ἰδέα κατάσχεται, ἀνάγκη αὐτοῖς οὐ μόνον τρισὶν εἶναι ἀλλὰ καὶ περιττοῖς’ (104\(d\) 5–7). Despite the dynamic language of ‘taking hold’, Socrates is not in fact asserting here that threeness

\(^4\) Socrates’ contention that his \(\alpha\gamma\iota\varsigma\) is in fact the only one also renders doubtful the above-mentioned interpretation offered by Vlastos concerning the relationship between the ‘autobiography’ and the introduction of the new \(\alpha\gamma\iota\varsigma\). If Socrates were initially led to reject natural science as a result of conflating ‘logical’ concerns with ‘physical’ ones, and if he were later to realize his mistake and then go on to formulate the idea of a purely ‘logico-metaphysical’ enquiry, he would no longer have any reason to reject ‘physical’ explanations as a source of confusion. In this connection, see especially 100\(c\) 10–d 3, which could not be given the sort of interpretation that Vlastos proposes with respect to Socrates’ earlier rejection of scientific explanation (at 96\(b\)–97\(b\)). The statements that Socrates is attempting to account for at 100\(c\)–d—those of the sort “\(x\) is beautiful”—could not plausibly be construed as \(a\ pri\ior\) truths for which any sort of physicalistic treatment would be absurd and for which a ‘logico-metaphysical’ analysis \(a\ pri\ior\) must be given. (Compare Vlastos, ‘Reasons’, 99–100.) When he discusses Socrates’ rejection of physical explanations, Vlastos contends that Socrates never goes so far as to charge his predecessors with the confusions that beset him in his youth (99 n. 64). But in that case, once Socrates has sorted out his own confusions he should have every reason to feel even more generous towards the sort of enquiry pursued by his predecessors.
is somehow productive of oddness. (Indeed, in what immediately follows, oddness itself is said to be what ‘effects’ a thing’s being odd—see 104 D 10, 12.) Instead of offering any analysis of the process by which there comes to be an odd-numbered group, Socrates is simply pointing out that any case of coming to be three is, by that very fact, a case of coming to be odd. Or as he goes on to observe, threeness can never become associated with the character opposed to oddness, namely evenness (104 E 1–2).

That sort of observation is precisely what motivates the ‘more subtle’ mode of answering, which is introduced at 105 B–C. Socrates elaborates it by means of several illustrations, one of them being that the question “What makes this body diseased?” may be given the subtle answer “fever” instead of the simplistic one “disease.” Some interpreters have taken him to be gesturing here at a type of causal analysis, but I do not think he intends to locate the causal basis of illness in fever. Instead, he considers fever to be but one sort of illness, as is made explicit in other contexts. In keeping with his remarks concerning the three-odd pairing, Socrates is noting that whenever fever comes to be present, disease does too and, therefore, that in some cases one may answer a question as to the basis for saying ‘x is diseased’ by pointing to the presence of fever. That is not in any sense a turn to a mode of explanation different from what was put forward at 100 B–D. Indeed, rather than cast his examples as part of a new aitia, Socrates describes them as but one more safe way of answering questions—a way that operates within the framework of his preferred mode of explanation. And although his characterization of the new answers as ‘more subtle’ is surely intended as a rejoinder to the superficially sophisticated

23 Earlier, at 104 D 1–3, Socrates does speak as if a certain character or Form can ‘compel’ (ἀναγκάζει) a collection to be either odd or even. Yet the sense of that remark is just what he goes on to state in the lines I have quoted: it is, for example, the necessary connection between threeness and oddness that allows us to speak of the former as compelling something to accept the latter.

26 See Rep. 610 B 1–2, and compare Alc. II 139 e–140 b; Xen. Mem. 3. 8. 3. 7.

27 On that interpretation, fever is a sufficient but not a necessary condition for disease. There is some question as to whether Socrates’ other illustrations of the ‘subtle’ way of answering should likewise be taken to describe conditions that are merely sufficient (as opposed to both necessary and sufficient). For helpful comment, see Gallop, Phaedo, 209–11. Even if Socrates is at points concerned with a relation that is necessary as well as sufficient, he need not be taken to think in terms of relations that are specifically causal in nature. Indeed, that sort of interpretation would scarcely work for the last example, which concerns the connection between oneness (μονάς) and being odd.
explanations of his materialist opponents (see 100 c 10; 101 c 8), there is no sign that Socrates means to recant his earlier rejection of causal explanations or to prepare the way for an eventual return to such explanations.  

III

If what has so far been argued is correct, the details of the passage strongly point towards a reading of the new *aitia* as metaphysical in character. Before developing that reading and trying to reconcile it with Socrates’ earlier descriptions of his concerns, let me pause to assess one further piece of evidence—namely, Aristotle’s testimony concerning the passage. For among his relatively few explicit comments on material from the Platonic dialogues, Aristotle declares several times that Socrates is here treating the Forms as if they were efficient causes of coming-to-be and passing-away.

Some interpreters have not hesitated to reject that testimony out of hand. Vlastos, for example, brushes it aside with the acerbic remark, ‘It is hard to believe that so patent a misreading of Plato’s doctrine could have been taken seriously by modern scholars.’ Yet others have been much more cautious. Rather than accept that Aristotle might have misunderstood the passage completely or else misrepresented what is said, they tend to presume that Aristotle is correct, at least in broad outline. Perhaps Plato’s Socrates does not really intend the Forms to be efficient causes, but surely Aristotle is right in thinking that the new *aitia* is somehow causal in nature.

28 Most scholars who reject a causal interpretation of 100 b–d similarly reject the idea that causal considerations re-emerge in any form at 105 b. Vlastos, however, takes a somewhat different view: although he thinks that the ‘clever *aitia*’ (as he misleadingly calls it) is still a form of ‘logico-metaphysical’ explanation, he also holds that Socrates considers it to have ‘causal implications’—namely, that any proper statement of physical law has the status of logical necessity. Vlastos mainly supports his case that Socrates is concerned (in part) with causal relations by means of the example I have mentioned, which he calls ‘a textbook example of a cause in Greek medicine’ (‘Reasons’, 105). But as I have already suggested, fever is here being understood merely as a species of sickness. Vlastos’s presumption to the contrary depends largely on his conviction that the ‘clever *aitia*’ must be ‘informative’ (103–4 n. 72). Yet there is nothing in the text to support that conviction. The mere description of the new way of answering as ‘clever’ need not indicate that it consists in discovering something previously unknown.


30 For recent examples of interpretations that continue to rely on the authority of Aristotle’s remarks, see Mueller, ‘Platonism’, 78 ff.; and Kelsey, ‘Causation’, 21.
In this case, the truth lies somewhere in the middle. Aristotle does distort the passage, but the distortion is not due to interpretative inattention, let alone to a will to falsify what Socrates says. Rather, Aristotle insists on reading the *Phaedo* from the vantage-point of his own philosophical concerns. He approaches the passage with his own theory of explanatory adequacy firmly in mind, and he is concerned to point out how Plato’s work falls short when measured against the Aristotelian standard. As we shall see, Aristotle is in fact relatively explicit that this is how he proceeds. It is mainly the expectation of a more dispassionate commentary that has fostered the interpretative extremes mentioned above.

Aristotle’s most extended discussion of the *Phaedo* is found at *GC* 2. 9, 335\textsuperscript{b}8–16:

> The third [source of coming-to-be, the efficient cause] must also be present. Everyone dreams of it, but no one actually states it [ἅπαντες μὲν ὀνειρώττουσι, λέγει δ᾿ οὐδείς]. Instead, some have deemed the nature of the Forms to be sufficient as an explanation [αἰτίαν] of coming-to-be. The Socrates of the *Phaedo* is like that: after blaming his predecessors for having said nothing coherent, he lays it down that among the things that exist there are Forms, on the one hand, and things that partake of Forms, on the other [τὰ μὲν εἴδη τὰ δὲ μεθεκτικὰ τῶν εἴδων]; and further, [he lays it down] that each thing is said to be in virtue of a Form, to come to be in virtue of taking a share in it, and to perish in virtue of losing it [εἶναι μὲν ἑκαστὸν λέγεται κατὰ τὸ εἴδος, γίνεσθαι δὲ κατὰ τὴν μετάληψιν καὶ φθείρεσθαι κατὰ τὴν ἀποβολήν]. If this is the case, he must think that the Forms are responsible for generation and destruction [ὡστε ἐι ταῦτα ἄλληθ, τὰ εἴδη οἴεται ἐξ ἀνάγκης αἰτία εἶναι καὶ γενέσθαι καὶ φθοράς].

The final statement is the crux of the interpretative difficulty, so let me begin with it and work backwards. As I translate, it is Aristotle who draws the inference that the Forms are efficient causes. Yet the statement is typically read otherwise, as reporting an inference that Socrates makes from the doctrines mentioned just before. In the revised Oxford translation, for instance, Joachim renders the lines as follows: ‘Hence he [Socrates] thinks that assuming the truth of these theses, the Forms must be causes both of coming-to-be and of passing-away.’\textsuperscript{31}

\textsuperscript{31} Joachim reinforces that in his edition of the text by remarking: ‘Aristotle is still paraphrasing the *Phaedo*’ (*Aristotle on Coming-to-Be & Passing-Away* [Aristotle] (Oxford, 1922), 249). A similar translation is provided by C. J. F. Williams (in his
Were Aristotle really asserting that Socrates himself deems the Forms to be causes of coming-to-be, his interpretation could be set aside for want of compelling evidence. But in fact Joachim’s alternative is less likely given the wording of Aristotle’s statement, since the placement of the verb οἴεται strongly suggests that it belongs only to the apodosis of the conditional statement and does not govern the statement as a whole. Even more significantly, the idea that Socrates explicitly identifies the Forms as efficient causes makes for a poor reading of Aristotle’s initial remark that everyone ‘dreams’ the efficient cause but nevertheless fails to ‘state’ it. Aristotle surely does not think that Socrates or anyone else adequately isolates the idea of an efficient cause. What he means instead is that each of his predecessors understands the need for such a cause only vaguely and commits himself to an account of it only indirectly, by the way he speaks of his favoured explanatory principle as if it could account for coming-to-be (and passing-away) tout court.32

Is that proposal borne out by the other things Aristotle says about Plato’s Socrates? His penultimate remark is that for Socrates a thing comes to be in virtue of participating in a Form and perishes by losing it. In so speaking, he must have in mind what is said at Phaedo 101 c. The noun that Aristotle uses for the participation-relation—μετάληψις—might be understood as an attempt to capture Socrates’ ingressive use of μετάσχεσις/μετασχεῖν, though one need not presume as much.33 Even if Aristotle is not being careful to

Clarendon commentary) and by E. S. Forster (in the Loeb edition). What I argue to be the correct translation is given by T. Irwin and G. Fine in their edition of Aristotle (Aristotle: Selections (Indianapolis, 1995), 162).

32 A roughly similar interpretation of Aristotle’s testimony is offered by Annas, who suggests that Aristotle means to exploit a confusion on Socrates’ part concerning the nature of his new αἴτια (see ‘Inefficient’, 318, 324–5). Although Annas prints Joachim’s translation and paraphrases it in the course of her discussion, she none the less seems to interpret the force of Aristotle’s remarks in a manner closer to the translation I defend. See p. 324: ‘Aristotle says of Socrates that if what he says is true, then he thinks that Forms must be efficient causes (335 B 15–16) . . . Aristotle is saying that the Phaedo passage commits Plato to making Forms efficient aitia; this may not have been Plato’s intention, but . . . [Aristotle] is abstracting from what he can guess of the author’s intention, and concentrating on what the text actually says.’ Compare Annas, Aristotle’s Metaphysics Books Μ and Ν (Oxford, 1976), 162.

33 Cornford supposes (Parmenides, 69 n. 1) that Aristotle’s μετάληψις is in fact equivalent to Socrates’ μετάσχεσις. His judgement is based partly on what he deems to be a similar use of μετάληψις at Plato’s Parm. 131 a 4. However, I find the present context much too indefinite to determine whether Aristotle in fact uses the word as Cornford thinks.
note the precise point that Socrates is making, there is no reason to think he goes to the opposite extreme of reading *Phaedo* 101c as if it were maintaining that the Forms are causal agents. In order to understand what motivates Aristotle’s subsequent judgement that Socrates must think of the Forms as efficient causes, we need to go back a bit further, to the immediately preceding lines. There, Aristotle observes that after criticizing his predecessors Socrates ‘lays it down that among the things that exist there are Forms, on the one hand, and things that partake of Forms, on the other’. With that observation, Aristotle is pointing out that Socrates admits only Forms and participants as the components of his new explanatory scheme. In all likelihood, he is thinking of 100c–d, where Socrates clearly states that he is abandoning all other theories and is relying in their stead solely on his talk of participation in a Form. Given what is said at 100c–d, and given the further fact that at 101c ‘coming-to-be’ is said to involve participating (or perhaps coming to participate) in a Form, Aristotle goes on to infer that Socrates is committed to viewing the Forms as efficient causes. Socrates may not say anything so explicit. But since the only remaining components of his explanatory scheme are participants, and since the latter cannot plausibly be deemed causes of their own coming-to-be, Socrates would have no choice but to allow that the Forms are the efficient causes of change. 

When the passage is understood that way, Aristotle’s remarks about Socrates are similar to the criticism he goes on to make of certain materialists, who likewise ‘dream of’ but do not ‘state’ the

---

34 That, I think, is the ultimate force of Aristotle’s initial judgement that Plato takes the nature of the Forms to be sufficient as a cause of coming-to-be (λειτυρ... αἰτίαν ἴνα πρὸς τὸ γίνεσθαι). According to Fine (*Forms*, 361–2), what Aristotle says does not in fact constitute a charge that Plato’s Socrates (effectively) treats the Forms as efficient causes. On the contrary, Aristotle is actually acknowledging that Plato sees no reason to mention such causes at all: ‘[Plato] might . . . believe that the existence of Efficient-causes is not a necessary condition for change. This is just what Aristotle’s introduction to the argument suggests he thinks Plato believes’ (162). That is a clever reading, but it is difficult to understand what sense it makes of Aristotle’s remark that Plato, among others, ‘dreams’ of the efficient cause. Bolton defends a similar interpretation by suggesting that Aristotle’s remark be understood as ‘a gibe, which plays on Plato’s own denigration of those who cannot recognize Forms as mere dreamers about true reality’ (*Discovery*, 98). Yet, Aristotle does not say simply that Plato and others are dreamers about reality, or about the true dimensions of an adequate causal theory. Aristotle is speaking about the efficient cause; and I do not see how what he says can be squared with any view that Plato openly rejects the need for efficient causes.
efficient cause. It is not that they have reflected on the idea of efficient causation and have thoughtfully concluded that matter may serve as an ultimate principle of activity in the manner required of a proper Aristotelian efficient cause (on which see 324.24 ff.). Rather, they maintain broadly that ‘from [matter] comes movement [ἀπὸ ταύτης [sc. τῆς ὕλης] γὰρ εἶναι τὴν κίνησιν]’ (335.16–17, cf. 326.9); and although what they expressly have in mind may simply be that various kinds of change can be understood in terms of the relative arrangements and motions of bits of matter, they nevertheless neglect to mention any other causal factors and thereby end up committing themselves to the proposal that matter actually initiates change.

If Aristotle’s remarks on the Phaedo are read as I am proposing, the conclusion drawn at lines 15–16 can readily be understood as part of his larger interpretative strategy. Aristotle may well recognize that Socrates means to concentrate on some sort of explanation other than an efficient one. He may even realize that Socrates is not trying to advance a causal theory at all and is instead interested in a notion of metaphysical explanation. However, Aristotle is concerned to discuss the way in which the passage relates to his own account of what is involved in an understanding of coming-to-be. Looking at the Phaedo from that perspective, Aristotle notes that in so far as Socrates eschews all explanatory devices other than the Forms (and their participants), he would have no way of explicating the process of change except by conceding that the Forms are efficient causes. In other words, Aristotle is commenting on what he takes to be the import of what Socrates says, not on what may be Socrates’ intentions in speaking as he does.

Taken that way, Aristotle’s remarks in On Generation and Corruption fit neatly with his well-known position in the Metaphysics.
that Plato explicitly acknowledged only the formal and the material cause in his expositions of the theory of Forms. Aristotle is not now changing his mind about Plato. Nor is he showing himself to be a poor reader of the dialogues. Instead, his comments are part of his familiar tendency to arrange material from his predecessors so as to show how it prefigures dimly the details of his own theories, or so as to set problems that are to be resolved by his theories. That is to say, Aristotle is interested solely in how the position taken in the *Phaedo* may be said to anticipate the ‘truth’ — the insight expressed by his own account of adequate explanation. The fact that his interpretation may not be faithful to Plato’s intent is of little moment.

If that reading is correct, Aristotle’s remarks do not tell in favour of any one interpretation of Socrates’ preferred *aitia*. And thus, there is no call for us to surrender the conclusion reached in Section II, namely that Socrates is not in fact concerned to explain how changes occur and is interested solely in giving an account of what obtains when something has come to be *two*, or *large*, or whatnot.

So far, then, what we have is a mystery. Socrates’ description of his new *aitia* seems to conflict with his earlier declarations that he is concerned to enquire generally into the explanation of coming-to-be, passing-away, and being \((96\alpha; 97\beta, c)\). Could Socrates be unaware of the conflict? And if he is aware of it, how could he ever think himself justified in adopting a mode of explanation that renounces all attempts to understand the process of change?

There is an answer to be had here, and it is one that shows Socrates to be more clearheaded than has often been supposed. Let

---

37 See \(A\ 6, \ 988'9–10\), and compare \(A\ 7, \ 988'34–'6\). Compare also \(992'25–'6\): ‘we [Platonists] say nothing about the *aitía* that is the origin of change’. Bolton provides a forceful review of Aristotle’s testimony in the *Metaphysics* (see ‘Discovery’, \(95\beta\)), but he goes wrong in supposing that some of Aristotle’s critical remarks in \(A\ 9\ (991'8–'11, \ '3–'9)\) supply clear evidence for the judgement that Aristotle never considers the Forms to have been understood as efficient causes. Aristotle can just as well be read as arguing that the Forms do not in fact play the role of efficient causes, which role they might well be expected to play given certain details of Plato’s theory, or given the way the Forms are spoken about in particular passages from Plato’s dialogues. That is precisely how I would interpret the discussion of the *Phaedo* at \(991'3–'9\).
me begin elaborating it by way of some brief remarks on the nature of the new aitia. As I have argued, interpreters who take Socrates to be concerned with some notion of metaphysical explanation are quite right to think as they do. But more needs to be said about what precisely is at issue.

Consider for a moment the famous characterization offered by Vlastos in his 1969 essay on the passage. As mentioned previously, Vlastos holds that, at least in the ‘safe’ version in which it is initially presented (cf. 100d–e), the new aitia is a purely ‘logical’, or ‘metaphysical’, or even ‘logico-metaphysical’ mode of explanation. What he means in characterizing it as ‘logical’ is that it consists in specifying the ‘definitions’ of various general terms and, thereby, in identifying the conditions for applying the terms to things that satisfy those definitions. But as Vlastos also holds, the conditions for applying a term do not derive simply from conventional usage. Instead, Socrates is concerned with what are sometimes called ‘real definitions’—linguistic accounts that describe the ‘essences’ of certain entities. Accordingly, Vlastos describes Socrates’ mode of explanation as at once metaphysical in character.38

For Vlastos, however, the metaphysical role of the new aitia is none the less subsidiary: ‘it is the logical function of the metaphysical entity that does the explanatory work of the “safe” aitia’.39 Unfortunately, when Socrates introduces his aitia, he makes no mention at all of a concern with definitions. At 100c, for example, he proposes to explain a thing’s being beautiful solely in terms of its participation in a Form.40 And although he had previously linked the talk of Forms to his concern with definitions (see esp. 75c 10–d 5; 78d 1–7), he now characterizes his reliance on the new aitia in famously self-deprecating fashion, as being ‘simple and artless

38 See e.g. ‘Reasons’, 91–2: ‘What is [Plato’s] Theory of Forms if not the claim that logical statements presuppose metaphysical ones and would be mumbo-jumbo without them? . . . The reason why we can speak significantly and truly of things being square or beautiful, he would insist, is that there exists an incorporeal, immutable, intelligible object, named “Squareness” or “Beauty”, in which corporeal, mutable, sensible objects occasionally “participate” and, when they do, are rightly called “square” or “beautiful”.’

39 Ibid. 92.

40 See Bostock, Phaedo, 149–51; and G. B. Matthews and T. A. Blackson, ‘Causes in the Phaedo’, Synthese, 79 (1989), 381–91. The latter go too far in maintaining that Socrates explicitly rejects any concern with definitions in what he says at 100c–101c. Rather than renounce definitions per se, Socrates turns away only from the attempt to supply reductive analyses of certain concepts, as I shall explain in what follows.
and perhaps simple-minded’ (100 d 3–4; cf. 105 c 1). Such a characterization would be difficult to reconcile with any proposal that Socrates has the nuances of a Form’s epistemic role prominently in mind.\footnote{In discussing what is said at 100 d, Vlastos tellingly appeals only to a Form’s metaphysical role (‘Reasons’, 94–5).}

Accordingly, it is best to suppose that Socrates is thinking in terms of a wholly metaphysical mode of explanation.\footnote{Just such an interpretation has also been adopted by Bolton, who presents it not as a way of remaining true to (some of) Vlastos’s insights but, rather, as an alternative to Vlastos’s proposal, late in his paper, that although the ‘clever’ version of Socrates’ αἰτία (cf. 105 b–c) is still ‘logico-metaphysical’ in character, it nevertheless has implications for the development of a causal theory. (See Vlastos, ‘Reasons’, 104–5, and Bolton, ‘Discovery’, 92–4.) Bolton explains the idea of a metaphysical αἰτία by referring simply to Aristotle’s doctrines in the \textit{Categories} and \textit{Metaphysics}.}
The task of describing what it consists in is complicated by the much-observed fact that he introduces the Forms without mentioning any arguments that would justify his turn to them. They are simply familiar to his interlocutors (ἐκεῖνα τὰ πολυθρύλητα, 100 b 4–5). I think that a good reason can be given for such a curious way of proceeding. It has to do with Socrates’ desire to emphasize the continuity of his larger concerns rather than to focus on the unique aspects of his own approach. Elaboration of the point will have to await an account of how precisely we should understand those larger concerns; but for now, let me explain briefly what I take the αἰτία to involve.

In what he says at 100 b ff. Socrates is thinking of the Forms as helping to supply the \textit{truth-grounds} or \textit{truth-makers} for certain statements of the language—paradigmatically simple subject–predicate statements such as “Charmides is beautiful” or “Simmias is large”. That is to say, Socrates is effectively beginning from linguistic items—true statements—that he regards as data for analysis. He wants to know what it is in virtue of which those statements are true, and he thinks of the Forms as the crucial elements of an adequate account. His reasons for so thinking are not discussed in the passage; but in other texts he holds that a Form is what supplies the explanation of how one general term can apply to a multiplicity of objects. The Form does so by being the single entity to which each of those objects is related, in which each ‘participates’.\footnote{For that way of thinking, see especially \textit{Rep}. 596 b 6–7, along with my ‘On \textit{Republic} 596 s’, \textit{Apeiron}, 39 (2006), 27–32. Cf. also \textit{Parm}. 130 b 3–131 a 2 and \textit{Rep}. 507 b 1–6. For further discussion of the conception of explanation at issue here, see my ‘The Anatomy of an Illusion: On Plato’s Purported Commitment to Self-Predication’, \textit{Apeiron}, 40 (2007), 159–98.}
The idea that Socrates begins from true statements might initially seem unlikely, since he introduces the new *aitia* at 100c–d by saying only that he is concerned with the explanation of a thing’s ‘being beautiful’. But note some further details of the surrounding context that may be adduced in support of such a reading. At 102b1–2, Phaedo sums up what has been said so far with the remark: ‘It was agreed that each of the Forms exists and that all other things have their names [αὐτῶν τούτων τὴν ἐπωνυμίαν ἴσχειν] because they partake of these Forms.’ Similarly, at 102c10 Socrates goes on to note that Simmias ‘is said to be both short and tall [ἐπωνυμίαν ἔχει σμικρός τε καὶ μέγας εἶναι]’ in virtue of certain relations between objects and Forms. In each case, the speaker is effectively recasting an ordinary judgement (“Charmides is beautiful”, “Simmias is short”, etc.) as if it expressed a relation between a linguistic expression—a general term—and an entity. The application of the general term to that entity is taken to be explained by the latter’s participating in a certain Form, which is named in some basic fashion by the general term.

At first glance, the preceding passages might seem to be mere periphrases, colourful but pointless circumlocutions with respect to some more primary way of describing Socrates’ concerns. Yet they are in fact much more: they will eventually prove crucial when it comes to looking broadly at the argument of the passage and reconciling the metaphysical *aitia* with Socrates’ early descriptions of what interests him. For now, let me confine myself to pointing out that one could hardly elaborate the metaphysical *aitia* by saying without further comment that it is concerned with ‘what is the case’, with being as opposed to coming-to-be. After all, Socrates’ predecessors were also interested in ‘being’ no less than in coming-to-be and perishing. And in his early attempts at similar sorts of enquiry, Socrates himself puzzled no small amount over the constitution of states of affairs—for instance, over questions as to why one thing is larger or greater than another (cf. 96d–e, to be discussed below). Socrates cannot readily differentiate himself from his predecessors by pointing to a distinct sphere of concern.

---

Compare also the discussion from 103e to 104b, where statements such as “Three is odd” are analysed in terms of the application to an entity (three, let me say) of a general term (“odd”). (Throughout that passage Socrates understandably has difficulty when it comes to the analysis of relational and numerical expressions. Why he chooses to rely on examples involving such expressions will be discussed shortly.)
He needs instead to describe a different methodology, a different way of pursuing the enquiry into being. That is precisely what is promised by his remarks at 97 b 6–7 (discussed in Section I).

If the foregoing proposal is correct, the problem of interpreting the ‘autobiography’ may now be cast as follows. In what sense might the project of analysing the truth-grounds for statements be deemed an alternative to Socrates’ earlier preoccupations—an alternative, moreover, that would allow Socrates to say he is still pursuing the explanation of ‘coming-to-be and passing-away and being’? Let me now address the problem directly by looking more fully at the development of the ‘autobiography’. To start, I shall need to examine in some detail Socrates’ criticism of his materialist predecessors. That will allow for a new understanding of Socrates’ overall concerns, one that will help explain his subsequent interest in teleology as well as his eventual turn to metaphysical analysis. As will emerge, Socrates thinks of the latter mode of explanation as the only way of addressing a deficiency that he considers inescapable for any materialist theory of the world.

V

Beginning at 96 c 3, Socrates describes how his youthful pursuit of natural science produced in him such confusion that he came to reject a variety of ordinary explanations for phenomena—explanations that had seemed obviously correct before he embarked on his scientific speculations. He was once convinced, for instance, that a man grows ‘by the addition of flesh to flesh and bone to bone’, and in general that a thing becomes large by the addition of a like material (96 c 8–d 5). Furthermore, he had once deemed it correct to say that one man is larger than another ‘by a head’, that ten things are greater than eight ‘in virtue of the accruing to them of two’, and that a two-foot length was greater than a foot-long one ‘in virtue of its exceeding the other by half its own length’ (96 d 8–e 4).\(^\text{45}\) Now,

\(^{45}\) Pace Vlastos, ‘Reasons’, 95 n. 50, I do not think that Socrates means to begin a new line of thought with the latter three examples. Those examples are in fact further cases of purportedly ordinary explanations that Socrates came to reject once he had immersed himself in scientific enquiry. Vlastos contends that the words immediately following the initial remarks at 96 c 2–d 5—namely, οὕτως τότε /ΑIΛbiΑταὤμην· οὐ δοκῶ σοι μετρίως; (d 5–6)—are meant to mark a contrast with what is said subsequently (at d 8–e 4) by giving ‘a seal of retrospective approval’ to the example used in the
however, he can no longer accept those accounts, nor can he find an adequate substitute within the realm of materialist explanation.

Commentators have been troubled by Socrates’ remarks, and understandably so. First of all, it is unclear why he should think that his scientific pursuits are threatening to his ordinary beliefs, especially since several of the latter have nothing obviously to do with causal explanation. Secondly, Socrates fails to explain just why he thinks that natural science conceived along materialist lines cannot offer any alternative explanation of the phenomena in question—or at least, of those among them that are suitable for causal explanation in the first place. As a result, a common reaction to what Socrates says has been to presume that he is somehow confusing radically different kinds of explanation and that he has no real objection to natural science per se. 46

Socrates’ remarks are hardly a model of lucid exposition. Still, much of the temptation to reject what he says as trivially confused disappears if one surrenders a presupposition common to many interpreters—namely, that Socrates is trying to attack materialism by challenging some concept of causation on which he thinks his predecessors relied. Instead, Socrates has another target altogether: he is attacking in broad fashion the materialist ideal of a reductive explanation. That is to say, he presumes that the materialist will claim to be able to explain all observed phenomena in terms of certain basic elements and processes, as if all ‘ordinary’ statements of fact could be translated into the vocabulary of the materialist first passage but not to those in the second. However, Socrates in fact stresses the continuity between the two passages by beginning the second with words closely echoing the ones that Vlastos quotes: ὅτε δὲ καὶ τάδε ἔτι. Αὐτῷ ὤμην γὰρ ἱκανῶς μοι δοκεῖν . . . . His point in both cases is simply that the examples he is discussing had once seemed unquestionable to him and would still seem that way to most people. Cf. 96 c 3–5 and Gallop, Phaedo, 171–2. (The issue is crucial to Vlastos’s understanding of the larger role played by Socrates’ criticisms of his predecessors; for further comment see sect. 1 and n. 24.)

46 Thus, Vlastos proposes that in his youth Socrates mistook conceptual questions for ones demanding a physical account (as discussed in sect. 1). Hackforth (Phaedo, 131), Crombie (Examination, 169), and Bostock (Phaedo, 141–2) all charge that Socrates simply jumbles together causal questions and those of other sorts (whether they be questions as to the degree of difference between two things or else ones appropriate to the philosophy of mathematics). Gallop tries to rescue Socrates at the price of rejecting any suggestion that Socrates is concerned with scientific speculation as such: his claim to have been blinded by scientific study is ironical, since ‘in reality, his problems are conceptual’ (Phaedo, 172).
theory. Socrates objects that for certain statements, such reductions are impossible.47

The examples he cites all concern cases in which the application of an expression such as “large” or “larger” is to be given a reductive analysis. In the first group, he effectively entertains the view that “becomes great/large” (πολύς/μέγας γίγνεσθαι, 96 D 4–5) may be explained according to the principle of like being added to like. In the second, he proposes understanding “is greater/larger” (μείζων εἶναι, 96 E 2, 4; πλείων εἶναι, E 2–3) in terms of some concrete measure in virtue of which one thing exceeds another. As mentioned above, the proffered explanations are not drawn from any particular Presocratic theory. They are supposedly ones that everyone would find acceptable, at least pre-theoretically. I imagine Socrates structures the discussion that way because none of the theorists with whom he is familiar actually undertook a reductive analysis of the term “large” (or of the phenomenon of largeness). When Socrates formulates his objection to reductionism, it is he himself who hits on the idea of attempting to analyse “large” reductively and of using that attempt as a problem case for materialism. Thus, he crafts his own examples, which he cleverly presents in the guise of everyday beliefs that he has been compelled to ‘unlearn’ (96 c 6).

Although the sample explanations are drawn from common ways of speaking, Socrates does not mean to imply that those explanations would be regarded ordinarily—by non-scientific speakers of the language—as samples of materialist analysis. He is maintaining only that after he had immersed himself in scientific study, such ways of speaking took on the air of materialist reductions.48 And

47 With this way of putting the issue, I do not mean to suggest that at 96 c ff. Socrates is thinking explicitly of words or statements rather than of ‘things’ or ‘phenomena’. The linguistic recasting simply helps to make the point clearer.

48 The example closest to everyday speech—that one man who is larger than another is larger ‘by a head’—is in fact the most unfit to be interpreted as a proposal for reductive analysis. That purportedly ordinary explanation is clearly the expression of a degree of difference, and nothing more. Yet I take Socrates’ point to be that once he began his scientific study, he could not help but regard that way of speaking as a proposal that ‘being larger’ should be explained in terms of a certain concrete mass, ‘a head’. Thus at 101 A–B, when he explains how his new mode of explanation escapes the problems of the scientific one, Socrates readily speaks as if according to his previous way of thinking it was quite literally a material thing—a head, and not just a unit of measure—that was responsible for largeness. The example is still strange, to be sure, but I think Socrates realizes its strangeness and is intentionally offering it as a comical illustration of his point. Note that at 101 B 3 Cebes laughs at what Socrates says.
once materialism came to seem troubling, he could not shake the idea that there was something wrong even with those ‘common’ forms of explanation. He ultimately became so confused that he was led to reject his prior beliefs altogether.

In the course of his remarks at 96c–e, Socrates does not discuss what troubles him about the idea of analysing “large” reductively. Yet, he goes on to describe what he finds wrong with an analogous and purportedly more basic example (96e 6–97b 3):

By Zeus, said [Socrates], I assure you that I’m far from presuming I know the explanation of any of these things, seeing as I can’t even accept, when someone adds one thing to another [ἐπειδὰν ἑνί τις προσθήκη ἐνὶ], either that the one to which the adding is done has become two, or that the one added has done so, or that the one to which the adding is done and the one added become two together on account of the adding of the one to the other. It would be astounding if when they were apart from one another [ὁτὲ μὲν ἑκάτερον αὐτῶν χωρὶς ἄλληλων ἦν], each was in fact one and they were not at that point a pair, but when they drew near one another [ἐπλησίασαν ἄλληλοι] this turned out to be the explanation of how there come to be two—the combination consisting in their being put close together [ὅταν σύνοδος τῶν πληγῶν ἄλληλων τεθημήν]. Nor, moreover, can I any longer believe that if someone splits one thing [κόπω τις ἐν διασχίζῃ], this in turn is the explanation of there having come to be two—the splitting. For in that event, there results an explanation of there coming to be two that is opposite to what was the case before: previously, it was because they were moved close together and one was added to the other [ἐπιθηκέται πληγῶν καὶ προσετίθετο ἄλληλον ἄλληλον ἔτερον ἔτερον]; now, it’s because they are being moved apart and one is being separated from the other [ἀπαγήσεται καὶ χωρίζεται ἄλληλον ἄλληλοι].

49 ἢ τὸ προστεθέν, a conjecture by Wyttenbach. The addition is not essential, but it is neat (a plausible case of haplography) and allows Socrates to canvas all the relevant possibilities. Against it, W. J. Verdenius has protested that ‘from a primitive point of view the thing to which something has been added has been enlarged, and not the reverse’ (‘Notes on Plato’s Phaedo’ [‘Notes’], Mnemosyne, 4th ser. 11 (1958), 193–243 at 229). But Socrates is attempting to drive home a paradox with regard to a certain theoretical account. In doing so, he need hardly be considered to speak from some ‘primitive’ common sense, if indeed there be such a thing.

50 ἐναντία γὰρ γίγνεται ἢ τότε αἰτία τοῦ δύο γίγνεσθαι, adopting (as do most editors) Heindorf’s ἢ instead of the ἢ of the manuscripts. The defence of the latter by Verdenius (‘Notes’, 229) and R. Loriaux (Le Phédon de Platon: commentaire et traduction (Namur, 1969–75), ii. 74) is unconvincing: there is nothing awkward in the lack of an article before ἄλληλοι, and in any case the emendation makes for a reading that is much less forced than one whereby Socrates maintains that the previous explanation actually transforms into its opposite (whatever precisely that might mean).
Socrates is here concerned with a physical phenomenon, the coming-to-be of two things, and not with an arithmetical operation considered in the abstract. He insists that the phenomenon of ‘twoness’ cannot be explained in terms of any element in the process of one thing’s being placed beside another. Nor can it be explained in terms of the process itself, a central objection being that an ‘opposite’ process—that of separation—would have to be cited in order to explain the very same phenomenon in other circumstances.

Socrates does not make it clear just why he considers that a problem for the materialist. Commentators have rightly seen his remarks here as part of a broad tendency to reject explanations involving ‘opposites’ (cf. 101 a–b), but the source of that tendency has remained murky.

I want to offer a suggestion that I believe will help

---

51 That is shown by the highly physicalistic language throughout the passage. There is little to recommend the proposal that Socrates is here confusing arithmetical and physical talk (Crombie, Examination, 169), or else that he is self-consciously exposing the perils of confusing the two (Vlastos, ‘Reasons’, 97). The only justifications for thinking that an abstract arithmetical operation is at all of concern here are (a) Socrates’ use of ἕν and δύο without further specification, and (b) his use of προστίθεσθαι, which would be the normal way of referring to the operation of addition. But as regards (a), Socrates does not qualify his use of number-terms because he wants it to be clear that his objection applies generally, to any talk of physical things, and he finds the neuter adjective sufficient for that. As regards (b), προστίθεσθαι of course also means “to be placed next to”; and in the absence of any other references to mathematics, there is no need to suppose it has the more abstract sense here.

52 At the outset of the passage, “two” is used as a predicate-term. But at 97 a 4–b 1 Socrates speaks simply of ‘two’s coming to be’ (τὸ δύο γενέσθαι (a 5), etc.), where (I take it) “two” is the subject of the infinitive. Socrates speaks in that fashion precisely in order to avoid identifying anything that comes to be two and, thereby, to call attention to the question under discussion—namely, whether processes of association or separation might be responsible for twoness.

53 The idea of opposition at work here and elsewhere in the passage is a loose one. What Socrates seems to have in mind is some idea of incompatible processes. For helpful comment, see R. J. Hankinson, Cause and Explanation in Ancient Greek Thought [Cause] (Oxford, 1998), 84–90.

54 A common approach has been to regard Socrates as relying on what has been called the ‘synonymy principle of causality’—roughly, the principle that one can explain a thing’s having a given quality only by reference to something that has the quality in equal or greater degree. For the principle, see J. Barnes, The Presocratic Philosophers, rev. edn. (London, 1982), 119 (cf. 88–9), and compare A. C. Lloyd, ‘The Principle that the Cause is Greater than its Effect’, Phronesis, 21 (1976), 146–56. For applications of some version of it to the Phaedo, see e.g. E. L. Burge, ‘The Ideas as Aitiai in the Phaedo’, Phronesis, 16 (1971), 1–13 at 5; Annas, ‘Inefficient’, 316; Bostock, Phaedo, 155–6; Hankinson, Cause, 91–2; Sedley, ‘Causes’, 123–4; and Dancy, Introduction, 301 fl. Kelsey (‘Causation’, 24) protests against the tendency to see the synonymy principle at work in our section of the Phaedo. As will emerge, I agree that it is not in fact what guides Socrates’ thinking.
Socrates’ New Aitia

illuminate the structure of the ‘autobiography’ as a whole, from Socrates’ attack on the reductive explanations of the materialists to his subsequent engagements with teleology and metaphysics. The suggestion concerns a general epistemic requirement that Socrates places on all talk of aitiai, whatever may be their nature. It is that in order for an aitia to be considered adequate, there must be an ‘intellectual transparency’ to the relationship between explanans and explanandum, in the sense that a proper description of the former will allow one to understand the occurrence of the latter. That is to say, an adequate aitia must enable one to see how the phenomenon to be explained follows from the terms of one’s theoretical account.55

It has been a theme of much recent scholarship, particularly in connection with Aristotle, that an epistemic requirement is at work in certain philosophical uses of the noun aitia (and the corresponding adjective aition).56 It is my argument that some such requirement is also at work in the present passage and that it is the source of Socrates’ tendency to presume that a proper explanans is broadly coextensive with its explanandum and, ultimately, that there is something deeply troubling about explanation by ‘opposites’. In the case of materialism, Socrates applies the epistemic demand by supposing that the materialist is committed to showing how a higher-level phenomenon may be said to ‘fall out’ from the details of a lower-level, mechanistic story.57 His objection is that for cer-

55 I do not mean to say that there needs to be an inferential connection between explanans and explanandum, such that a given explanans will in every case give rise to the explanandum in question. Rather, the connection between the two may well be a defeasible one. (On the point, see Hankinson, Cause, 91.) Unfortunately, Socrates does not say enough here to provide a basis for venturing a more precise account of the explanatory relation.

56 Thus, the Aristotelian αἰτίαι discussed at Physics 2. 2 and Post. An. 2. 11 are generally acknowledged to be not merely causes (in some post-Humean sense) but explanations, perspicuous answers to certain why-questions. For further discussion, see the judicious remarks on Aristotle in Annas, ‘Ine¶cient’, 319–23 (along with the further references in her note 19). For a recent account of the varying treatments of Platonic αἰτίαι, see D. Wolfsdorf, ‘Euthyphro 10 Α 2–11 Β 1: A Study in Platonic Metaphysics and its Reception since 1960’, Apeiron, 38 (2005), 1–71 at 52–62.

57 Such an idea may well have played a role in some of the materialistic theories with which Plato was probably familiar. Consider Democritus’ crude but suggestive attempt to explain the sense of taste with reference to features of the atoms that come into contact with the tongue: ‘Democritus . . . makes sweet what is round and large in size; sour what is large, rough, polygonal, and not rounded; sharp, just as the name suggests, what is sharp in body and angular and bent and thin and not rounded . . .’ (Thphr. CP 6. 1. 6–6. 8 = 129 A 129 DK; cf. Thphr. De sens. 65b–67 = 125 A 125 DK). There is of course a significant explanatory gap between the features of the atoms and the experience of sweetness, sharpness, etc. But what is interesting here is Democritus’
tain phenomena the reductionist project inevitably fails, since one cannot supply an account that would appropriately link \textit{explanans} and \textit{explanandum}. In the passage quoted above, Socrates finds several difficulties in the proposed explanation of how there come to be two things. First, there is no adequate epistemic connection between the process of association and phenomenon of twoness. Second, and even more seriously, if indeed the latter phenomenon may be explained equally well by an ‘opposite’ process (the separation of one thing from another), then there can be no connection of the appropriate sort. The problem here is not merely that there is an incompleteness in the proffered explanation. It is that the very idea of an intelligible explanatory connection has dramatically been violated. After all, if one were to allow explanations by ‘opposite’ processes in different circumstances, one would effectively be supposing that the phenomenon in question might be reduced to a disjunction of incompatibles. And in that case, there would be no hope of linking \textit{explanans} and \textit{explanandum}. Since Socrates can think of no further theoretical refinement that would address the problem, he takes the example of twoness to constitute a deep difficulty for the whole project of materialist explanation.\footnote{On that reading, one can readily see the relevance of the objection concerning twoness to the task of explaining how one thing becomes larger or greater than another. If Socrates cannot explain how ‘two’ comes to be from ‘one’, he certainly cannot explain what it is for anything to become large as a result of some process of aggregation. The idea that things become large by aggregation is precisely the attempt to minimize the gap by telling the atomic story in a way that would readily accommodate our descriptions of our perceptual experiences. Also instructive in this connection is the Democritean account of hardness and heaviness, where the analysis is presented in terms of structural features of a thing’s ‘molecular’ composition rather than in terms of features of the component atoms: ‘What’s compact is hard, what’s loose is soft . . . Yet the position and arrangement of empty spaces differs somewhat in the hard and soft from in the heavy and light. That is why iron is harder but lead heavier. Iron is unevenly composed and has a good deal of empty space in many places, though it is compact in parts . . . . Lead has less empty space and is evenly composed and uniform throughout. That makes it heavier but softer than iron’ (Thphr. \textit{De sens.} 62–68 A 135 DK).}

Socrates cannot, of course, demonstrate that there are no other ways of explaining twoness; but he clearly thinks of association and division as exhausting the likely candidates. His practice of casting his remarks autobiographically—as a story of how he was forced to abandon materialism rather than as a story as to why the materialist project must fail—neatly allows him to avoid the issue of whether the proponent of materialism might ever come up with a compelling response.
sort of assumption that Socrates had made in his youth, but his forays into materialism have now led him to question it. For, the use of the phrase “becomes large” cannot be explained with reference to one or more of the elements involved in the process. Nor can it be explained in terms of the process itself, since in certain circumstances (when “large” is used in the sense of “numerous”) a process of division or separation might conceivably be said to supply the explanation.

The other sort of youthful assumption that Socrates has now been led to abandon involves accounts of why one thing is larger or greater than another. Here, the relevance of the objection concerning twoness is less easy to see. Socrates returns to the matter later on, however, when he shows how his new aitia avoids the pitfalls of reductive analysis (100b–101b). He actually mentions two problems with a reductive account of “is larger”, the first being that whatever is introduced to explain that phrase can at once be said to explain “is smaller”. After all, that measure (or material factor) in terms of which one thing may be said to exceed another is precisely the measure (or material factor) in terms of which the latter falls short of the former. Yet a single explanans cannot coherently be used to account for opposite phenomena—or so Socrates maintains. The second problem is that in some cases the materialist will be forced to explain one thing’s being larger than another by means of an explanans that is in fact opposite to the phenomenon of largeness. As Socrates playfully puts the point, a small factor—such as a head—will in some cases be used to explain the largeness of a large thing (or more precisely, to explain why that thing is larger than something else). And Socrates finds that absurd.

Once again, with the emphasis on opposition, he is pointing to cases in which there is a failure of explanatory connection between explanans and explanandum. From a modern standpoint, his examples are frustrating in so far as they display an ignorance of the logic of relations—or in the example concerning twoness, of sets—along with a correlative tendency to neglect any considerations of the context in which one makes judgements of largeness, or twoness, or whatnot. But in attempting to understand Socrates’ point, it is important not to be too hasty in passing over the examples he uses. It is precisely what we see as their limitations that

59 On some ambiguities in Socrates’ discussion of the point, see the helpful remarks in Gallop, *Phaedo*, 184–6.
make them attractive to him in the first place. Socrates’ argumentative strategy here is bold, even breathtaking. He wants to argue not simply that materialist explanations are crude or implausible but, even more strongly, that they are completely unworkable. To do so, he must identify a case in which there is no hope of establishing a proper link between the terms of the theory and the phenomenon to be explained. For that purpose, the numerical and relational examples seem ideal. After all, there would seem to be no plausible way of accounting for phenomena such as ‘twoness’ and ‘largeness’ by pointing to one or more features of a given process or state of affairs. In order to make that point in as vivid a fashion as possible, Socrates notes that in such cases the explanations that would typically be offered cannot avoid the incoherence of admitting opposite explanantia, or of positing an explanans that would suffice to explain opposing explananda, or finally of giving rise to an even more straightforward opposition between explanans and explanandum.

I now want to show how the explanatory demand that is ultimately at issue here may be found at work in Socrates’ encounter with teleology and in his eventual turn to metaphysical analysis. This will put us in a position to answer our ultimate question—the question as to why Socrates might think it appropriate to associate radically different sorts of explanation under the same broad rubric.

Sedley objects to the idea that Socrates is concerned with a notion of explanation (rather than causation) by arguing that, if indeed this were the case, one would expect to see a tendency on Socrates’ part to redescribe various proposed αἰτίαι for maximum explanatory effect. Yet ‘[i]n no case does Socrates replace a rejected cause, such as the bones and sinews rejected as the cause of his sitting, with a redescription of the same item. Rather, each time he substitutes a reference to a quite different item, in this case Socrates’ judgment about what is best’ (‘Causes’, 122). However, the fact that Socrates does not merely redescribe a proposed αἰτία is easily understood in the light of his larger strategy, which consists in rejecting certain broad types of explanation in favour of others. Sedley goes on to offer a further objection: ‘Socrates anyway assumes that a satisfactory cause must be able to survive such redescriptions, at least in the following case: he excludes a head as the cause of something’s largeness on the ground that, a head being something small, this causal account would entail something small’s being the cause of largeness (101 a–b).’ But Socrates uses the phrase “by a head” as a tongue-in-cheek illustration of a degree-of-difference account involving material quantities (compare n. 48). His point is that it is a small quantity (a mere head) that here supplies the supposed explanation of largeness. Nothing about that example need be taken to involve a general assumption about the validity of redescriptions or the referential transparency of causal contexts.
VI

After detailing his frustrations with materialist enquiry, Socrates dwells at length on his attraction to teleology. His initial encounter with Anaxagoras’ work gave him the idea that intelligence is responsible (aitios) for everything, and this in turn suggested to him that everything is arranged ‘for the best’. Unfortunately, a closer acquaintance with Anaxagoras’ book dashed his hopes of learning explanations of that sort. Failing to meet with properly teleological accounts elsewhere or to learn them on his own, Socrates was forced to move on, ending up finally at the theory of Forms.

On many readings of the ‘autobiography’, it is puzzling that teleology should comprise a phase of the discussion—one that is attractive as an alternative to materialism but that is none the less abandoned in favour of another mode of explanation. Some interpreters have accordingly downplayed the importance of the teleological interlude, casting it as an aside bearing no real connection to the main line of argument. Others have gone to the opposite extreme of minimizing the importance of Socrates’ declaration that he is unable to produce workable teleological explanations. Instead, they have proposed, such explanations may in one fashion or another be deemed the ultimate fruit of Socrates’ intellectual endeavours. One advantage of the reading developed thus far is that it allows us to see why the remarks on teleology occur precisely as they do. Given Socrates’ reasons for rejecting materialism, the turn to teleology can readily be understood as a new attempt to satisfy the broad ideal of an adequate explanation.

In discussing the appeal of teleology, Socrates expresses the hope that it will prove the proper route to the sort of explanations he has been seeking all along: ‘If someone wanted to find the explanation of how each thing comes to be or perishes or is, he must find how it’s best for it to be, or to suffer or do anything at all’ (97 c 6–d 1). What Socrates hopes he will thereby discover is an explanans that would be genuinely explanatory, in the sense of offering full rational insight into why things happen the way they do. Such insight would begin with an account of the shape and position of the earth and would

---

61 For brief discussion, see my remarks in sect. 1, along with n. 11. The former tendency is well illustrated by Vlastos, who dismisses the teleological interlude from his otherwise elaborate and careful reading with the sole remark, ‘its message is familiar and, superficially at least, quite clear’ (‘Reasons’, 82).
proceed from there to the relative motions and the orientations of the heavenly bodies. In the end, one would be able to understand what is best for each thing in a way that makes reference to the common good for everything (98 B 1–3). In other words, one would ultimately be able to supply explanations that demonstrate clearly how each *explanandum* cannot be otherwise than as it is.

From the perspective of this new way of thinking, Socrates is able to offer a diagnosis of the confusion into which his materialist predecessors fell. Because they neglected the good when giving their causal explanations, they proved unable to appreciate what it would be to discover that which is truly responsible (*aition*) for a given phenomenon. Instead, they fixed their attention solely upon a mere condition for adequate explanation, ‘that without which what is responsible could not be responsible’. Socrates’ illustration of his point is highly tendentious, involving as it does a case of intentional human action—his own decision to remain in prison. Yet it is important to keep in mind that Socrates’ attraction to teleological explanation rests largely on what he considers to be the antecedent failure of the programme of materialist reduction. And Socrates bolsters his case for teleology towards the end of his discussion by pressing the problem of adequate explanation at the level of celestial phenomena, which of course will present difficult cases for any theory that does not rely on ‘for-the-best’ accounts.

As it turns out, Socrates cannot garner any satisfying teleological theories from Anaxagoras, and so he is forced to embark on his ‘second voyage’, which involves leaving behind such theories altogether. He turns instead to what we have already seen to be a project of metaphysical explanation. Our overarching problem has been to understand the sense in which Socrates might be inclined to view this as yet another route to the explanation—of coming-to-be, passing-away, and being. The preceding reflections afford us a way of appreciating how a turn to metaphysics might well be consistent with the larger concerns that have inspired Socrates all along.

---

62 99 B 2–4: ἄλλο μέν τί ἐστι τὸ αἴτιον τῷ ἄνθρωπῷ, ἄλλο δὲ ἐκεῖνο ὅσον ὁθὸν τὸ αἴτιον οὐκ ἂν ὁποῖον τὸ αἴτιον ἄνθρωπῷ. Much has been made of the fact that Socrates uses a nominalization of the neuter adjective (*τὸ αἴτιον*) rather than the noun *αἰτία*. He does so because he is contrasting what he now considers to be the crucial element of an adequate explanation with the conditions needed in order for that element to be efficacious. Had he tried to express his point using the word *αἰτία*, he would inappropriately have made it seem as if he is distinguishing the proper explanation from something needed to complete it—something without which it could not truly be an explanation.
If indeed his thinking is guided by an ideal of explanatory adequacy, Socrates may well view his metaphysical explanations as satisfying a demand that his earlier theorizings had failed to satisfy. After all, the talk of Forms makes for a connection between *explanans* and *explanandum* that is properly ‘transparent’ and that does not fall victim to the sorts of objection that plague materialist analysis. If one wants to know why something ‘is large’, for example, one will appeal to a Form that is peculiar to cases ‘of largeness’ and that will never be involved in explaining the opposite phenomenon. Likewise, if one wants to explain something’s ‘becoming large’, one will speak of a thing’s *coming to participate* in the particular reality, the *ἰδία οὐσία* (101.3), that serves to account for the application of the term “large” to an object.

As a mode of explanation, this initially seems rather peculiar. It is of course completely non-causal in nature and, compared to the theories Socrates turns away from, it threatens to seem utterly vacuous—*εὐήθης*, as Socrates puts it (100.4). Indeed, it sounds uncomfortably like what is sometimes called a ‘Molière’ explanation. Yet the superficial appearance of emptiness can be mitigated by some further reflection on what might render the metaphysical aitia so attractive in the first place.

As we have seen, Socrates’ attack on his materialist predecessors is effectively an attack on the idea of a reductive explanation. Hence, it makes sense that Socrates would feel himself driven to a completely non-reductive alternative, one that proceeds by acknowledging a theoretical commitment for each of the phenomena to be explained. When I discussed the workings of the metaphysical mode of explanation (Section IV), I described it as proceeding from true statements to the grounds for their truth. Specifically, I suggested that Socrates concentrates on the predicate-expressions of certain statements and wants to know the ground for applying each such expression to a number of objects. His answer involves a Form in which the objects participate. The potential appeal of that pattern of thinking is now easy to see. By means of it, Socrates is able to make good on the idea that there is an entity unique to the application of a given general term; and thereby, he is able

---

63 The phrase derives from Molière’s *Le Malade imaginaire*, where the ability of opium to induce sleep is explained in terms of its *virtus dormitiva*. 
to secure the proposal that his metaphysical *aitia* is in fact wholly non-reductive in character. But what of the fact that the new *aitia* is not in any sense a causal form of explanation? That would be troubling only if one were to presume that Socrates’ ultimate concerns are in fact causal in nature. Such a presumption may have seemed correct on an initial reading of the ‘autobiography’, but it has now emerged that what really engaged Socrates all along was an epistemic ideal that he once thought his causal theorizings could satisfy and that he presently thinks cannot be satisfied except by metaphysics. Socrates’ desire to cast the metaphysical *aitia* as the only alternative to his earlier pursuits helps explain why he would present it in a way that stresses the continuity with what has gone before rather than the unique features of his own approach. Thus at 100c–d he declares that his new *aitia* offers the only viable explanation of why something ‘is beautiful’, or even ‘becomes beautiful’. However, Socrates well understands that his new *aitia* does not in fact offer any account of how changes occur, which was the central concern of the theories he has left behind. So at 101c he carefully qualifies the sense in which his new mode of explanation supplies an account of becoming. In effect, it explains only the ‘be’ part of a ‘come to be’ statement. The process of change is not actually at issue. Interestingly, Socrates is not at all apologetic about that limitation. On the contrary, he firmly enjoins Cebes to embrace his own mode of explanation by telling him with evident sarcasm to bid farewell to all materialist theories, which are subtleties suitable only for ‘those more sophisticated than yourself’ [τοῖς σεαυτοῦ σοφωτέροις].

Still, one might pose a problem. If indeed Socrates is clear about what he is arguing, why should he be at pains to stress any sort of continuity with his earlier concerns? Even allowing that he is guided by an overarching ideal of adequate explanation, he is presently proposing to satisfy that ideal in quite a different way from be-

---

64 I should note here that there is a kind of reductive analysis that might be said to attend the project of supplying the truth-grounds for statements. It involves distinguishing between ‘defined’ and ‘undefined’ items of one’s vocabulary and, thereby, identifying a basic set of terms relative to which metaphysical commitments are to be determined. However, I take it that at *Phaedo* 100 b ff. Socrates is discussing expressions that he would consider indefinable, or irreducible. Thus, the latter notion of reductive analysis is irrelevant to his present concerns. Even when Socrates elsewhere wonders about the scope of his metaphysical commitments (as at *Parm.* 130c–d), he does not have in mind any set of theoretical principles that would readily determine his thinking one way or another.
fore. The guiding questions have effectively shifted along with his answers to them. Surely it would have been better—not to mention more modest—had he made the shift explicit by renouncing any further concern with the explanation of ‘coming-to-be and passing-away’.

In thinking about that problem, it helps to remind ourselves of the larger context in which the ‘autobiography’ is situated. The metaphysical aitia is ultimately attractive to Socrates in so far as it provides a scaffolding for the final proof of the soul’s immortality, a proof meant to address Cebes’ objection that the soul is merely a long-lasting thing which will eventually perish. As Socrates considers the task of proving the soul’s immortality, he may well think that he needs to proceed by first banishing materialism from the intellectual landscape. After all, materialism stands in stark opposition to the theory of Forms and thus threatens to destroy the appeal of any proof based on that theory. The means that Socrates adopts for attacking a materialist world-view are extraordinarily creative. The very theory that will provide the basics of his proof of immortality also turns out to be the one that definitively vanquishes materialism—at least when both are compared according to the same explanatory requirement and the latter is found wanting. One can thus see why Socrates would feel driven to compare materialist and metaphysical explanations with one another and why he would look for a general rubric under which to fit all of his concerns. The rubric that he settles on—the explanation of ‘coming-to-be and passing-away and being’—is undoubtedly somewhat awkward, but Socrates needs something elastic enough to encompass all of his concerns while being vague enough to accommodate different explanatory methods. And in so far as metaphysical explanation will be what eventually helps show the immortality of the soul, Socrates can perhaps content himself with the thought that metaphysics has a limited relevance to reflections on coming-to-be and passing-away. Far from being a mere Molière explanation, the theory of Forms turns out to be the crucial device for inaugurating what Socrates deems the correct vision of the world, one that accords a central place to the immortal soul.

In this connection, I might note a further reason why Socrates may feel driven to recommend metaphysical analysis by casting it as an alternative to other modes of explanation. It is that he cannot simply begin with the theory of Forms and proceed from there to
the immortality of the soul. The reason is that the metaphysical theory is woefully incomplete, as Socrates himself acknowledges at 100d when he confesses to being at a loss regarding the proper explanation of the participation relation. Had he tried to defend the theory on its own terms, it would have seemed no more compelling than the competing theories of his predecessors. Indeed, we know from the opening of the Parmenides how vexing an adequate account of participation must have been for Socrates. There, the older Parmenides has to remind the young Socrates not to let his difficulties concerning participation undermine his confidence in the very existence of Forms (135b–c). By the dramatic date of the Phaedo, it would seem, Socrates is no closer to having a viable account of participation. But he has effectively learnt the lesson of Parmenides, and he now seeks to highlight the attractions of the theory of Forms by showing why it is preferable to several unworkable alternatives.

As I suggested earlier, Socrates’ drive to cast his aitia as a compelling response to the failings in the theories of his predecessors may help to explain why he does not linger over its peculiarities at 100b–d. Yet once he has finished comparing it with the theories he has left behind, the discussion can more easily make room for an acknowledgement of what is unique about the new aitia—namely, that it turns on the way we speak about the world. That acknowledgement is worked into the conversation through an intrusion of the dialogue’s narrative frame (102b) and is then unobtrusively

\[^{65}\text{The desire to avoid a discussion of the participation relation may be part of the motivation for the methodological comments (the famed method of hypothesis) adumbrated at 100a 3–7 and amplified somewhat obscurely at 101d 1–102a 1. A number of scholars have taken the methodological comments to be part and parcel of the second voyage, but I agree with those who deem it an additional element of Socrates’ discussion. After all, the methodological remarks are not limited to issues of explanation: they are understood to be perfectly general in scope, applicable καὶ περὶ αἰτίας καὶ περὶ τῶν ἄλλων ἁπάντων (100a 6). However, the ‘second voyage’ is concerned specifically with matters of explanation, as is shown by the linguistic parallels between 99d 1–2 and 100b 3–4. At 99d 1–2 Socrates proposes to exhibit the way in which he has been pursuing his second voyage (τὸν δεύτερον πλοῦν . . . ἐπίδειξιν ποιήσωμαι;). Then, at 100b 3–4, he begins to take up that proposal by offering to exhibit the mode of explanation which he now pursues (ἐρχόμαι . . . ἐπικείμενον τῆς αἰτίας τὸ εἶδος ἐπηρεαστέον). \(\text{Compare also 100b 8.}\)\] Unfortunately, there is little consensus among interpreters as to how the method of hypothesis should be understood or where precisely it is being applied. I cannot treat the relevant issues here, but for some discussion of recent interpretations, see Kanayama, ‘Methodology’, 51 ff.\]
incorporated into Socrates’ illustrations and applications of the metaphysical theory (102 df.).

In the end, Socrates’ whole strategy is exceedingly clever. Perhaps it is simply too clever, since it forces him to build his argument on the awkward objections to materialism that he offers at 96 d–97 b. But what is most worth emphasizing is that, far from being the confused jumble that a number of interpreters have taken it to be, the passage carefully selects and arranges the materials of Presocratic thought so as to recommend a radically new philosophical enterprise. Its creativity and argumentative boldness are impressive indeed, even for a philosopher who is no stranger to bold proposals. The sheer variety of the themes that are sounded and the way in which they are carefully arranged make our text one of the richest announcements of a programme in the entire history of ancient thought. When considered in that light, it is truly a fitting tribute to Socrates as he approaches the hour of his death.

California State University, Long Beach

BIBLIOGRAPHY


