No cover image available

Desire, Practical Reason, and the Good

Sergio Tenenbaum (ed.)

https://doi.org/10.1093/acprof:oso/9780195382440.001.0001

Published: 2010 **Online ISBN:** 9780199870158 **Print ISBN:** 9780195382440

Search in this book

CHAPTER

3 Plato on the Desire for the Good

Rachel Barney

https://doi.org/10.1093/acprof:oso/9780195382440.003.0003 Pages 34–64

Published: July 2010

Abstract

The chapter discusses Plato's oft-repeated assertion that all human desire is for the good. This claim is perplexing in a number of respects. For one thing, it seems to come in two forms, both of them apparently empirically false (and mutually incompatible): (1) all desire is for what the desiring agent views as good ('the apparent good') and (2) desire can only be for what really is good. The chapter argues that the thesis amounts to the claim that desire is realist or objectivist in orientation: In desiring, we attempt to respond appropriately to the detection of value in the world. Thus, without being vacuous (for it rules out many alternative conceptions of desire), Plato's thesis sets no empirically refutable limitations on what objects may be desired. His point is rather that whatever is desired is desired on the basis of an evaluative belief. The tension between the two forms of the thesis is resolved by taking (2) as a clarification of (1), intended to make its realist import explicit: We desire 'the apparent good' not under that description, but simply in the sense that we pursue the real good as best we can detect it. The thesis is thus comparable to some epistemological arguments in the Sophist and Cratylus: Plato holds that a certain commitment to realism is built into our everyday cognitive practices, desiring included.

Keywords: Plato, desire, good, realism

Subject: Philosophy of Action, Philosophy, Moral Philosophy

Collection: Oxford Scholarship Online

In a number of dialogues, Plato affirms in various ways that human desire is for the good. In the *Protagoras*, Socrates proposes that "no one who knows or believes there is something else better than what he is doing, something possible, will go on doing what he had been doing when he could be doing what is better" (358b7–c1).[1] Likewise, "No one goes willingly toward the bad or what he believes to be bad; neither is it in human nature, so it seems, to want to go toward what one believes to be bad instead of to the good" (358c6–d2). In the *Meno*, Socrates argues that no one really desires or wants [*epithumein*, *boulesthai*] what is bad (77a–78c). In the *Gorgias*, he claims that when we act, what we want [*boulesthai*] is some beneficial

Downloaded from https://academic.oup.com/book/4008/chapter/145635901 by University of Toronto user on 11 March 2024

outcome; if it does not ensue, then the agent does what seems good to him, or as he sees fit [*dokei autô*], but not what he *wants* [*bouletai*] (466a–468e). Diotima and Socrates agree in the *Symposium* that the desire [*epithumia*] for happiness or good things is the supreme love or longing [*erôs*] in everyone (205d); in fact, "What everyone loves [*erôsin*] is really nothing other than the good" (205e7–206a1).

This *Desire thesis*, as I will call it, forms a key part of the so-called 'Socratic intellectualism' of the early dialogues, along with several other paradoxical claims: that the virtues consist in a kind of knowledge, and thus form a unity; that *akrasia* or weakness of will is impossible; and that all wrongdoing is involuntary [*akôn*]. The Desire thesis seems to be foundational for the others, for it (arguably) precludes *akrasia*, and, in conjunction with the fact that what is really good for us is virtue, entails both that moral knowledge suffices for virtuous action and that wrongdoing can only be caused by ignorance. It is often claimed that in Book IV of the *Republic*, Plato rejects the Desire thesis, and with it this whole package of Socratic moral theory.[2] Recognizing—rather late in life, one might think—that some desires are not directed toward the good at all, Plato introduces the lower parts of the tripartite soul to house these ''good-independent'' desires, thus repudiating the Desire thesis and its intellectualist implications. Yet in Book VI of the *Republic*, Socrates goes on to say that the good is "what all soul [*hapasa psuchê*] pursues [*diôkei*] and does everything for the sake
p. 35 of" (505e1–2). And in the late *Philebus*, Socrates says of the good that ↳ "everything that has any notion of it hunts for it and, desiring [*boulomenon*], reaches out [*ephietai*] to get hold of it and secure it for its very own, caring nothing for anything else except for what is connected with the acquisition of some good" (20d7–10). The corollary that wrongdoing is involuntary is also repeated emphatically in the very late *Timaeus* (86b–87b).

So we have reason to suspect that the Desire thesis is a sustained and foundational principle of *Platonic* moral psychology. But whether the various statements I have quoted really do boil down to a single, consistently held thesis is far from certain. And indeed there seem to be two distinct versions of the thesis in play. One, which I'll call the *Appearance thesis*, is that all desire is for the apparent good—that is, for an object the desiring agent *takes* to be good. It is this claim, that desire is *sub specie boni*, which recurs in later philosophers such as Aquinas and Spinoza, and remains a live and controversial option today. The other, the *Reality thesis*, looks odder and more distinctively Platonic: it is that human desire can only be for *what, in fact, is good*. It is the latter that in the *Gorgias* leads to the shocking corollary that wrongdoing is involuntary [*akôn*]—specifically, that Archelaus, the brutal yet glamorous tyrant cited by Polus as a paradigm of flourishing injustice, does not do what he wants.[3]

It's some measure of the sheer confusingness of the issues here that, depending on how they are presented, the Reality thesis can seem either to follow trivially from the Appearance thesis or to contradict it. On the one hand, it seems natural to say that if some desire is perspicuously described as being for *x*, then it is for what really, genuinely is *x*. After all, if I say that I want to eat an apple, you probably do assume, without further inquiry, that I want to eat a real apple—as J. L. Austin noted, in the absence of some particular puzzle or contrast, the term 'real' doesn't seem to add anything in everyday contexts.[4] And when Plato (or for that matter Aristotle) says that wealth is desired as an apparent good, what he means is that the desiring agent takes it to be a *real* good—'apparent goods' aren't some natural kind that we might prefer to real ones, but are just the class of things thought to be genuinely so. So the Reality thesis can be read as essentially a gloss on the Appearance one—and indeed I will eventually argue that this is the right way to take it. On the other hand, consider again the desires of Archelaus the tyrant. Suppose that Archelaus avows a desire to be the most powerful man in Macedonia. The Appearance thesis diagnoses that desire as a desire for an apparent good; power without wisdom is not really a good, Socrates argues, but Archelaus desires it because he thinks it is. But the Reality thesis simply rejects Archelaus' self-description: given that power without wisdom is not a genuine good, it can't be the object of his desire *at all*. So the two theses give directly conflicting results, and the Reality thesis seems to preclude precisely that fallibility in selecting objects of desire that the Appearance thesis attributes to us.

Downloaded from https://academic.oup.com/book/4008/chapter/145635901 by University of Toronto user on 11 March 2024

In rough outline, it is clear what Socrates wants to say about Archelaus. Given that it, in fact, frustrates his
p. 36 desire to obtain the good, Archelaus' wrongdoing must be understood as *misguided* action, based on a faulty
↳ evaluation of the object that motivates it. As Santas explains in a classic discussion, the agent who pursues what is not really good is like one who, desiring pepper (the 'intended object' of his action), reaches out for what is actually the salt shaker (the 'actual object').[5] In such a case, if we grasp what is going on, we may feel licensed to warn the agent: 'You don't really want that!' Given Plato's relentless objectivism about the good, such trivial misidentifications represent one end of a whole spectrum of 'miswanting,' with the ostensibly satisfied tyrant at the other. There will of course be important differences between the tyrant and the confused condiment-seeker, for the tyrant's false beliefs are not merely ad hoc perceptual errors. They may comprise a whole network of mistakes about value, involving deep conceptual confusions and failures of self-knowledge; and attempts to correct them are likely to meet with enormous psychological resistance. But all these complex psychological ramifications have their origin in what is nonetheless the same *kind* of cognitive mistake. The question that remains is whether the 'actual object' fixed on by that mistake is properly counted as an object of desire at all. If it is, then the Reality thesis and the corollary that wrongdoing is involuntary seem to be given up; if not, then desire no longer seems to occupy its standard functional role in the motivation of Archelaus' behavior.

This chapter attempts to work out what each of these two versions of the Desire thesis means, as well as the puzzling relation between them. Ideally, such a reading of the Desire thesis should meet a number of desiderata. First, it should take both theses seriously and literally as presented—that is, as psychological laws, not as eristic gambits, disguised normative claims, or exaggerated generalizations. Second, it should respect the fact that both theses are presented as global claims about conation in general, not just desire in some special restricted sense—in contrast, for instance, to Aristotle's account of *boulêsis*, rational desire, in *Nicomachean Ethics* III.4. Though some interpreters detect a distinction between *epithumein* and *boulesthai* here (as I will discuss in section I), Plato himself seems to vary his terminology in order to emphasize its indifference: as we saw in the first paragraph above, he speaks not only of wanting [*boulesthai*] and desiring [*epithumein*] the good but also of loving [*erân*], pursuing [*diôkein, Rep.* 505e1], choosing [*haireisthai, Prot.* 358d2–3], hunting for, aiming at [*thêreuein, ephiesthai, Phil.* 20d8], and "being willing to go toward" [*ethelein ienai epi, Prot.* 358c6–d2] it. This gratuitous variation is a strong signal that we should understand the Desire thesis as one about motivation across the board: thus, following Plato, I will, for the most part, use 'want,' 'desire,' and so forth interchangeably.

An ideal interpretation of the Desire thesis would also show how Plato might reasonably take the Appearance version to be an intuitively plausible principle, so that it may in some contexts be proposed without argument (*Rep.* 505e1–2, *Prot.* 358d), while in others Plato supports it with what looks more like dialectical clarification than demonstration (*Gorg.* 467c–68c; *Symp.* 205e). It would show how the Reality
p. 37 thesis warrants ↳ the scandalous inference that Socrates draws from it in the *Gorgias*: that actions that obtain the bad—e.g., the unjust decrees of the tyrant Archelaus—are involuntary [*akôn*] (509e). And it would explain how, in both the *Meno* and the *Gorgias* (as we will see), the Appearance and Reality theses are not only treated as compatible: the Reality thesis is introduced by being *added on* to the Appearance one, as if it were a trivial variant or a snappy corollary.[6] Thus a fully satisfactory reading would depict the two theses forming a coherent unity, with the Reality thesis being easily derived from the Appearance one. Finally, an optimal reading would explain how Plato could continue to hold some version of the Desire thesis even after writing *Republic* IV.[7]

So far as I can see, not all of these desiderata can be met; and most recent interpretations have given up on at least one of them. Thus both Rachana Kamtekar and Heda Segvic have recently argued for taking Plato's talk of 'wanting' [*boulesthai*] as sharply distinct from ordinary desire [*epithumein*]. On Kamtekar's reading, which relies heavily on an analogy with our real latent beliefs as revealed by the Socratic elenchus, the thesis holds that while all people desire [*epithumein*] good things, vicious people desire bad ones as well; but

Downloaded from https://academic.oup.com/book/4008/chapter/145635901 by University of Toronto user on 11 March 2024

wanting [*boulesthai*] is oriented to the good alone.[8] This 'wanting' is a real if mysterious psychological phenomenon, a kind of latent teleological orientation to the good revealed less in our avowed desires than other conative behavior (for instance, our being *satisfied* or not by the attainment of some perceived good). Heda Segvic has also developed an account of the Reality thesis as expressing a special Socratic conception of 'wanting' [*boulesthai*]. Like perception, wanting is a kind of successful interaction with the world:[9] to want something in this sense involves *knowing* that its object is good, and that one cannot want the bad is simply part of the concept. Both of these readings are carefully argued and illuminating on many points; but neither is easily reconciled with Plato's own statements of the Desire thesis, which certainly sound like naturalistic, lawlike claims about the causality of conation in general as normally understood. The leading representative of a more literalist reading of the Desire thesis is Terry Penner, and this chapter is greatly indebted to his work.[10] However, Penner's reading (or evolving family of readings) takes Plato's views on desire to be enmeshed with a complex set of commitments regarding reference and psychological states; it is hard to see the theses so read as intuitively graspable and widely accepted ones, which again is what Plato's own presentation of them seems to call for. Penner's reading also involves doing away with the Appearance thesis altogether—a heavy cost since it is, on the face of it, better attested than the other version.

My own solution, briefly, will be to take the Appearance thesis *au pied de la lettre*, while opting for a deflationary reading of the Reality thesis. The latter, I will suggest, is not intended as an independent thesis at all but rather as an *interpretation and clarification* of the Appearance thesis: It is *not* to be taken in the
p. 38 strong and literal sense that would render Plato's ↳ argument self-defeating, by in the end denying that we are fallible in our desiring. Rather, taken together, the two versions of the Desire thesis express a position I will term *cognitivism* about desire. As the Appearance thesis says, I always desire what seems good to me. But, as the Reality thesis clarifies, that does *not* mean that I desire objects *under the description* 'what seems good to me,' taking my subjective responses to be constitutive of value. Rather, in desiring I do my best to track what is antecedently valuable, insofar as I can detect it. Properly understood, the Desire thesis is really a claim about the priority of cognition to motivation, and depicts the latter as world-guided in just the same way as the former.

In section I, I begin with detailed analysis of the *locus classicus* of the Appearance thesis, the rather tricky argument at *Meno* 77–78. Section II tries to explain what 'good' means in the Appearance thesis, and introduces this 'cognitivist' reading. Section III turns to the *Gorgias* and the Reality thesis, and considers the thorny question of how the Appearance and Reality theses are related. Section IV then places the Desire thesis as a whole in the broader context of Plato's realism: its real force, I will argue, is to claim that there is a certain commitment to objectivity built into our ordinary ways of believing and desiring.

I

The starting point of Socrates' argument in the *Meno* is Meno's proposal to define virtue, *aretê*, as "to desire [*epithumein*] fine things [*ta kala*] and have the power to acquire them" (77b), a suggestion that Socrates reformulates in terms of desiring good things [*agatha*]. (The sort of good things a virtuous person wants to secure are specified in Meno's later remarks: health and wealth, gold and silver, honors and offices in the city [78c].) Socrates then rejects the first part of the definition as idle. Desiring good things is not a perspicuous criterion for virtue, for no one ever does otherwise.[11]

Socrates begins by distinguishing two putative groups who might be described as desiring what is bad: those who do so believing the bad things to be good (call these group A), and those who do so knowing them to be bad (group B). Meno affirms the existence of both (77c2–5). Socrates then divides the latter group into those who desire bad things believing that those things *benefit* their possessors (group B1), and those who

Downloaded from https://academic.oup.com/book/4008/chapter/145635901 by University of Toronto user on 11 March 2024

recognize that they are harmful (group B2) (77d1–4). He gets Meno to agree that those who believe that the bad things will benefit them do not really know that they are bad: in effect, group B1 really must be subsumed under group A (77d4–7). In a somewhat mysterious move to which I will return, Socrates adds that those who desire what they believe to be good *desire good things*; that is, there is really no group A either (77d7–e4). Finally, Socrates turns to deal with group B2. Those who believe that the bad things they desire
p. 39 harm their possessors must know ↳ that they will be harmed by possessing them, that they will be miserable to the extent they are harmed, and that they will thereby be made unhappy (77e5–78a4). But nobody wants [*boulesthai*] to be unhappy (78a4–5). So nobody wants to possess bad things: there is no group B2. Given that B1 and A have already been shown to be empty, there is therefore no one at all who wants what is bad (78a5–b2).

My switch from 'desire' to 'want' above reflects a shift in Socrates' usage from *epithumein* to *boulesthai*, which raises a basic question as to whether the same attitude is under discussion throughout this passage. The two terms certainly have different connotations, and, in some contexts, different senses.[12] Thus a number of interpreters, including Kamtekar and Weiss, take Socrates to here distinguish two distinct attitudes with different objects, with *boulesthai* restricted to the genuinely good.[13] But any such reading comes at a high price, making the argument as a whole a matter of deliberate fallacy. For when Meno finally grants that "no one wants [*boulesthai*] bad things" (78b1–2), it is clearly an admission that his opening thesis has been refuted. Socrates even presses the point in a way that rephrases Meno's earlier claim: "Were you not saying just now that virtue is to want [*boulesthai*] good things and be able to get them?" (77b3–4). Their shared conclusion, that the first part of Meno's definition of virtue should be dropped, follows from these steps. But if the two terms are not legitimately interchangeable, then none of this genuinely follows.

So it is better to take the difference between *boulesthai* and *epithumein* here as one merely of connotation.[14] The final steps of the argument are, after all, structured just as we would expect them to be if *epithumein* had been used throughout (which is not, of course, to deny that Socrates aims to exploit the differing connotations of the terms in order to maximize the plausibility of his argument to Meno);[15] and as I noted earlier, such terminological vacillation is, in fact, characteristic of Plato's presentation of the Desire thesis. In a passage of the *Symposium* that echoes this one in many respects (204d–6b),[16] there is again a switch from *epithumein* to *boulesthai* (205a2ff.) when the object is happiness in general; but Diotima also mixes and matches the terms *en passant*, speaking of a *boulêsis* and *erôs* to have good things forever (205a5) and of an *epithumia* for good things and to be happy (205d1–3). And Plato's affirmations of the Desire thesis often bypass attitude terms altogether in favor of more behavioral talk about what we 'go toward' or 'pursue,' casting the good as object of *all* the motivations that cause intentional action (*Protagoras* 358c6–d2; *Republic* 505e1; *Gorgias* 468b1; *Philebus* 20d8).[17]

So read, the *Meno* argument proceeds by a reasonably straightforward exhaustion of alternatives.[18] The key eliminative moves come as the groups B1, A, and B2 are successively redescribed and excluded from the ultimately empty set of those who desire the bad. The most transparent of these moves is the first, at 77d4–7. No one can be correctly described as 'desiring the bad believing it will benefit him,' presumably because
p. 40 to ↳ think of something as beneficial *just is* to think of it as good. As this reveals, the argument is structured around a framework of conceptual connections embedded in everyday Greek usage, which are presented as uncontroversial here and in other similar contexts.[19] Most of Socrates' moves here function primarily as reaffirmations and clarifications of this framework, which Meno is happy enough to accept. As it eventually emerges, good things [*ta agatha*] are as such beneficial [*ôphelima*] while bad ones [*ta kaka*] are harmful [*blabera*]; and the beneficial and harmful as such contribute to our happiness [*eudaimonia*] and unhappiness [*kakodaimonia*], respectively. Desire, meanwhile, is what I'll term an *appropriative* attitude, an impulse to *obtain* some object (or a standing disposition to have such an impulse): when we desire, as Socrates says, what we want is for the object of desire to become ours (77c7–8; see also *Symp*. 204e3–4, 206a6–8; *Philebus* 20d9). Of course, none of this tells us anything about what objects *are* good and bad to obtain. Rather, all

Downloaded from https://academic.oup.com/book/4008/chapter/145635901 by University of Toronto user on 11 March 2024

these connections are conceptual or formal, and can be shared by interlocutors with radically different substantive conceptions of what the good and happiness consist in. (*A fortiori*, 'good' here, as in Plato's other expressions of the Desire thesis, clearly does not mean *morally* good—though Socrates does also hold that, as a matter of fact, the life of virtue is what our good consists in. Hence 'evil' is a very misleading translation for *kakon*, despite the awkwardness of 'bad' in plural uses.)

This framework is also in play in Socrates' second move of exclusion, when he rejects the possibility of desiring the bad while recognizing it as harmful. This seems to presuppose that, if I do not want the effect, I cannot want the cause; and it might be objected that this is (if anything) a normative rather than a descriptive principle. For surely desire does not always respect causal relations in a rationally coherent way: I can recognize that something will cause me unhappiness and, irrationally, want it nonetheless—not want the unhappiness, perhaps, but want the very thing that causes it.

This objection could be blocked if we could assume that all our particular desires trickle down, so to speak, from One Big Desire, our desire for happiness—that is, that whatever I want, I want strictly as an instrumental *means* to happiness.[20] The difficulty is that, despite the conceptual relations between goods and happiness I have just sketched above, this does not appear to be what Plato has in mind. Rather, the thesis seems to be concerned with immediate, piecemeal responses to things found attractive in their own right—bright shiny objects like health and wealth, honor and high office (see *Meno* 78c–d, 87e; *Euthyd.* 278eff.). Perhaps we could see these as (on Meno's view) constituent or ingredient means to happiness, or even as a rather crass 'objective list' account of what happiness is. But since it is generally accepted that such goods can fail to benefit (*Euthyd.* 280b–81e), their relation to happiness is presumably mediated by the experiential benefits they are intended to secure, such as pleasure, satisfaction, freedom, the beautiful
p. 41 [*kalon*], etc. I will refer to these general ↳ kinds of value as 'mid-level' goods.[21] And in any case, neither Meno nor Socrates seems to suppose that any thought about my happiness—that is, my overall long-term welfare or flourishing—must play a causal role in my desire either for particular goods *or* the mid-level benefits they secure. In most of the relevant texts, Socrates does not even seem to be discussing comparative or 'all things considered' desires.[22] For the purposes of the Desire thesis, desire is evidently conceived as a simple, direct, two-place relation between an agent and a perceived good such as power or fame. And this gives us an important clue for interpreting the Appearance thesis. Plato is *not* claiming that every desire must be mediated by some thought about our own long-term welfare, in a way that would exclude self-destructive or wanton desires; his claim is only that our appropriative impulses proceed from cognition of their objects *as* valuable in some way.

So these two exclusionary moves, eliminating groups B1 and B2, really work the same way. Their point is to nail down Meno's assent, in the face of his initial wavering, to the identity of the good, the beneficial, and the happiness-inducing, as interchangeable descriptions of the object of desire. Meno's initial inclination is to see the three as capable of coming apart to some extent, presumably because reflective thoughts about what is beneficial and about happiness might generate desires in conflict with others not so mediated. But under just a little conceptual pressure from Socrates, Meno is prepared to join him in treating the three as equivalent.

The remaining turning point in the argument is its most mysterious and controversial step: namely the exclusion of group A, immediately following its subsumption of B1, when Socrates disqualifies from 'desiring the bad' those who desire objects they *think* are good:[23]

> It is clear then that those who do not know things [that is, the things they desire] to be bad do not desire what is bad, but they desire those things that they believe to be good but that are in fact bad. So that those who have no knowledge of these things and believe them to be good clearly desire good things.

Downloaded from https://academic.oup.com/book/4008/chapter/145635901 by University of Toronto user on 11 March 2024

Despite its breezily inferential air, that crucial second sentence doesn't really follow in any obvious way. Plato seems to be engaging in legislation about the ascription of desires, to the effect that they are to be ascribed using a description of the object that the desiring agent would avow (viz., 'good'). In terms of the distinction introduced by Santas,[24] where there is a gap between the 'intended object' of a desire (the description the agent would avow) and the 'actual object' (what we observe him going for), we should identify desires in terms of their intended objects only. But no argument has been given, or even hinted at, for us to accept that stipulation. As McTighe and Vlastos have pointed out, we might consider the case of Oedipus.[25] Oedipus wants to marry Queen Jocasta (under that description); he doesn't want to marry his mother (under that description); unfortunately for him, he cannot marry a description and they are, in fact, the same person. If we accept the Socratic stipulation, it is simply wrong to say that Oedipus desires to
p. 42 marry his mother. But there are contexts in ↳ which it might seem correct and informative to say exactly that, and Socrates has given us no reason not to.

This question is complicated by another, related puzzle raised by the same passage. Whether the Reality thesis, as well as the Appearance thesis, is in play anywhere in the present argument is a matter of interpretive controversy.[26] It *should* be: Meno, like Callicles and Thrasymachus, is an advocate for the political life, with wealth and power as his canonical 'good things' (78c), and it seems most unlikely that he would endorse as happy anyone who manages to get whatever odds and ends *he* happens to think good. More explicitly, at 77c5 Meno affirms that he includes among those who want bad things people who do so thinking that those things are good. So for Socrates really to exclude all of Meno's proposed ways of desiring the bad, as he must do to warrant dropping the first part of his definition of virtue, the Reality thesis *must* be in view. Now the most plausible point at which to locate an assertion of the Reality thesis in the *Meno* argument is in this same mysterious elimination of group (A). For we may read it as affirming that everyone who wants what he *thinks* good is, *ipso facto*, properly described as wanting what *is* good—and not what is bad, even if the 'actual object' he pursues is bad. In that case, Socrates' claim that group A should count as desiring the good (full stop) is not best read as a stipulation about the priority of descriptions the desiring agent would avow (not, at any rate, unless we also stipulate that nothing is properly desired under any description *other* than 'good'). For the Reality thesis claims that a desire for the tyrant's license to do injustice, say, should *not* be ascribed to the person who avows it. Like everyone else, he *really* desires only what really is good. Admittedly, much more would need to be said to spell out a viable view here. In particular, we would need a characterization of the difference between avowed 'desire' and *really* desiring that renders the distinction intelligible, a problem I will turn to in the next section. Be that as it may, Socrates' inference here at 77e2–3 seems to amount to a kind of slippery-slope transition from the Appearance thesis to the Reality one. To desire what seems good to us is to desire it *qua* good, and this is *really* to desire what really *is* good.

To sum up the story so far: Plato in the *Meno* intends to assert both the Appearance thesis and the Reality thesis, and both are intended as theses about human conation across the board. He presents the former as more or less intuitively obvious, by showing that just a little conceptual clarification induces Meno's assent to it; and he seems to think, for reasons that are so far mysterious, that the Reality thesis follows unproblematically from the Appearance one.

II

To see what Plato is driving at here we need to begin by taking a closer look at his understanding both of
p. 43 desire and of goodness. As I noted ↳ earlier, desire is here conceived as an appropriative impulse, a cause of actions intended to obtain the desired object for the agent. We might be tempted to construe the Appearance thesis as *identifying* this appropriative attitude with a cognitive state.[27] But the *Meno* passage consistently presents thinking good and desiring as two distinct psychological operations, one of which is causally prior

Downloaded from https://academic.oup.com/book/4008/chapter/145635901 by University of Toronto user on 11 March 2024

to the other. And this seems right given Plato's other commitments. For one thing, it is an important principle for Plato, made explicit in relevant passages of the *Symposium* (204a) and *Philebus* (34d–35d), that we can only desire what we *lack*.[28] But since the belief that something is good can be held just as easily about an object one already possesses, the desire for a thing cannot be identical with that belief.

So the Appearance thesis is not a direct ancestor (at least not a legitimate one) of those views that explain desire as *itself* an evaluative belief or a perception-like experience of value.[29] Nor is it a claim about the conditions required for us to interpret something as a desire in the first place; nor a set of necessary and sufficient conditions for desiring.[30] In fact, it does not seem to me an attempt to analyze desire itself at all. Rather, it is a simple *causal* claim about the mechanics of human motivation, to the effect that we cannot desire something without first finding it good. This is a claim about the limits of human nature (*Protagoras* 358c6–d2, cited in the opening paragraph above), for which the ultimate explanation presumably lies in the providential teleology of the *Timaeus*.

As to what desire itself consists in, Plato has little to say beyond characterizing it as an appropriative attitude. But in the *Republic*, he explains that a soul experiencing a desire 'takes aim at it' or 'draws it to himself'; or again 'nods assent to it as if in answer to a question' (437b1–c6). In the opposite states, the soul 'pushes and drives things away' (437c6–d1). So desire seems to be quite literally a psychological inclination, an inner impulse or lunging toward an object at some remove—a kind of internal rehearsal for the motion of the body in voluntary action. The Appearance thesis thus amounts to the claim that an inner, action-causing inclination of this kind is causally dependent on a certain kind of cognitive state, specifically the belief that its object (that is, the object to be obtained by the corresponding external movement) is good.

We can now turn to consider more closely what that belief amounts to. This can best be clarified by considering two familiar lines of objection to the Appearance thesis. First, we might think that some desires are too *primitive* to count as dependent on a belief about the goodness of their objects. Second, it seems that a desiring agent might have a conception of the good but, perhaps quite deliberately, not desire what fits it —that is, a desire might be deliberately bad-seeking or *perverse*.

Whether any version of the Desire thesis can account for the desires of perverse and primitive agents remains a matter of deep philosophical controversy. As G. E. M. Anscombe noted, even Milton's Satan, with his resounding cry "Evil be thou my Good,"[31] may reasonably be asked: ↳ 'What's the good of its being bad?' And some perfectly intelligible, even familiar answers can be supplied: "condemnation of good as impotent, slavish, and inglorious … the good of making evil my good is my intact liberty in the unsubmissiveness of my will."[32] This points toward a strategy well suited to the Platonic version of the Desire thesis, namely the explanation of perverse desires as only superficially deviant instances of desire for intelligible and even familiar 'mid-level' values. Satan desires what is *liberating*; other ostensibly bad-oriented agents might be found to desire the objects they do under the generic description 'pleasant,' say, or 'vindicating' or 'surprising.' And this prospect can help us to understand the role and content of the 'good' in the Desire thesis. Goodness here operates as a formal concept—the highest genus uniting such mid-level values, the positive evaluative valence they all have in common.

p. 44

So read, the Appearance thesis presents human desire as a response to two in principle distinct cognitive operations. One is the taking of some object to have a certain property—or, better, a set of properties, nested at different levels of generality and culminating in one or more 'mid-level' values: this is salty-and-thereby-delicious-and-thereby-pleasant, for instance. The other is the taking of this hierarchy of properties as *good*. There may or may not be an explicit or propositional judgment involved in these acts of cognition; but they are both acts of *classification*, and thus imply the possibility of universal judgments. *This* has *this* feature, and *this* feature is good: implicitly, *any* relevantly similar object would count as having this feature, and *anything* with this feature would count for us as to that extent good. And Plato's point seems to

Downloaded from https://academic.oup.com/book/4008/chapter/145635901 by University of Toronto user on 11 March 2024

be that in principle, this provides an explanatory schema for the explanation of *all* human desire, which cannot take place without these cognitive acts.

If this is right, we should not see Plato as insisting that the belief that precedes desire must be *explicitly* about the good (let alone about the morally good, the good all things considered, or happiness). It need only be an evaluative belief of positive valence, picking out *as* valuable some property possessed by the object in view. Something is good by virtue of its participation in some mid-level value: by being pleasant, honorable, virtuous, liberating, sacred, and so forth. And to say that *these* values are good is not (or not only, and not exactly) to say that they contribute to our happiness; rather, it is to say that they carry a positive evaluative charge—a location on a table of values—that triggers pursuit as the appropriate response. The residual puzzle, of course, is whether this is a sufficiently robust conception of 'good' for the Appearance thesis to have any real force. What, if anything, constrains our selection of mid-level values? And what exactly does it mean for us to classify them *as good*, if this is to be something distinct from and causally prior to desiring them—*and* also distinct from thinking of them as part of our happiness (as I insisted in rejecting the One Big Desire hypothesis)? Absent an account of what the positive evaluative valence here amounts to, 'good' threatens to collapse into 'desirable,' and from there into the merely desired.

p. 45 We can reach the same point from the other direction, through consideration of primitive desires. That our most basic physical appetites are good-independent seems to be a point raised and treated as an objection to the Desire thesis by Plato himself in *Republic* IV. Socrates here goes out of his way to argue that the appetite of thirst is not for good drink but for drink simpliciter: "Thirst itself isn't for much or little, good or bad, or, in a word, for drink of a particular sort … thirst itself is in its nature only for drink itself" (439a4–7). This stipulation provides a crucial step in Plato's argument, since it secures the claim that when a thirsty person decides not to drink, it must be because of a *second*, distinct part of his soul, which rejects the drink as a result of rational calculation. Still, exactly what Socrates means to claim is not so obvious.[33] For he here explicitly excludes the possibility of thirst being for hot or cold drink, or much or little drink; yet he also allows that "where heat is present as well as thirst, it causes the appetite to be for something cold as well, and where cold for something hot" (437d8–e2). In other words, he recognizes perfectly well that the actual appetites we experience often *are* qualified, insisting only on the abstract point that those qualifications form no part of *thirst as such*. This seems a point of metaphysics rather than psychology, amounting to an application of the general principle that correlatives are either both unqualified or both qualified. In this, it is akin to a number of other Academic-looking quasi-digressions in the central books of the *Republic*, such as the analysis of *dunameis* in Book V (477c-e). Indeed, it looks rather like a precursor of that analysis, defining a particular kind of *dunamis*, thirst, strictly in terms of what it is 'set over,' namely drink.

So it is not clear that Plato's discussion of thirst really intends to reject the Desire thesis. A further reason to doubt that it does so is that (as I noted earlier and will discuss in section III) Book VI of the *Republic* presents what looks very much like a restatement of the Appearance thesis. Obviously it would be preferable to find a reading on which the two passages are compatible. One possible solution is to say that just as a city may take wealth or freedom as its good (*Rep.* 562b–c), thirst is the appetitive drive that constitutively *takes drink as its* good. This involves attributing to the appetitive part of the soul a certain amount of cognitive equipment, sufficient for performing the two operations I sketched above; but then there is strong independent evidence in the *Republic* for doing so.[34]

Now at this point we seem to have come around again to the problem raised by my discussion of perverse desires, the problem of what exactly it *means* to say that some desire depends on taking its object *as* good. For it sounds like little more than saying that, for instance, the appetitive soul takes pleasures to be desirable, or just that it does constitutively desire them. So the Desire thesis seems to be threatened by a dilemma, or a Scylla and Charybdis. If 'good' is construed in wholly formal and nonrestrictive terms, so that perverse and primitive desires are no counterexample, then the thesis risks collapse into the tautological-
p. 46 sounding claim that ↳ desire is for the desirable or the desired. In that case empirical adequacy comes at

Downloaded from https://academic.oup.com/book/4008/chapter/145635901 by University of Toronto user on 11 March 2024

the cost of vacuity. The framework of mid-level values may still be useful, for it applies to primitive and perverse agents as easily as any others, and indeed provides a way of defining them. We may say that a perverse agent is one guided by a bizarre, irrational or incoherent set of mid-level values, while a primitive agent is one governed by a narrow and inflexible set of them (e.g., 'drink' and nothing else), and incapable of rational reflection thereon. But that does not yet tell us what work the concept *good* is doing here.

It will help at this point to compare the Desire thesis to a similar-sounding view that Plato clearly does *not* intend. According to Thomas Hobbes, "Whatsoever is the object of any mans Appetite or Desire; that is it, which he for his part calleth Good: And the object of his Hate, and Aversion, Evill.... For these words of Good, Evill, and Contemptible, are ever used with relation to the person that useth them: There being nothing simply and absolutely so" (*Leviathan*, Part I, ch. 6). In other words: "The common name for all things that are desired, insofar as they are desired, is good; and for all things we shun, evil" (*De Homine* I.xi.4).[35]

Now this subjectivist account of Good and Evill is, I take it, exactly what Plato is *denying* when he says that desire is for the good.[36] For though he and Hobbes agree in identifying the good with the desired in extension, they differ on the all-important question of explanatory direction. Plato insists that we only desire what we antecedently believe to be good; Hobbes, that the description 'good' doesn't correspond to any real or even projected property of things, but is simply attached by fiat to whatever we happen to desire. On the Desire thesis, contra Hobbes, desiring is a teleological business governed by intrinsic norms of success. When we desire *there is something we are trying to get right*, and we get it right when we desire what is really valuable. To anticipate a bit, this will be the key to understanding how the Reality thesis can be not only compatible with but derivative from the Appearance thesis. The force of the Reality thesis is simply to make this claim about world-guidedness explicit, clarifying that in pursuing what *seems* good to me I aim at what really *is* good. Together, the two versions of the thesis present a position I will refer to as *cognitivism* about desire: the claim that our desires are causally contingent on positive evaluative beliefs about the objectively real, antecedently valuable properties of their objects.[37]

So read, the Appearance thesis is neither implausibly restrictive nor vacuous. It cannot be restrictive, since it sets no limit to the *range* of objects on which desire may fall. At the same time, the merely formal role of 'good' here does not render the thesis vacuous. For it is used by Plato to make the substantive and highly controversial claim that desiring is an activity structured by norms, in which we undertake to be properly guided by the way the world is. Put in other terms, his claim is that reasons, in the form of evaluative perceptions and cognitions, govern and explain our desires, and not vice versa.[38]

It is not hard to see why Plato might present the Appearance thesis so understood as intuitively intelligible and widely acceptable, as he does in the *Protagoras* (358b–c), *Republic* (505d–e), and *Philebus* (20d). (I have also argued that the argument for it in the *Meno* is really just a matter of conceptual clarification.) Long before Plato's time it had been taken as obvious that we naturally pursue what seems beneficial, so that self-harming actions require special diagnosis as involuntary or abnormal. In the *Iliad*, when the warrior Glaucus exchanges his gold armor for bronze, Homer explains: "The gods took away his wits" (VI.234–36). In Gorgias's *Defense of Palamedes*, Palamedes declares: "All people perform all actions for the sake of these two things: either to gain some profit or escape a loss" (19).[39] Gorgias's *Encomium of Helen*, the most intriguing pre-Platonic work on moral responsibility, is a very enigmatic and controversial work, but it can be read as an extended exploration of the same principle, arguing that any manifestly self-destructive action, such as Helen's flight to Troy, must have been caused in such a way as to count as involuntary. Strikingly, when Socrates first proposes the Appearance thesis in the *Protagoras*, it is as an uncontroversial starting point for argument, in less need of defense than the hedonism that accompanies it; and it is accepted without demur by the circle of sophists gathered round—otherwise a rather contentious group (358c; see also 335d–338b, 351b–e). So we might plausibly tell a story according to which a commonsensical, traditional presumption that voluntary action aims at the good of the agent came to be made explicit by the sophists (a group including Socrates, in the eyes of his contemporaries), and adopted as

Downloaded from https://academic.oup.com/book/4008/chapter/145635901 by University of Toronto user on 11 March 2024

an exceptionless first principle for the philosophical explanation of action. In that case, what is distinctively Platonic here is not, after all, the Appearance thesis as such, but rather his use of it *as a vehicle for cognitivism*—his faith that, properly understood, this commonsensical view commits us to an objectivist conception of the good as object of desire. To see how he thinks it does so, we first need to look closely at the Reality thesis.

III

Though I have argued for its fleeting presence in the *Meno*, the *locus classicus* for the Reality thesis is *Gorgias* 466a–68e. Here Socrates insists to an outraged Polus that orators and tyrants have no real power, at least if power is presumed to be a good thing; for to do as one sees fit without intelligence is a bad thing (466e–67a). That orators and tyrants act without intelligence is made out a bit later on when Socrates argues that injustice in one's soul is the worst condition anyone can experience (474c–81b), which implies that anyone who seeks out the power to do injustice with impunity is acting unintelligently. But before then Socrates argues for the even stronger claim, not strictly necessary for his broader argument, that the unjust tyrant or
p. 48 orator does not even do what he *wants*. Socrates starts ↳ from noting cases in which it would sound wrong to say that someone wants [*boulesthai*, used consistently throughout the argument] to do the action he performs, and right to say that he wants some distinct end to which it is a means: taking prescribed medicine in order to become healthy, for instance, or undertaking a dangerous sea voyage to make money. Socrates then gets Polus to agree to the general principle that when we act for an end, what we want in acting is that end, which is always some good (467d6–e1).

Socrates initially flip-flops on the question of whether I should be said to derivatively want an action I perform as a means to some end, or *only* the end itself (467d6–e1; cf. 468b8–c1). The appropriative conception of desire suggests the latter, since the objects of desires are properly speaking the 'things' to be 'acquired' by our actions, not the actions themselves. Still, Socrates soon allows that "we want to do these things if they are beneficial, but if they're harmful we don't" (468c). Evidently actions can inherit derivative standing as objects of desire, from the goods for the sake of which we perform them. What is important for Socrates' argument is that there is always a conceptual *gap* of some kind between an action and the prospective benefit in virtue of which we perform it, so that an action can always fail to attain its end and thus fail to have value. This seems fair enough as a characterization of teleological action—that is, action performed 'for' a distinct end, no matter how narrow the distinction between the two. Even if I play basketball simply for the joy of playing basketball, the result I desire is something distinct from the action I can perform, and the two can come apart: This fallibility seems to be part of what it means for an action to have an end. (Of course, whether all voluntary actions *are* teleological in this sense is another question.) The questionable move is the further inference that when the performance of an action fails to obtain the value desired, the action itself fails to count as desired. Exactly how Socrates reaches this inference is controversial,[40] but, schematized, the key moves of his argument seem to run as follows (468b–e):

> A. 'It's for the sake of what's good that those who do all these things do them': Archelaus wants to kill his enemies if it is better for himself to do so. [*Appearance thesis*] (468b7–8)
>
> B. If it is better for himself, Archelaus wants to kill his enemies; if it is not better for himself, Archelaus does not want to kill his enemies. (468c3–5)
>
> C. If Archelaus kills his enemies, and it is not better for himself to do so, Archelaus does not do what he wants [*Reality thesis*]. (468d1–7)

Socrates' argument here appears to turn on an equivocation at (B),[41] the dangerous ambiguity of which is signaled by Polus' marked reluctance to assent at 468c6–7. For claim (A) presents *anticipated* benefit as a

Downloaded from https://academic.oup.com/book/4008/chapter/145635901 by University of Toronto user on 11 March 2024

cause of present desire, as per the Appearance thesis; claim (C) presents *real* future benefit as a *criterion* for ascriptions of present desire; and (B) seems to pivot between the two.[42] Archelaus might well assent to (B),
p. 49 or more ↳ simply to the conditional: 'If it doesn't benefit me, I don't want to do it.' But in doing so he would mean to endorse what is really a future-oriented subjective principle equivalent to (A): if I don't *think* it will benefit me, I *won't* form the desire to do it. Socrates instead infers a present-tense claim about the status of the desires that Archelaus in fact avows. So, though the presentation is not quite so elliptical as in the *Meno*, the *Gorgias* too seems to illegitimately infer the Reality thesis from the Appearance one by equivocation.

Worse, (A) and (C) seem to give conflicting diagnoses of Archelaus' condition. The Appearance thesis (A) implies that Archelaus' unjust actions are caused by his desire for wealth, which he takes to be good. That seems plausible enough; but how then does he *not* do what he desires in so acting? Moreover, if we are to accept the Reality thesis as Plato's final diagnosis here, it needs to be supplemented by two things that are difficult to supply. The first is an account of what *does* cause Archelaus' action. What is this 'seeing fit' that motivates misguided action, and how does it motivate us? It seems that an attitude of 'seeing fit' performs all the functional roles of a desire without being a desire—a rather fishy status given that, as we saw, the Desire thesis is supposed to be about conation across the board. The other is some account of how these mysterious real desires for the real good actually contribute to our psychological economy. Without explanations on these two points, Plato seems just to be feebly stipulating that motivations of which he disapproves should not be counted as real desires; and that we all really do desire whatever he thinks we ought to. It would be hard to see this as anything other than a confusion of the descriptive and the normative.[43]

I will return in a moment to consider our interpretive options here, but it is worth noting first that this is far from the end of Socrates' argument. At this point it remains an open question what the good that we desire consists in. Socrates' conclusions at 468c–e are scrupulously conditional, viz that the unjust tyrant does not do as he wishes *if* those unjust actions are bad for him; and this shows only that *it is possible* to hold political power yet not do what one wishes. His later, scandalous conclusions that the tyrant Archelaus is miserable, and that the unjust act involuntarily, depend on two further tranches of argument that are soon provided: the empirical claim that Archelaus is in fact unjust, established by Polus's recounting of his story at 470c–73d, and the argument at 473d–77e by which Socrates establishes that injustice in the soul (if one escapes punishment for it) is the worst evil one can possess. The conclusion that Archelaus is miserable then follows (479a–e); and only later, in a somewhat inaccurate moment of retrospect to Callicles, does Socrates throw in the corollary about involuntariness:

> Do you think Polus and I were or were not correct in being compelled to agree in our previous discussion when we agreed that no one does what's unjust because he wants to, but that all who do so do it involuntarily [*akôn*]? (509e)

p. 50 In fact, at no earlier stage of the argument had Socrates said anything about the voluntary: but the equation of the not-wanted and the involuntary is apparently supposed to be unproblematic.[44]

For our purposes the crucial phase of the argument is Socrates' fancy footwork at (A)–(C) above, with (C) understood as a negative formulation of the Reality thesis. I will briefly note what seem to me some of the more promising interpretive options here: most are owed to or inspired by Terry Penner, but I will not here engage with the complexities of his accounts of them, and will just skim the surface of considerations for and against each. One such option would be to construe the relation of a desire to its object as *referential*, and as constituted by successful relation to that object.[45] Given the reading of the Appearance thesis that I offered earlier, on which a desire depends on our taking its object to have some positively valued property, we might think that the object of a desire is properly *the goodness of the value instantiated by that object*. For instance, strictly speaking, what Archelaus desires is 'the benefit of the pleasures of tyranny.' Since that

Downloaded from https://academic.oup.com/book/4008/chapter/145635901 by University of Toronto user on 11 March 2024

object does not exist (tyranny does not have pleasures, or if it does they are not beneficial), the desire is null and void; it fails to be a desire, just as a would-be sentence with a nonreferring subject term fails to be a sentence (*Sophist* 262e). We can perhaps make this more plausible, or at least intelligible, by thinking of a desire as a 'power,' a *dunamis*, and thus by nature fixed on its correlative object (*Rep.* 477a–d). (Perhaps we can no more desire the bad than we can see the audible or hear colors.) However, though these potential connections to other Platonic principles are intriguing, it is hard to see how this reading can meet the twin desiderata noted above. If desires for the bad fail to be desires at all, what does cause Archelaus's action when he does injustice? (Of course, if his unjust action is *involuntary*, one might argue that there is really *no* action here and thus no causally efficacious motivation either: but I will soon argue that this is too strong a sense to give to Plato's 'involuntary.') And what role does his alleged desire for the good ever actually play in his agency?

Alternatively, and more straightforwardly, we might insist that desire for the good—the real good—is directly in play in causing *every* action. For the *only* desire we have is a general standing desire for 'the good, *whatever it may turn out to be*.'[46] When I 'see fit' to φ, the cause of my action is simply the conjunction of that standing desire with, as a sort of minor premise, a belief that φing will satisfy it. This does fairly well at satisfying our two desiderata, for it gives desire for the real good a genuine psychological role, and gives an account of 'seeing fit' without introducing good-independent conation. On the other hand, this option involves attributing to Plato the One Big Desire picture that I rejected earlier—and that in an especially strong form, with no genuine conative 'trickle-down' to any desiderata more specific than happiness. Besides sounding un-Platonic, this is hardly an intuitively plausible picture of human desire. We are all
p. 51 familiar with desires having the form 'I want x, whatever x may turn out ↳ to be'—to go to the best restaurant in Chinatown, say, quite independently of any beliefs one might have about which restaurant that is. And, as this reading helps to bring out, the Desire thesis implies that we all have a general, unspecified desire of this kind for 'the good, whatever it may turn out to be.' Still, we experience these unspecified desires as distinctive in *kind* [47]—it is hard to accept that *all* our desires are structured in this way, still less that all really amount to a *single* desire having this form. Moreover, on this reading as on the first, there is *nothing* of which the Appearance thesis is true—we have, properly speaking, *no* conative attitude toward the apparent-but-not-real goods that we mistakenly pursue. Yet it is hard to see how the Appearance thesis is dispensable. It figures, after all, as premise (A) in the argument above, and seems to be the basis for the Reality thesis in the *Meno* as well.

Another alternative would be to say that whatever desire Archelaus acts on when he commits injustice is *trumped* by a stronger or deeper desire that conflicts with it. This possibility is suggested by the parallel case of belief as disclosed by the Socratic *elenchus*.[48] It is no accident that the *Gorgias* argument for the Reality thesis leads into a methodological debate in which Socrates comes as close as he ever does to explaining his dialectical method, using as a case study his plan to elicit from Polus an affirmation that (contrary to his initial avowals) it is worse to do injustice than to suffer it (468e—74b). Plato thus frames the Reality thesis and the elenctic display as twin exercises in depth psychology. We may well *neither* avow what we really believe *nor* pursue what we really desire, and either lack of self-knowledge will lead us to self-frustrating behavior. And there is more than a parallel here. Socrates' two claims are meshed together, for the evaluative beliefs about doing and suffering injustice that he scrutinizes in the *elenchus* are actually phrased in motivational terms: Socrates and Polus are arguing about which course of action the other would *really* 'prefer,' 'want,' 'welcome,' and 'take' (474c), and who is really to be envied. And it is, of course, the Desire thesis that licenses this taking of conative attitudes as proxies for beliefs about better and worse.

Now in the discussion about the *elenchus*, Socrates presents his endeavor as one of showing his interlocutor what he really believes, deep down: "For I do believe that you and I and everybody else consider doing what's unjust worse than suffering it, and not paying what is due worse than paying it" (474b). That is, here as always, the *elenchus* is supposed to result in the interlocutor giving up on his initial mistaken avowal

Downloaded from https://academic.oup.com/book/4008/chapter/145635901 by University of Toronto user on 11 March 2024

rather than its latent contrary, which is to be accepted as in some sense his true position all along (476a, 480a). But Socrates will go on to insist, against the equally misguided and refutable Callicles, that the price of his false avowed moral beliefs is a kind of *disharmony*—a lifelong mental conflict and psychological incoherence (481e–82c). And it is hard to see how that can be the case unless his *false* beliefs are
p. 52 psychologically real as well, even if the latent ones are somehow deeper or more truly representative. There ↳ is obviously scope for a parallel claim in the case of desire. For our standing desire for happiness could both conflict with and be reasonably assumed to trump all our more specialized desires, without the reality of either coming into question. (Nothing in the Desire thesis, as I understand it, precludes conflicting desires; it only insists that they must be based on conflicting evaluative beliefs.)

I think there must be a grain of truth to this interpretation (and will shortly spell out what it seems to me to consist in). At any rate, Plato's intertwining of the cases of desire and belief here cannot be accidental. However, Socrates does not explicitly diagnose Polus as having conflicting desires. (Perhaps Plato does not yet see how such conflicts are possible: in Book IV of the *Republic*, conflicting desires will be the basis for the individuation of psychological parts, and there is no account of such parts in the *Gorgias*, though there may be hints in their direction.) It might also be objected that if Archelaus has a real (albeit trumpable) desire to exile his enemies, it is hard to see how Socrates can claim, as he later will, that his action is *involuntary*. That it does not represent what he *most* wants, and indeed conflicts with it, may make his action less than fully endorsed and autonomous, perhaps even akratic. But it is not obvious that we should understand Socrates' claim of involuntariness as meaning no more than that (though I will argue shortly that this is probably right).

In short, it is hard to come up with a fully satisfactory explanation of how the Reality thesis is to be understood. As I have already suggested, the best solution seems to be to understand the Reality thesis in somewhat deflationary terms, as a clarification. It specifies that 'good' in the Desire thesis is to be understood in objective rather than subjective terms, and thus makes explicit the cognitivist conception of desire that I offered as a reading of the Appearance thesis in section II. This is why it is introduced so casually in the wake of the Appearance thesis in the *Meno* and the *Gorgias* alike: not because it is being fallaciously inferred, but because it is to Plato's mind merely a disambiguation. And an intuitively reasonable one at that—after all, in garden-variety contexts we assume that someone whose desire is appropriately described as a desire for an *x* desires a *real x*, not a merely apparent one.

A passage in *Republic* VI provides support for this reading. In leading up to the analogy of the Sun, Socrates explains that knowledge of the good must belong to the Guardians. Socrates goes on to explain the special status of the good as follows:

> In the case of just and beautiful things, many people are content with what are believed to be so, even if they aren't really so, and they act, acquire, and form their own beliefs on that basis. Nobody is satisfied to acquire things that are merely believed to be good, however, but everyone wants the things that really *are* good and disdains mere belief here.
>
> —That's right.
>
> p. 53 ↳ All soul [*hapasa psuchê*] pursues the good and does everything [*panta prattei*] for its sake. It divines that the good is something but it is perplexed and cannot adequately grasp what it is or acquire the sort of stable beliefs it has about other things, and so it misses the benefit, if any, that even those other things may give. (505d5–e5)

Socrates initially presents it as a distinctive feature of desires for the good that they are for the real thing, as if desires for justice or beauty might be for their merely apparent instantiations. But this is immediately followed by the affirmation that the soul does *everything* it does for the sake of the good.[49] So, in fact, *all* our

Downloaded from https://academic.oup.com/book/4008/chapter/145635901 by University of Toronto user on 11 March 2024

desires have this orientation to reality. There is no contradiction here, since as Plato brings out there *is* still a sense in which someone might fairly be said to desire 'apparent justice.' This would be a true if potentially misleading way of describing a desire—not a distinctive kind of desire for justice, nor a desire for some defective species of justice, but a desire for the real goods to be obtained by possessing 'apparent justice'—i.e., by possessing a *reputation* for justice. Such goods would include security from punishment, presumably, and the esteem of one's neighbors;[50] and these are things that, unlike justice, the agent takes to be *really* good. No desire is *most* perspicuously described as a desire for an apparent x, and if I avow a desire for x *simpliciter* you are licensed to assume that I want the real thing. But in some cases it may still be useful to describe a desire as one for an apparent x, where apparent x's are reliable means to, or proxies for, real (and really valued) y's.[51]

The *Republic* passage thus makes explicit the connection between the Appearance and Reality theses. The Appearance thesis claims that we always pursue what we think good: 'good' is always a perspicuous (if somewhat underspecified) description of the object of desire. *Republic* VI tells us that if 'x' is a perspicuous description of the object of my desire, then my desire is for a *real* x. And that means that my desire—always—is for the real good. What *that* means in turn is that the Appearance thesis cannot be reduced to the subjectivist claim that I call whatever I happen to pursue good. Rather, I pursue what seems good *as an attempt* to obtain what really is so. None of this requires Plato to deny the psychological reality of the desires we avow and are moved by, including those oriented to bad objects; his claim is just that there is something we are *trying to do* when we desire, at which we fail when we desire the bad. (That is why Socrates can casually allow, at the end of the *Meno* argument, that wretchedness is a matter of wanting [really] bad things and getting them. This is not a lapse from the Desire thesis properly understood, but a corollary to it [78a7–8].)

The obvious objection to this interpretation of the Reality thesis is that it is insufficiently radical. For so read, the thesis does not properly entail that Archelaus does not do what he wants, at least not in such a way as to render his actions involuntary. As Santas pointed out regarding the person who confuses the salt and
p. 54 pepper shakers, *in a sense* he is not ↳ doing what he wants to do; but Vlastos remains equally right that in a different sense or respect an Oedipus, say, also *does* want to perform the misguided action. Whether a desire is best described in terms of the 'intended' or the 'actual' object seems to be, as I noted in section I, somewhat context-dependent. And if desires may at least sometimes be usefully ascribed in terms of a mistaken actual object, Socrates is surely not licensed to state without qualification that Archelaus acts unwillingly [*akôn*].

How far Plato can be defended from a charge of fallacy or confusion on this point is a delicate question. But my deflationary reading of the Reality thesis suggests that we might also take the corollary about involuntariness in a weakened sense. And there are some independent grounds for doing so. For it is important that we not retroject on Plato Aristotle's very restrictive account of involuntariness in *Nicomachean Ethics* Book III, which seems deliberately designed to clamp down on a broader and more flexible earlier usage. The very fact that Archelaus 'does as he sees fit' would be sufficient to class his actions as voluntary by Aristotle's standards; but in a way Plato can agree. For in denying that the tyrant does what he wants, Plato clearly does *not* mean that his action is not attributable to him, in the manner of a reflexive physical movement or some completely inadvertent behaviour. On the contrary—it is crucial to Plato's argument in the *Gorgias* that the 'involuntary' wrongdoer is morally responsible for his actions, for which he will be punished in the afterlife. The function of the involuntariness claim is to bring out that there is *a further threshold* of free agency, one higher than mere 'doing as one sees fit,' which the wrongdoer fails to meet.[52] He fails to meet it because he makes a mistake in forming the more determinate desires by which he aims at the good, and thereby fails to do what he *most deeply* wants to do. He thus deserves pity and reeducation as well as punishment and condemnation.[53] This is, I take it, the important truth brought out by the reading of the Reality thesis on which the tyrant's unjust desires are 'trumped' by other, deeper ones in

Downloaded from https://academic.oup.com/book/4008/chapter/145635901 by University of Toronto user on 11 March 2024

conflict with them—in particular, by the underspecified standing desire we all have for 'the good, whatever it may really be.'

For Plato, at least part of the point of the Reality thesis is to warrant the positing of this higher threshold for agency. The idea is that we can measure the proximate, behaviorally manifest desires of the wrongdoer against his standing desire for the genuinely good; insofar as they veer off target we can diagnose him as failed by his own lights. The tyrant may in a sense do what he wants, but *he does not want as he wants to want*. To that extent his action is inadvertent, self-frustrating, and unfree, and the desire itself is inauthentic, false to his own aims in desiring.[54] We would be more likely to call this higher threshold *autonomy*, and at least some philosophers today would explain it in terms of efficacious second-order desires: Plato calls it acting willingly, or doing what one wants.[55]

Downloaded from https://academic.oup.com/book/4008/chapter/145635901 by University of Toronto user on 11 March 2024

p. 55

IV

I have argued for a reading of the two formulations of the Desire thesis as together expressing a cognitivist conception of desire. On this reading, the thesis belongs to a group of Platonic positions and arguments that might all be loosely described as giving an objectivist or realist account of central features of human agency. In the *Cratylus*, for instance, Socrates argues that naming is an activity that can be performed correctly or incorrectly, by arguing that this is true of agency in general:

> So an action's performance accords with the action's own nature, and not with what we believe. Suppose, for example, that we undertake to cut something. If we make the cut in whatever way *we* choose and with whatever tool *we* choose, we will not succeed in cutting. But if in each case we choose to cut in accord with the nature of cutting and being cut and with the natural tool for cutting, we'll succeed and cut correctly. (386e–87a)

This argument takes as its explicit starting point the rejection of Protagorean relativism; but it can also function as an independent argument for that rejection.[56] For it is criterial for agency that not everything can count as success: I must aim at some determinate outcome that my performance may or may not bring about. There is a parallel here with Socrates' arguments against Protagoras in the *Theaetetus*. The famous *peritropê* argument claims that Protagoras is committed to endorsing as true the claims of those who reject his theory (170a–71c); the precise flow of the argument is notoriously difficult to spell out, but Myles Burnyeat has argued plausibly that it aims to show that some commitment to objectivity is built into the act of assertion itself.[57] And a parallel point is made in more restricted terms by Socrates' final argument against Protagoras, which turns on the possibility of predicting wrongly (177c–79b). I *cannot* make a prediction about some matter of which my current state stands as truth-maker. A prediction can only count as such if it incorporates some risk of falsification by the future. Both asserting and predicting are only intelligible as *fallible* practices, and Plato's claim is that this fallibility implies at least a limited realism; that is, commitment to a truth understood to outrun our capacity to shape it.

What the Desire thesis claims is that, as with naming, acting, predicting, and asserting, so too with desiring. I have already noted the parallels between desire and the assertion of belief in the context of the elenchus; desire is even more closely akin to action and prediction, for it is inherently forward-looking. Recall that Socrates in the relevant passages identifies the good with the beneficial, so that the claim that something is good is really a kind of prediction about its effects if acquired. Like all these other activities, desiring is purposeful and world-guided, for when I desire there is something I am trying to do. Thus desiring too is an activity that can go well or badly for us; and we can expect it to go no better than our thinking.

p. 56

On all these fronts Plato is exploring the idea that a certain commitment to an objective reality is built into our ordinary understanding of cognition and action as fallible. And it is natural to see cognitivism about

desire as fundamental here, for if what we want is the real good, this had better be what our thought is about and what our actions aim at. At the same time, cognitivism itself stops short of offering any purchase against an enlightened subjectivism or hedonism. Archelaus can agree that in pursuing the apparent good, what he desires is *really* to be benefited; he can still insist that, as a matter of objective fact, his good is to be found in subjective states about which, in the long run, he cannot be mistaken. But this is already a long way from the infallible subjectivity of Hobbesian desire. If Plato is right about the structure of acting and desiring, some possible accounts of value are excluded: and the way lies open to a fully realist conception of the good.[58]

Notes

1. Quotations from Plato are as translated by the various hands in the Hackett *Complete Works*, in some cases with revisions (John Cooper with D. S. Hutchinson, eds., *Plato: Complete Works*. Indianapolis: Hackett, 1997).

2. See, for instance, Irwin 1995: 223–43.

3. A further wrinkle is that, strictly speaking, both theses can be presented in either a positive version (all desire is for the good; everyone desires the good) or a negative one (no desire is for the bad; nobody wants what is bad). Plato seems to treat these formulations as interchangeable, and I will do the same. In some contexts he also, more dangerously, seems to conflate not wanting to ϕ with wanting *not* to ϕ, as if there were no difference for his purposes between the absence of a desire and an aversion as possible responses to what is not good.

4. Austin 1979: 86–89.

5. Santas 1964: 149-–57; and 1979: 185–89.

6. The *Protagoras* even seems to flop back and forth between the Appearance and Reality theses (my emphases): "No one who *knows or believes* there is something else better than what he is doing, something possible, will go on doing what he had been doing when he could be doing *what is better*" (358b7–c1); "no one goes willingly toward the bad or what he believes to be bad; neither is it in human nature, so it seems, to want to go toward what one believes to be bad instead of to the good" (see 358c6–d2). But these passages are probably better read as expressions of the Appearance thesis, with reference to beliefs occasionally omitted for the sake of concision.

7. As do Carone 2001 and Moss 2008.

8. Kamtekar 2006: 150.

9. Segvic 2000: 9–11. For some objections, cf. Kamtekar 2006: 138. Terry Penner's objections to 'special sense' readings also seem applicable (1991: 149).

10. See Penner 1987, 1988, and esp. 1991; Penner and Rowe 1994 and 2005; also Reshotko 2006; and Rowe 2007. Penner's
p. 57 arguments are so complex and detailed that it is virtually impossible to do them justice *en passant*; I fear I ↳ cannot properly engage with them here, but only draw on them as useful for developing my own view.

11. How exactly the argument is supposed to work has been much debated. For differing analyses, see Santas 1964: 150–57; Penner and Rowe 1994; Kamtekar 2006: 150–53; Weiss 2001: 32–38; and Scott 2006: 46–53. Penner and Rowe seem to me importantly right in insisting that the Reality thesis must be in play here if Meno's position is really to be engaged (16 with n. 21). But what Meno affirms at 77c5 is that people desire bad things *both* under that description *and* believing them good. So if he is to be refuted, *both* options must be eliminated, i.e., the Appearance and Reality theses must both be made out. Penner and Rowe (18–22) rely on what seems to me a very strained reading of 77d7–e2 in order to eliminate reference to people who desire bad things thinking them good (see Kamtekar 2006: 152–53).

12. For Aristotle, *boulêsis* and *epithumia* are technical terms for rational desire and nonrational appetite, respectively, and there are precursors for such a distinction in Plato (*Charm*. 167d–e; *Prot*. 340b). In the *Gorgias, boulesthai* is consistently used in relation to the Desire thesis at 466eff., while later on Socrates and Callicles discuss appetites, *epithumiai*, without any apparent regard for the earlier thesis (though see Carone 2004). Their discussion also seems to look forward to the

Downloaded from https://academic.oup.com/book/4008/chapter/145635901 by University of Toronto user on 11 March 2024

Republic in taking those appetitive desires to have a distinctive 'location' in the soul (493bff.). Still, these contexts are all casual or *ad hominem*. It is certainly plausible that Plato's deployment of *boulesthai* and *epithumein*, in the *Meno* and elsewhere, reflects a division of labor in contemporary usage, and that Socrates gets some persuasive mileage from their differing connotations. It is much harder to believe that the early dialogues operate with a theoretical typology of human desires that Plato at no point articulates or defends.

13. Kamtekar 2006: 127, 150–51; Weiss 2001: 35–39. See also Santas 1964: 152n15; Segvic 2000: 14–15.

14. Even in the *Republic*, every part of the soul, reason included, has its own *epithumia* (580d7), and *boulesthai* is used of the lower parts (439b1)—as if to challenge any quick assumptions we might make that the motivations of the different parts are different kinds of thing. As Lorenz (2006: 45–46) notes, even in the *Republic, epithumia* does not *mean* 'low appetite,' but something more like '(intense) desire' (see Carone 2001: 122n30).

15. The only possible exception is one that Kamtekar emphasizes, Socrates' rhetorical question at 78a7–8: "For what else is being miserable but to desire [*epithumein*] bad things and to possess them?" She takes this to imply that the results of Socrates' argument must be compatible with the claim that some people *do epithumein* what is bad, and that therefore *boulesthai* must be something different (150n44, 153). I take him rather to be emphasizing that, on reflection, Meno's own definition of happiness at 77b2–5, to which this line clearly alludes, supports the conclusion that no one could (lucidly) want bad things.

16. Features common to the two passages include the following: (1) a preliminary replacement of *kala* by *agatha* as the object of desire (*Symp*. 201c, 204c–d; cf. 202c10–d5; *Meno* 77b6–7; cf. also *Lysis* 216c); (2) the specification of desire as desire to secure for oneself (*genesthai hautô; Symp*. 204e4, 206a6–8; *genesthai autô, Meno* 77c7–8); (3) the assumption that we become happy by ↳ acquiring good things (*Symp*. 205a1, d2; implicitly, *Meno* 77e–78b); (4) the claim that all desire is alike in being for the good (205a–b, d–e; *Meno* 78b4–6); and (5) as noted above, the treatment of *epithumein* and *boulesthai* as interchangeable (*Meno* 77e–78a)—and *erân* as well, at certain points in the *Symposium*.

p. 58

17. *Philein* is rather different, since its paradigmatic usage is for a non-appropriative, (ideally) reciprocal attitude of affection toward another person. So, despite its overlap on many points with the dialogues discussed here, I will avoid making use of the *Lysis*, which offers a discussion (one that is in any case perplexing and apparently aporetic) of the object of *philia*.

18. As Scott (2006: 48n4) notes, the structure of argument, by elimination of alternatives, is Gorgianic.

19. See, for instance, *Euthyd*. 278e–82d; *Symp*. 204e–5a. In the *Republic*, when Thrasymachus challenges Socrates to say what he thinks justice is, he adds a preemptive attack: "And don't tell me that it's the right, the beneficial, the profitable, the gainful, or the advantageous, but tell me clearly and exactly what you mean; for I won't accept such nonsense from you" (336c–d). The implicit accusation has little basis in the *Republic* so far. Presumably we are to understand that Socrates habitually builds arguments around such formal conceptual connections, in a way that strikes others as evasive on the substantive questions.

20. See Vlastos 1991: 203–9 on the 'Eudaemonist Axiom.' As Vlastos notes, for Plato (and more explicitly for Aristotle), intrinsic goods seem to be those that are valued both as ends in themselves (i.e., severally and independently) *and* as constituents of happiness. (Vlastos compares enjoying a movement of a symphony both in itself and as part of the whole composition.) There is thus a conceptual difference between thinking of and valuing some object as a part of happiness and simply as good: Presumably each description has motivational force.

21. Properly speaking, as Socrates argues (to general assent) at *Meno* 87d–88d and *Euthydemus* 280c–81e, we are happy not by *possessing* good things but by being *benefited*, which involves using good things correctly. To do so requires intelligence—which, in turn, suggests that intelligence is the real source of the benefit and, therefore, the only unconditionally good thing. However, Plato never brings this argument to bear on his account of desire itself, which continues to be presented in strictly appropriative terms.

22. There is no doubt a slippery slope between thinking something good, thinking it better than some alternative, and thinking it best, especially in the context of occurrent desires in particular deliberative contexts. However, it seems to me significant that Plato usually prefers non-comparative locutions. (The exceptions would be *Protagoras* 358b7–d4 and arguably some moments of the *Gorgias* [466e2; but the 'better' at 468b2, b6, and d3 is in comparison to not doing the action in question, not to a less beneficial alternative].)

23. See Kamtekar 2006: 151–52.

Downloaded from https://academic.oup.com/book/4008/chapter/145635901 by University of Toronto user on 11 March 2024

24. Santas 1964: 154–56.

25. See McTighe (1984: 205–6) and Vlastos (1991: 148–54). Vlastos sees this move—and with it the whole of the Desire thesis—
p. 59 as a tragic mistake on Socrates' part, derived from a failure to distinguish between the actual and ↳ intended objects of desire. He adds that, had Socrates properly grasped the import of his own views, he would realize that (emphasis in original) "*he has no reason to deny that those wicked tyrants and their ilk do desire* those horrible things—assassination, etc.—which 'seem best' to them: under their misdescriptions of those actions as 'good' they most certainly do desire them" (153); thus "his famous doctrine that all wrongdoing is involuntary would dissolve" (154).

26. See Penner and Rowe 1994: 16n21; and Kamtekar 2006.

27. See Moss 2008 for the argument that, for Plato, nonrational desires are perceptions of value. See also Segvic 2000: 35.

28. This principle does important ethical work for Plato. In the *Philebus*, it shows that desire presupposes memory (since the desiring creature must have some awareness of the opposite of its current state), and thus that desire belongs to the soul, not the body (33c–35d). In the *Symposium*, it shows that the gods are not lovers of wisdom, and neither are those who mistakenly think they are wise already (204a). See also *Lysis* 215e, 217e–18b, 221e on desire as directed toward an opposite or what one is deficient in.

29. Such views would include conceptions of desires as cognitive but not doxastic: i.e., as identical with cognitive states, perhaps perceptual or perception-like ones, which fall short of full-blown belief. I have argued elsewhere on independent grounds that for Plato, to be 'appeared to' or to 'have an impression' just *is* for some part of the soul to adopt a belief, however preliminary and unreflective it might be (Barney 1992: 286–87). So the 'non-doxastic' option is not really available to him—though one might still see his view as akin to modern ones that, lacking the apparatus of the partitioned soul, parse desire in non-doxastic terms.

30. So far as I can tell, no Platonic text claims that we *always* desire what we believe to be good, still less that the strength of our desire for something always perfectly tracks the *degree* to which we believe it good. This seems a matter of common sense, at least in the case of general beliefs and occurrent desires for particular objects. I may believe that fish delicacies are very good indeed, yet feel no desire for them by the time the tyrant's banquet gets to the fifteenth course.

31. *Paradise Lost* IV.110.

32. Anscombe 1963: 75.

33. See Carone 2001: 116–21 and Moss 2008. On the broader question of whether the *Republic* repudiates the Socratic moral psychology of the earlier dialogues, see also Rowe (2007: 18): "Plato remains throughout essentially a *Socratic*."

34. Of course, it is an open philosophical question just what cognitive resources are required for beliefs, concepts, and so forth; but it has been widely accepted since Moline 1978 that the lower parts of the soul in the *Republic*, which regularly communicate and politick with the rational part, must be seen as having significant independent cognitive resources. For the depiction of the lower parts as having beliefs, see *Rep*. 442c–d, 574d–e, 605c (see also Barney 1992: 286–87; Lorenz 2006; Carone 2001: 117–21; Moss 2008). This is not to deny that the psychology of the *Republic* is in part driven by the problem of how to understand desires *we experience as* independent of, and resilient against, our considered beliefs
p. 60 about the good. I cannot discuss this question in detail here, ↳ but it seems to me that the psychology of the *Republic* should be understood as Plato's solution to some of the puzzles raised by the Desire thesis in earlier works, including the *Gorgias* (the status of the *epithumiai*, 492d–95a) and the *Protagoras* (the denial of *akrasia*, 352b–58c). Far from being a renunciation of the Desire thesis, the theory of the tripartite soul rescues it by showing that, *if conjoined with a suitably complex and psychologically realistic account of the agent who desires*, it can account for all the diverse phenomena involved in nonrational desire and motivational conflict.

35. Hobbes evidently intends this subjectivist Desire thesis as one about conation in general, since he holds that "of the voluntary acts of every man, the object is some Good to himselfe" (I.14). And Hobbes's claim that "Aristotle, and other Heathen Philosophers define Good, and Evill, by the Appetite of men" (IV.46) suggests that he believes (or wants the reader to believe) that this version of the thesis has a respectable ancient ancestry (presumably *N.E*. III.4).

36. This way of putting it raises the question of whether Plato might have had some view along Hobbesian lines already in his sights. In section IV I will say a bit about the Desire thesis as part of Plato's broader realism, which certainly was in part a response to the subjectivism of Protagoras and others. But whether Protagoras himself might have intended something

Downloaded from https://academic.oup.com/book/4008/chapter/145635901 by University of Toronto user on 11 March 2024

like the Hobbesian view as part of his 'Man is the measure' thesis seems to me unclear from Plato's own account of it in the *Theaetetus*. Rudebusch (1999: 27ff.) takes Polus's initial position in the *Gorgias* to be a kind of 'ethical Protagoreanism.'

37. Of course, there are other positions that might be appropriately called 'cognitivism,' including those on which desires just *are* evaluative beliefs; and these would contrast with Hobbes' view in much the same way. But, as I noted earlier, there is no reason to saddle Plato with this problematic conception of desire; his consistent practice of referring to thinking good and desiring as separate operations, one of which is the cause of the other, tells strongly against it.

38. This cognitivist conception of desire is brought out by Plato in a series of etymologies in the *Cratylus* that link desire and belief. Like *doxa*, 'opinion,' *boulê*, 'planning' has to do with trying to hit (*bolê*) some target, and '*boulesthai*' ('wanting') and '*bouleuesthai*' ('deliberating') signify aiming at something (*ephiesthai*). "All these names seem to go along with '*doxa*' in that they're all like '*bolê*,' like trying to hit some target" (420c5–6). Kamtekar (2006: 146–47) has a good discussion of the significance of this passage. As I argued earlier, for Plato a desire is not a belief; but at a sufficiently high level of generality, desiring *resembles* believing, as a different *kind* of attempt to 'grasp' or 'latch on to' the world as it is.

39. Unfortunately, the following sentence, which seems to be a diagnosis of self-harming actions, is apparently textually corrupt.

40. See, for instance, Penner 1991: 173–88.

41. See McTighe 1984: 206–7. Segvic 2000 defends the argument, on the grounds that Socrates has all along been speaking of 'good things' as the ends of action, rather than things thought to be good. But she admits that in that case Polus's assents depend on his misunderstanding the argument (44n39). Penner 1991 (149) argues that it is "an entirely successful refutation."

p. 61 42. To put the problem in different but perhaps equivalent terms, there seems to be something dubious in the way that the condition embedded in the content of Archelaus' desire in (A) is extracted from its scope in (B) and (C). This has been pointed out to me by John MacFarlane, whose work with Niko Kolodny on conditionals may well be relevant to a full understanding of the logical difficulties here; I hope to pursue these issues elsewhere. A further complication is that, as Sergio Tenenbaum points out to me, Socrates' conclusion should be that Archelaus wants *not* to do what will not benefit him—which is compatible with his *also* having a conflicting desire to do it, perhaps under a different description. I do not think that Plato means here to illicitly rule out the possibility of conflicting desires. Rather, he is for simplicity's sake conflating not wanting to ϕ with wanting not to ϕ, to avoid having to continually distinguish the three categories harmful-beneficial-neutral and the corresponding attitudes desire-aversion (i.e., desiring *not*)-neither.

43. Or perhaps a flight to a fanciful metaphysics of postulated 'real selves'; see Kamtekar 2006: 137 and McTighe 1984: 195ff. on the 'neoplatonic' reading.

44. The reference to a previous agreement with Polus is often taken (for instance, by Dodds 1959 and Zeyl 1986 *ad loc*.) as referring to 467c–68e; but this is at best an oversimplification, since the result there is still conditional (that is, injustice has not yet been shown to be bad).

45. This is influenced by, but not quite the same as, the proposal of Penner 1991, according to which actions should be individuated, for Plato's purposes, as incorporating *all* their consequences. See Kamtekar 2006: 141–43 for some serious difficulties with this reading: notwithstanding these, the account is suggestive and, I think, Platonic in spirit in that it vividly depicts the vicious person as a prisoner of *fantasy*, deluded by various cultural clichés about the sources of happiness (see Penner 1991: 188–89 on the 'Private Benjamin' problem).

46. See the fuller and more complex version in Penner and Rowe 1994: 3–10.

47. Moreover, an important group of texts suggest that Plato not only recognizes such desires as having a distinctive structure, but assigns them a distinctive role. The texts I have in mind relate to prayer. In the *Laws*, the Athenian observes that we all pray that our desires will be satisfied, but this is mistaken. We should pray not that all things follow our desire, but that our desire follows our rational judgment (687e). Socrates' prayer at the close of the *Phaedrus* leaves up to gods the determination of what external goods are appropriate to him (279b–c). Xenophon's Socrates prays the same way, "since the gods know best what things are good" (*Mem*. I.3.2; see *Cyr*. I.6.5). The (inauthentic) *Alcibiades* II teases out a Socratic argument for the claim that we should all pray in just this way. It is clear why such prayers are ethically appropriate for a Socratic. To pray is to officially register a desire; and to desire in this deliberately underspecified way expresses an

Downloaded from https://academic.oup.com/book/4008/chapter/145635901 by University of Toronto user on 11 March 2024

epistemic humility appropriate to our human limitations. So what Penner 1991 takes as the essential structure of human desire according to Plato seems rather to represent a Socratic ethical *norm*.

48. The parallel is emphasized by Kamtekar (2006: 143–48), who concludes that "as people latently believe truths, so they latently want good things" (148).

p. 62 49. All this is controversial: for contrasting interpretations, see Kamtekar 2006, esp. 154n52; Irwin 1977: 336n45; and Segvic 2000. Interpreters who take Book IV to assert the existence of good-independent desires must take the present passage as weaker in at least one of two ways. First, the crucial phrase *pasa psuchê* can be taken as referring not to *all* soul but to *every* human soul, and thus as referring to the desires of the rational part only. Second, the 'does everything for' [*panta prattein*] can be taken nonliterally, as an idiom loosely meaning 'goes to great lengths for.' Both readings are possible, but seem less natural than those I offer here.

50. This argument in *Republic* VI recalls a passage from Glaucon's earlier speech against justice in Book II. With a paradoxical rhetorical flourish, Glaucon imagines the defenders of injustice casting it as a kind of pursuit of truth: "A really unjust person, having a way of life based on the truth about things and not living in accordance with opinion, doesn't want simply to be believed to be unjust but actually to be so" (362a4–6). In the Book VI passage, Socrates recalls and reaffirms this point in earnest. The unjust man is indeed a seeker of the real thing—the real thing that he values, which is, of course, his own good.

51. This is different from the sense in which Plato or Aristotle may speak of a desire for wealth, for instance, as one for an apparent good; here the sense is that the agent *himself* takes wealth to be a real good, whereas 'apparent justice' in *Republic* VI is what will look like justice to other people. Granted this asymmetry, in both cases the point stands that a desire for an apparent x can always be redescribed more perspicuously as a desire for a real something—the real good in the case of desire for the apparent good, real security and esteem in the case of apparent justice, and so forth.

52. It is also worth noting that in later dialogues Plato defends the view that wrongdoing is involuntary by a somewhat different route. In the *Timaeus* and *Laws*, he is primarily concerned to describe bad states of *character* as involuntary, rather than wrong actions, on the grounds that no one would deliberately receive into his soul the evil of vice (*Laws* 731c–e, 860d–61d; *Timaeus* 86d–e). In these later dialogues, Plato's position seems to be the one attacked by Aristotle in *Nicomachean Ethics* III.5, that people only act wrongly on the basis of bad character, and bad character is itself involuntary.

53. In the *Protagoras*, Socrates insinuates the thesis that wrongdoing is involuntary into his exegesis of a poem by Simonides. He comments: "I am pretty sure that none of the wise men thinks that any human being willingly makes a mistake or willingly does anything wrong or bad. They know very well that anyone who does anything wrong or bad does so involuntarily" (345e). Here too, the idea of *mistakenness* is central.

54. See Plato's (notoriously difficult and controversial) account of false pleasures in the *Philebus*. Here too Plato insists that an affective or motivational state, in this case pleasure, can *directly* incorporate cognitive error. (At 38a–40c he explicitly acknowledges that this will strike most people as a category mistake, and bites the bullet.)

55. This way of looking at the Desire thesis is presumably unavailable to Plato since, at *Charmides* 167e, he seems to hold that second-order desires are impossible. Plato's reasoning there is dubious, but it is also not clear that the apparatus of
p. 63 higher-order desires would really be helpful to him here. As I will ↳ try to bring out in section IV, Plato's point is really one about the teleological, and thus fallible, character of desire at *any* level; to attempt to capture this formal feature of all desire through the relation of higher-order desires to lower-order ones seems misleading, and would invite an infinite regress.

56. For further discussion, see Barney 2001: 42–44.

57. Burnyeat 1990: 30–31.

58. This chapter began life as a paper for a seminar taught by Alexander Nehamas at Princeton University in 1991, and has been changing continually ever since. I am indebted for improvements to audiences at the Texas Workshop in Ancient Philosophy, the University of Chicago, the University of Toronto, the University of Pittsburgh, Lehigh University, Universität Köln, and the Townsend Working Group in Ancient Philosophy at the University of California, Berkeley; I fear I have lost track of all the individuals who have helped me to shed various mistakes, but they certainly include Tad Brennan, Matt Evans, Michael Green, Rachana Kamtekar, Jessica Moss, Alexander Nehamas, Terry Penner, Kieran Setiya,

Downloaded from https://academic.oup.com/book/4008/chapter/145635901 by University of Toronto user on 11 March 2024

and Roslyn Weiss. More recently, John MacFarlane, Gurpreet Rattan, and Sergio Tenenbaum have all raised puzzles about which I need to think more, and hope to address in the future.

Downloaded from https://academic.oup.com/book/4008/chapter/145635901 by University of Toronto user on 11 March 2024

References

Anscombe, G. E. M. 1963: *Intention*; 2nd ed. Ithaca, N.Y.: Cornell University Press.
Google Scholar Google Preview WorldCat COPAC

Austin, J. L. 1979: Other minds, reprinted in *Philosophical Papers*; 3rd ed. Oxford: Clarendon. 10.1093/019283021X.001.0001
Google Scholar Google Preview WorldCat COPAC Crossref

Barney, R. 1992: Appearances and impressions, *Phronesis* 37: 283–313. 10.1163/156852892321052533
WorldCat Crossref

Barney, R. 2001: *Names and Natures in Plato's* Cratylus. London: Routledge.
Google Scholar Google Preview WorldCat COPAC

Burnyeat, M. 1990: Introduction, *The Theaetetus of Plato*. Indianapolis: Hackett.
Google Scholar Google Preview WorldCat COPAC

Carone, G. 2001: Akrasia in the *Republic*: Does Plato change his mind? *Oxford Studies in Ancient Philosophy* 20: 107–48.

Carone, G. 2004: Calculating machines or leaky jars?: The moral psychology of Plato's *Gorgias, Oxford Studies in Ancient Philosophy* 26: 55–96.
WorldCat

Dodds, E. R. 1959: *Gorgias*. Oxford: Oxford University Press.
Google Scholar Google Preview WorldCat COPAC

Hobbes, T. 1934: *Leviathan*. London: Dent.
Google Scholar Google Preview WorldCat COPAC

Hobbes, T. 1972: *De Homine*. Garden City, N.Y.: Anchor.
Google Scholar Google Preview WorldCat COPAC

Irwin, T. 1977: *Plato's Moral Theory*. Oxford: Clarendon.
Google Scholar Google Preview WorldCat COPAC

Irwin, T. 1979: *Gorgias*. Oxford: Clarendon.
Google Scholar Google Preview WorldCat COPAC

Irwin, T. 1995: *Plato's Ethics*. New York: Oxford University Press. 10.1093/0195086457.001.0001
Google Scholar Google Preview WorldCat COPAC Crossref

Kamtekar, R. 2006: Plato on the attribution of conative attitudes, *Archiv für Geschichte der Philosophie* 88: 127–62. 10.1515/AGPH.2006.006
WorldCat Crossref

Lorenz, H. 2006: *The Brute Within: Appetitive Desire in Plato and Aristotle*. Oxford: Clarendon.
Google Scholar Google Preview WorldCat COPAC

McTighe, K. 1984: Socrates on desire for the good and involuntariness of wrongdoing: *Gorgias* 466a–468e, *Phronesis* 29: 193–236. 10.1163/156852884X00012
Crossref

Moline, J. 1978: Plato on the complexity of the psyche, *Archiv für Geschichte der Philosophie* 60: 1–26. 10.1515/agph.1978.60.1.1
WorldCat Crossref

Downloaded from https://academic.oup.com/book/4008/chapter/145635901 by University of Toronto user on 11 March 2024

Moss, J. 2008: Appearances and calculations: Plato's division of the soul, *Oxford Studies in Ancient Philosophy* 34: 35–68.
WorldCat

Penner, T. 1987: *The Ascent from Nominalism: Some Existence Arguments in Plato's Middle Dialogues*. Dordrecht: Reidel. 10.1007/978-94-009-3791-8
Google Scholar Google Preview WorldCat COPAC Crossref

p. 64

Penner, T. 1988: Socrates on the impossibility of belief-relative sciences, *Proceedings of the Boston Area Colloquium in Ancient Philosophy* 3: 263–325.
WorldCat

Penner, T. 1991: Desire and power in Socrates: The argument of *Gorgias* 466A–468E that orators and tyrants have no power in the city, *Apeiron* 24, no. 3: 147–202.

Penner, T., and Rowe, C. 1994: The desire for good: Is the *Meno* inconsistent with the Gorgias? *Phronesis* 39: 1–25. 10.1163/156852894321052225
Crossref

Penner, T. 2005: *Plato's Lysis*. Cambridge: Cambridge University Press.
Google Scholar Google Preview WorldCat COPAC

Plato. 1997: *Complete Works*. Edited by John Cooper and D. S. Hutchinson. Indianapolis: Hackett.
Google Scholar Google Preview WorldCat COPAC

Reshotko, N., 2006: *Socratic Virtue: Making the Best of the Neither-Good-nor-Bad*. Cambridge: Cambridge University Press. 10.1017/CBO9780511482601
Google Scholar Google Preview WorldCat COPAC Crossref

Rowe, C. 2007: *Plato and the Art of Philosophical Writing*. Cambridge: Cambridge University Press. 10.1017/CBO9780511482625
Google Scholar Google Preview WorldCat COPAC Crossref

Rudebusch, G. 1999: *Socrates, Pleasure, and Value*. New York: Oxford University Press.
Google Scholar Google Preview WorldCat COPAC

Santas, G. 1964: The Socratic paradoxes, *Philosophical Review* 73: 147–64. 10.2307/2183333
WorldCat Crossref

Santas, G. 1979: *Socrates: Philosophy in Plato's Early Dialogues*. London: Routledge & Kegan Paul.
Google Scholar Google Preview WorldCat COPAC

Scott, D. 2006: *Plato's* Meno. Cambridge: Cambridge University Press.

Segvic, H. 2000: No one errs willingly: The meaning of Socratic intellectualism, *Oxford Studies in Ancient Philosophy* 19: 1–45, reprinted in Segvic, H. 2009: *From Protagoras to Aristotle: Essays in Ancient Moral Philosophy*. Princeton, N.J.: Princeton University Press.
WorldCat

Vlastos, G. 1991: *Socrates: Ironist and Moral Philosopher*. Ithaca, N.Y.: Cornell University Press. 10.1017/CBO9780511518508
Google Scholar Google Preview WorldCat COPAC Crossref

Weiss, R. 2001: *Virtue in the Cave: Moral Inquiry in Plato's Meno*. Oxford: Oxford University Press. 10.1093/0195140761.001.0001
Google Scholar Google Preview WorldCat COPAC Crossref

Zeyl, D. J. 1986: *Gorgias*. Indianapolis: Hackett.
Google Scholar Google Preview WorldCat COPAC

Downloaded from https://academic.oup.com/book/4008/chapter/145635901 by University of Toronto user on 11 March 2024