

Tobias Hoffmann (ed.)
*Weakness of Will from Plato to the
Present* (Catholic University of America
Press 2008)

1 Weakness and Will in Plato's *Republic*

KENNETH DORTER

The central problem in determining Plato's attitude toward moral weakness lies in the apparent discrepancy between what Socrates says in dialogues like the *Protagoras*, *Laches*, *Charmides*, and *Meno*, on the one hand, and Book 4 of the *Republic*, on the other. The former are characterized by moral intellectualism, the view that since we always want good things for ourselves, once we know that something is good we will act in accordance with that knowledge; they argue that knowledge has an intrinsic power too great for it to be enslaved by inferior principles like appetite and spiritedness.¹ While the *Protagoras* scornfully ascribes to "the multitude" the view that knowledge alone is not strong enough to rule over its rivals—that without the help of something like moral strength it can be enslaved by our emotions (352b–c)—the *Republic* seems to defend precisely the view that the *Protagoras* dismissed as that of the multitude. According to Book 4, it is possible for us to know what is good and yet fail to act on that knowledge because we are too weak to master our temptations or fears, so that if our knowledge is not accompanied by moral strength or self-mastery, our rational faculty can indeed be overmastered by its inferiors:

Self-control [σωφροσύνη] is surely some kind of order; the self-mastery [ἐγκράτεια] of certain pleasures and appetites, as they say, using the phrase "master of oneself" [χρεῖτω αὐτοῦ]—I don't know how—and other such phrases that are like traces that it has left behind. . . . Yet isn't the expression "master of oneself" ridiculous? He who is master of himself would also be subject to himself, and he who is subject master. The same person is referred to in all these statements. . . . But the saying seems to me to want to say that in the same person there is something in

I would like to thank Tobias Hoffmann for his helpful comments on an earlier version of this paper.

1. Some passages say this not only about knowledge but also belief (e.g., *Protagoras* 358b–d, οἴόμενος οἴεται), which complicates the issue because it is more difficult to maintain that mere belief is too noble to be enslaved. We will consider the claim about belief after discussing the claim about knowledge.

the soul that is better and something that is worse, and when the part that is better by nature is master of the worse, this is what is meant by speaking of being master of oneself. . . . But when, on the other hand, because of bad upbringing or bad company the better part which is smaller is mastered by the multitude of the larger, we blame this as something shameful, and call it being subject to oneself and licentious. (430e-31b)²

In the soul of the tyrant, explicitly, "the most reasonable part of it is dishonorably and wretchedly enslaved" (577c).

It is tempting to resolve the inconsistency by supposing that Plato has simply changed his mind and come to recognize, like Aristotle,³ that we quite regularly act against our better judgment under the pressure of our appetites and fears. There are two reasons to be cautious about that solution, reasons that seem to point in opposite directions. On the one hand, the dialogues that assert moral intellectualism all end at an impasse; they fail to resolve the issue to which they are devoted, and leave us uncertain about precisely what we are meant to learn. Conceivably, we may be meant to learn that the inquiry foundered because the doctrine of moral intellectualism leads to a dead end. In that case the *Republic* would present positively what the other dialogues imply negatively: the need to accommodate the possibility of weakness of will.

On the other hand, the *Republic* presents its self-mastery thesis in a way that undermines our confidence in it as much as if it too had led to an impasse. Just prior to his introduction of the concepts of self-control and self-mastery, Socrates says, "Two things still remain to be discerned in our city, self-control and . . . justice. How then might we find justice without having to bother any more about self-control?" (430d). He seems to be suggesting that the most adequate conception of justice would not require reference to self-control (or self-mastery), but Glaucon replies that he does not want to skip over self-control in any case, so whatever lay behind Socrates' question remains unexplained. Socrates renews his misgivings more forcefully midway through the ensuing discussion: "in my opinion we will never get an accurate answer using our present methods of argument." But when he adds that "perhaps we can get an answer that's up to our previous standard," Glaucon is satisfied, so the source of Socrates' dissatisfaction once again fails to be explained (435c-d; cf. 504b). After

2. Translations are my own.

3. "It is problematic how someone with correct understanding can lack self-mastery. Some say this is not possible for someone who has knowledge; for it would be strange, as Socrates thought, if when someone possessed knowledge something else should master it and 'drag it around like a slave.' Socrates in fact used to attack the account altogether, on the grounds that there is no such thing as lack of self-mastery; for no one understands himself to act against what is best, but they do so only through ignorance. Now this account clearly goes against the evidence" (*Nicomachean Ethics* 7.2.1145b21-29).

these warnings it is not surprising that Socrates sums up their eventual conclusion with something less than enthusiasm: "if we said that we discovered the just person, and the courageous one . . . I think we would in some way not completely seem to be lying" (444a).

It seems, then, that if the intellectualistic claims of dialogues like the *Protagoras* are thrown into question by their aporetic conclusions, the self-mastery thesis of the *Republic* is made to look equally uncertain in other ways, and it is conceivable that the *Republic* may ultimately not be opposed to moral intellectualism after all.⁴ As I intend to show in this paper, no interpretation can do full justice to the *Republic's* position that does not accommodate Socrates' insistence that the treatment in Book 4 is inadequate, and the way that the conception of knowledge changes in the next three books, especially after the Divided Line. The conception of knowledge presented in Book 7 will be sufficient for virtue—no longer vulnerable to temptation and fear or in need of supplemental reinforcement—which means that the *Republic* is not ultimately at odds with the *Protagoras*; the discrepancy appears only if we take Book 4 to be the *Republic's* final word on the subject, instead of seeing it as a preliminary treatment that is superseded in what follows.⁵

The first quotation shows that the concept of self-mastery requires (1) a conception of the soul or self as divided, and (2) the identification of one part as better and another as worse. In what follows we shall consider, first, how Socrates establishes in Book 4 that the soul has three parts, and then the arguments in Book 9 for regarding rationality as in some sense the best of them (both claims are contentious, and their contentious nature will be pointed out by Socrates himself). Only then will we return to the argument for the tripartition of the soul to discover

4. Cf. Samuel Scolnicov, "Reason and Passion in the Platonic Soul," *Dionysus* 2 (1978): 35–49, at 45, and *Plato's Metaphysics of Education* (London: Routledge, 1988), 102 and 112; W. K. C. Guthrie, *A History of Greek Philosophy*, vol. 4 (Cambridge: Cambridge University Press, 1975), 435–36; Emile de Stryker, "The Unity of Knowledge and Love in Socrates' Conception of Virtue," *International Philosophical Quarterly* 6 (1966): 428–44, at 440–41; and Horace Fries, "Virtue Is Knowledge," *Philosophy of Science* 8 (1940): 89–99, at 91 and n. 5.

5. Explorations of this issue usually focus on Book 4. See, e.g., Christopher Bobonich, *Plato's Utopia Recast: His Later Ethics and Politics* (Oxford: Clarendon Press, 2002), 235–58; Gabriela Roxana Carone, "Akrasia in the *Republic*: Does Plato Change His Mind?" *Oxford Studies in Ancient Philosophy* 20 (2001): 107–48; Terence H. Irwin, *Plato's Ethics* (New York: Oxford University Press, 1995), 223–24; Amélie Oksenberg Rorty, "The Limits of Socratic Intellectualism: Did Socrates Teach Arete?" in *Proceedings of the Boston Area Colloquium in Ancient Philosophy*, vol. 2, ed. John Cleary (Lanham, Md.: University Press of America, 1987), 317–30, at 328–29; Lloyd P. Gerson, "Platonic Dualism," *Monist* 69 (1986): 352–69, at 359; Glenn Lesses, "Weakness, Reason, and the Divided Soul in Plato's *Republic*," *History of Philosophy Quarterly* 4 (1987): 147–61, at 148; and Charles Kahn, *Plato and the Socratic Dialogue: The Philosophical Use of a Literary Form* (Cambridge: Cambridge University Press, 1996), 240.

why Socrates disparaged it, and how it is superseded by the conception of knowledge introduced in Books 5–7. Because of the *Republic's* “arch” structure, it is in those central books that the discussion reaches the highest level, after which it explicitly returns to the discussion of Book 4 (543c). Why Plato proceeds in this way will be considered subsequently.

Since the concept of will does not appear directly in Plato (the term usually translated as “weakness of will,” ἀκράτεια as at 461a, simply means “weakness”), we will formulate the problem in terms of Socrates’ statement that moral weakness (“being subject to oneself”) means that the better part in us is overpowered by the worse. In that case, for our purposes Plato’s question in Book 9, “Which is our best part?,” can be taken to mean, “Which is our true will?”⁶

1. The Tripartite Soul

Books 2 and 3 develop Socrates’ model of the soul writ large, the depiction of the formation of a city as a visible analogue of the soul, to facilitate our understanding of the nature of justice (368c–69a). When the city evolved into three classes—productive workers, military auxiliary, rulers—the definitions of the virtues were based on the relationship of the classes to one another. Consequently the definitions will be transferable to the soul only if it has an analogous triadic structure. It is obvious that the distinctive characteristics of the three classes—appetite in the productive class, spiritedness in the auxiliaries, rationality in the rulers—are in each of us since we all enjoy pleasures, have a degree of ambition, and care whether things are true or false. It is also obvious that they sometimes lead to conflicting goals and desires; but that does not mean they must be discrete parts within us. It is doubtful, for example, that we ever experience rationality when it is not accompanied by some emotion, or emotion that is completely devoid of rationality, so our experience does not seem to support a model of the self in which these functions are anything like separate parts. Socrates raises the problem himself: it may be, he says, that our soul acts as a unity when it learns, gets angry, and desires, rather than doing each of these with a different part of itself (436a). My hand can reach for objects of pleasure, swing a tennis racket in competition, and turn pages in pursuit of truth, but that does not mean it has different parts for each of the three functions. Why should it mean that in the case of the soul? To eliminate that possibility

6. In the moral context of weakness of will, will is not necessarily distinguished from reason, as it is in the debate between voluntarism and intellectualism (i.e., whether reason or will predominates in us).

Socrates constructs an argument around the principle of opposition, an ontological equivalent of the principle of contradiction:

(1) The same thing cannot be opposed to itself in the same respect, in relation to the same thing, at the same time. So, if this happens in the soul we are not dealing with one thing but many (436b).

(2) Standing still while moving our arms and head is not an exception because we are not standing still and moving in the same respect: our feet are at rest and our arms are moving (c-d).

(3) The same is true of a spinning top that stays in one place: "it's standing still with respect to its axis . . . and moving in a circle with respect to its periphery" (d-e).

(4) Sometimes we are thirsty but decide not to drink, so there is something in us that tells us to drink, namely, appetite, and something else that tells us not to, namely, rationality, the two of which must then be different species in our soul (439c-e).

(5) Sometimes we get angry at ourselves because of our appetites, in which case our spiritedness fights against our appetite as one thing against another (439e-40d).

(6) When we rationally restrain our anger, then spiritedness and rationality are in opposition (441a-b).

We will consider the cogency of this argument in section 6, but for now we can say that it provides at least *prima facie* grounds for believing that appetite, spiritedness, and rationality are distinct parts of us. Our next step is to see how Socrates establishes that rationality is the best of the three, and must be obeyed for self-mastery to occur—in which case rationality is what would correspond to our true will and be the referent of weakness of will. That conclusion is by no means self-evident. Most people are governed by their appetite, the next largest group by spiritedness, and the smallest by rationality, so the concept of weakness of will is seen in different ways by different people. Appetitive or spirited people sometimes feel that if they listened to the warnings of reason and refrained from doing what their appetites or spiritedness prompted, it would show not strength of will but timidity. Thrasymachus has nothing but contempt for the "weak," "fearful" people who do not try to maximize their pleasures and power (344c), a view that Glaucon echoes as a devil's advocate (359a-60d), and that resurfaces again in Book 8 (549c-50a). Rational people may appear weak to the spirited, and unadventurous to the appetitive. Throughout the history of philosophy hedonists have championed appetite over rationality, and, especially since Schopenhauer and Nietzsche, voluntarists have championed something like spiritedness over rationality. Outside the philosophical

community there is even stronger disagreement with the claim that the pursuit of truth is preferable to the pursuit of pleasure or power. Here again Socrates himself calls attention to the problem: although to rational people rationality seems like what is best in us, to appetitive people appetite appears to be best and rationality is good only as an instrument of our appetites, while to spirited people spiritedness seems best and rationality is good only if it brings us honor (580c–81c).

To establish his claim that reason alone deserves to be our ruling faculty—our will—(and that therefore the just life, the rule of reason, is the best life), Socrates offers three arguments to the effect that only a life that follows reason will bring true and lasting happiness, while the others lead only to ambiguous and temporary satisfaction. In terms of the question of weakness of will, it makes sense to call reason our true will if it is the only drive that can bring us fulfillment.

2. First Argument: Fitness to Rule

Socrates' first argument compares the inner life of the most just and the most unjust person. The most just person is ruled by rationality, while the most unjust is ruled by the most extreme of the three forms of appetite. The three forms are distinguished by first dividing appetite into necessary and unnecessary appetites, and then unnecessary appetites into lawful and lawless species. Necessary appetites are those that are indispensable or beneficial, like the desire for healthy food. Unnecessary appetites are neither indispensable nor beneficial, and are "harmful both to the body and to the soul with respect to wisdom and self-control," like the desire for unhealthy food (558d–89c). Socrates' description of lawless appetites could have come straight out of Freud:

They are probably in all of us, but when restrained by the laws, and by the better desires together with rationality, in some people they are gotten rid of completely or only a few weak ones remain, while in others they are stronger and more numerous. . . . They are aroused during sleep when the rest of the soul slumbers—the rational, gentle, and ruling part of it—but the bestial and wild part, full of food or drink, springs up and, pushing away sleep, seeks to go and satisfy its dispositions. You know that in this condition it dares to do everything, as though released from and rid of all shame and wisdom. It shrinks from nothing, and tries to have sex with its mother, as it believes, or with any other person, god, or beast; nor does it shrink from any murder or refuse any food. In a word, it stops short of no folly or shamelessness. . . . There is then a terrible, wild, and lawless form of appetite in each of us, even in some of us who seem to be most moderate, and this becomes clear in our sleep. (571b–d, 572b)

Once we relax control over our appetites, things that were once unthinkable become at first conceivable and then irresistible, as we become driven by an appetite for the excitement and adventure of novelty. The belief that every appetite no matter how extreme deserves to be gratified cannot help but be felt eventually as an insatiable craving. But those who are in the grips of an insatiable craving cannot be happy; they are rather the most unhappy and least free kind of person, slaves to their most violent appetites. Socrates had already argued in Book 7 that the life of rationality leads to a vision of goodness that brings complete happiness (516b–19d). Now, by contrast, we find that the most appetitive life has the opposite effect and leads to the greatest unhappiness. The argument gives us one reason to believe that if weakness of will means that the worse part of us rules over the better, it is those ruled by appetite who most fit this description.

3. Second Argument: Criteria of Truth

The next argument could hardly begin more hesitantly: "Look at this second one if indeed you think there's anything in it" (580d). This is where Socrates points out the relativity we noted earlier, that appetitive, spirited, and rational people all love their defining pleasure—material things, honor, or wisdom, respectively—so if we ask them which of the three kinds of life is most pleasant they would each choose their own: lovers of material things will not care much about learning or honor except where they are materially profitable; lovers of honor have contempt for money, which they consider vulgar, and for learning, which they regard as smoke and nonsense (some things never change), except where these can bring them honor; and lovers of wisdom regard the other two not as true pleasures but only as necessities, which they would ignore if possible. How then can we decide which of the three is correct (580c–81c)?

It seems odd that Socrates would ask which way of life is most pleasant, since he was only challenged to show that just people are *happier* than unjust ones, and Socrates is the last person we would expect to equate happiness with pleasure, especially after he pointed out that some pleasures are bad (505c).⁷ It would be more in keeping with his preceding analyses to maintain that even if the life of a voluptuary is more pleasant than that of the philosopher, the philosopher's life is more noble, and

7. Cf. N. R. Murphy, *The Interpretation of Plato's Republic* (Oxford: Clarendon Press, 1951), 209–11, 223; Julia Annas, *An Introduction to Plato's Republic* (Oxford: Clarendon Press, 1981), 306–14.

nobility is closer to the good (and happiness) than is pleasure. Socrates insists, however, that this argument is not about nobility or goodness but only about pleasure and pain (581e–82a). He wants to show that even if he accepts the terms of the lovers of appetite his conclusion still follows. When people object against hedonism that they are not motivated by pleasure but by nobility or morality, hedonists reply that the moralists have not chosen morality instead of pleasure, but that being moral gives them the *most* pleasure—that people who say they prefer something else to pleasure are really only saying that they prefer a less common form of pleasure to a more common one. Accordingly, Socrates now classifies all three parts of the soul as species of appetite that pursue corresponding species of pleasure (580d), and seeks to show that even by this criterion the life of rationality will still turn out to be the best (as Epicurus later concluded as well).

Socrates suggests three criteria for determining the truth of the matter: experience, intelligence, and rationality (582a). In the main argument, which deals with experience (the other two are dealt with quickly afterward at 582d), Socrates asks which of the three types of people has the most experience of all three kinds of pleasure: appetitive, spirited, and rational. They conclude that while lovers of wisdom have ample experience of all three, lovers of material things and lovers of honor have never experienced the pleasure of contemplating reality (582a–d). Philosophers are the only ones in a position to compare the three on the basis of experience, and so their verdict has the greatest claim to truth.⁸

A hedonist might object to this that Socrates' three criteria—experience, intelligence, and rationality—all belong to the rational part of the soul, so even though the argument takes pleasure as its standard, its criteria are those of rationalism. But this would be problematic only if those criteria were arbitrary, whereas they are in fact the most reasonable ones that could be appealed to. Could the hedonist object to this reply on the same grounds—that it appeals to the criterion of reason-

8. Cf. John Stuart Mill's similar argument in *Utilitarianism*, chap. 2, par. 6–7. R. L. Nettleship believes the argument is unsatisfactory because "a man who had no experience of a kind of pleasure which he was asked to believe was better than his own could not be convinced by the experience of another"; see *Lectures on the Republic of Plato*, 2nd ed. (London: Macmillan, 1901), 321. However, the goal of the argument is not to convince appetitive and spirited people that their belief about what is most pleasant is wrong, but to provide evidence to show which of the three beliefs has a stronger claim to truth. Theodor Gomperz claims that the argument "overlooks the fact that greater susceptibility to one kind of pleasure is usually coupled with a smaller capacity for enjoying other kinds" (*Greek Thinkers: A History of Ancient Philosophy*, vol. 3 [London: John Murray, 1905], 100), but he offers no evidence that—assuming his generalization to be correct—the capacity itself is smaller rather than the pleasure seeming less important by comparison.

ableness? In the *Philebus*, when the life of reason and the life of pleasure compete against each other, and Socrates proposes to decide between the two claims rationally, Philebus replies that as far as he is concerned, pleasure will be the winner regardless of what happens in the argument (12a-b), and refuses to take any further part. It is an understandable course for him, for why should the competing claims of hedonism and rationalism be judged by the criteria of reason? If Philebus finds the pursuit of pleasure more pleasant than the pursuit of truth, then for him that is the only criterion that counts. There is no common ground on which Socrates can meet someone like that, and in the *Philebus* Plato shows that he understands this. Hence the caution with which Socrates introduced this argument: "See if you think there's anything in this." The most that can be said against Socrates' procedure here is that he has not demonstrated all of his premises, but it is never possible to demonstrate all premises, and we must be satisfied if we begin with what is self-evident in the sense that to deny it is to deny the possibility of rational inquiry. Socrates' assumption meets that standard.

If this argument presupposes rationality as the criterion, that cannot be said of the next one, which will proceed not from the character of knowledge but from the character of pleasure itself. The three arguments by which the rational life defeats the appetitive life correspond, in fact, to all three of the values manifested by the three parts of the soul and the city. The first argument was based on a civil war model wherein not only do the three parts of the city and the soul compete for supremacy, but a further power struggle takes place within the appetitive part, either between rich and poor classes or between necessary and unnecessary appetites. In the end the just life is preferred because in the unjust soul the better parts are enslaved by the worse. That argument relied on the categories of spiritedness—warfare, enslavement, and domination—while the present argument relies on the rational categories of experience, intelligence, and rationality. The third and final argument will rely on the categories of appetite: which life has the most pleasure and the least pain.

4. Third Argument: True and False Pleasure

Socrates distinguishes three conditions: pain, pleasure, and an intermediate calm. When people in pain call it pleasant to be free of pain, they are mistaking calm for pleasure: as they move from the bottom to the middle they mistake the middle for the top (583d). Since it is also true that when someone ceases to feel pleasure the state of calm feels

painful (the psychological basis of addiction), the intermediate state can be both painful and pleasant. But it is not "possible for what is neither to become both" (583e), so this state is not true pleasure (or true pain).

It is often objected that if people find this condition pleasant, it makes no sense to say they are not feeling pleasure: pleasure is a subjective state and only the individual can pronounce on its presence or absence.⁹ But the argument rests on more than telling people that they do not feel what they think they feel. Socrates' point is that because these perceived pleasures depend on a prior condition of pain, a life devoted to them is equally a life devoted to pain, and thus only an apparently pleasant life. Still, if one person is content because of pseudopleasures like relief from pain, and another is content because of real pleasures, both are nevertheless content. If ignorant people are happy because they do not know any better, are they any less happy than wise people?

The answer is implied in the next stage of the argument. Ignorance and being unwise are empty states of the soul, as hunger and thirst are empty states of the body—the latter filled by food and drink, and the former by knowledge and reason. We are more truly filled, Socrates says, when we are filled with what more truly is, namely, the form (εἶδος) of true opinion, knowledge, reason, and all of virtue, rather than the class of food, drink, and nourishment in general. The latter never *is*, but is in a constant state of becoming, while the former "is always the same, and immortal, and true" (585b–d). Consequently it fills us with something that is stable and pure, and thus with more reality than other pleasures do (586a). It is not enough to say that as long as we feel pleasure, the basis of that feeling is unimportant—for pleasure turns out to be not just a feeling but a state of being and a change in our condition; and we can scarcely say that changes in our condition make no difference as long as what we feel is pleasant. Guided by rationality, we will even be able to pursue the pleasures that the realms of appetite and spiritedness have to offer, and enjoy them in their truest forms (586d–e). Under the guidance of rationality, spiritedness and appetite will devote themselves to pursuits that are most likely to achieve the desired results, rather than if our spirited pursuits were reckless and self-defeating, or our appetitive pursuits sought pleasure in ways that also bring the most pain. Here again what Socrates is proposing is the conception of pleasure maximization through rationality that was later championed by Epicurus.

9. Cf. R. C. Cross and A. D. Woozley, *Plato's Republic: A Philosophical Commentary* (London: Macmillan, 1964), 266, and Guthrie, 541. But as Annas points out, "it is up to us to argue, and not just assert, that pleasantness can *never* be an objective matter"; see *An Introduction to Plato's Republic*, 314, and her discussion at 308–14 generally.

5. Problems with the Tripartite Soul

At this point the *Republic's* position seems clear. Weakness of will is possible because our soul is divided into three parts, one of which deserves to be identified as our will because it is what is best in us. When we ignore the best part in order to follow one of the inferior parts, we are weak-willed, or, in Plato's terms, subject to ourselves rather than masters of ourselves. Contrary to what most people believe, the best part is rationality, so that is our true will or best self.

But we have not yet dealt with Socrates' reservations about the first argument, the division of our soul into three parts. We must now try to understand the reasons for his warnings and see where that leaves us. Are there any obvious limitations with the tripartite model of the soul that might explain Socrates' repeated warnings? Most obvious was the assumption that if these are parts of the soul there are just three of them. The demonstration argued why there are not fewer than three, but gave no reason why there might not be more than three. In fact, a few pages later Socrates casually refers to the three parts of the soul together with "any others that may be in between them" (443e).¹⁰ It is not hard to imagine borderline cases between the frustration of unfulfilled appetite and the anger of spiritedness, or between the striving of spiritedness and the striving for truth. Is there any limit to the number of parts in the soul, or could we keep finding additional intermediate cases until the soul looks more like a continuum than a troika? Does it even matter whether or not there is a finite number of parts?

As the dialogue proceeds, not only does Socrates allude to intermediate parts between the original three, but each of the original parts is further subdivided. At the end of Book 6, in the Divided Line section, rationality is divided into four parts. Spiritedness too is subdivided, sometimes taking the form of anger, sometimes love of honor, and sometimes love of victory (436a, 581b, 586c-d). As for appetite, in Book 9 Socrates goes so far as to say that there is *no common quality* to all the things called appetite: "because of its multiple forms, we could not designate it with one appropriate name, but we named it by what was biggest and strongest in it" (580d-e). In the course of the dialogue the soul looks less and less like a triad, and more and more like a continuum that at one end pursues the most brutish pleasures, and at the other end the most transcendent truth.

Let us look more closely at steps 2 and 3 of the argument for tripar-

10. Similarly, the appetitive, spirited, and rational classes of the city that serves as a visible analogue of the soul are only "the primary three classes" (581c) because there can be any number of intermediate cases and subdivisions among them.

tion, the example of the man waving his arms and the example of a spinning top. Both were examples of apparent contradictions that were resolved by distinguishing different parts, in anticipation of the way different parts would be attributed to the soul. In the first case, although the man was both standing still and moving, it was his feet that were standing still and his arms and head that were moving. If we think of the soul on that model, then it is composed of parts that are absolutely distinct: if an arm, leg, or even the head were lost, the others would still remain. On that model appetite, spiritedness, and rationality are three entirely distinct parts of the soul, which is how the soul appeared when the model was first formulated. But the example of the top is different. If we removed the periphery of a top, whatever is left would still be the periphery—we cannot remove it but only narrow it. Similarly if we tried to remove the axis by drilling a hole through the center of the pin, the axis would remain but it would be hollow. Periphery and axis are not separable parts like arms and legs, but abstractions that stand to each other as relations. If we apply this model, the tripartite soul looks quite different. If appetite, spiritedness, and rationality are related to one another like the axis and periphery of a top, then they too are abstractions and relative to one another, more like locations on a continuum than discrete parts.

But if the parts of the soul are like locations on a continuum, how can we account for the oppositions that Socrates pointed to? In the discussion of step 4, Socrates illustrates the relationship between thirst and drinking with two kinds of examples of how one thing can stand in relation to another. In the first, greater is in relation to less, double to half, heavier to lighter, and so on (438b-c). In the second, knowledge of building stands in relation to building, while knowledge of medicine stands in relation to healing (438c-d). The second model obviously fits the context: thirst is the species of appetite that is related to drinking, as medicine is the species of knowledge that is related to healing. But the first model seems irrelevant. How can we say that thirst is related to drinking the way the double is related to the half? Either Socrates lost track of what he was talking about or the example serves a less obvious purpose. But he must have thought it was important because he gave us nine versions of it.¹¹ In fact it gives us a model of how things can be called opposites simply because they are different degrees on a continuum. The examples show that if the soul is more like a continuum than a triumvirate, we can still account for the oppositions that arise in

11. Greater to less, much greater to much less, formerly greater to formerly less, about-to-be greater to about-to-be less, more to fewer, double to half, heavier to lighter, faster to slower, hot to cold, "and all similar cases."

it. When rationality and appetite are in conflict, for example, we can say that the opposition is between, on the one hand, greater rationality and less appetite, and, on the other hand, greater appetite and less rationality, or between short-term gratification ("about-to-be less") and long-term gratification ("about-to-be greater").

At this point all these implications are merely implicit, but they become explicit later, although Socrates does not call attention to the way the model has changed. In another passage that could have been taken from Freud's discussion of the libido, Socrates says,

We surely know that when the appetites strongly incline toward some one thing they are thereby weakened toward others, like a stream from which there is a diversion into another channel. . . . So when they flow toward learning and all such things, they will be concerned, I suppose, with the pleasures of the soul itself by itself, while those that come through the body it will abandon, if someone is . . . a true philosopher. (485d-e)

Rationality was called the love of learning and was opposed to appetite, the love of pleasure, but now we are told that the love of learning is itself a kind of appetite. Here all three "parts" of the soul are regarded as channels through which eros (to use the word that Socrates employs in leading up to this quotation) flows toward different kinds of gratification (485b1, b8, c7). Later Socrates makes it even more explicit that rationality and spiritedness can be conceived as species of appetite, when he says that each part of the soul has its distinctive kind of appetite (580d), and he even calls philosophy (the love of wisdom) a kind of appetite (561c-d). If we conceive of the soul as a continuum extending between appetite at one end and rationality at the other, we can conceive of everything on it as a degree of appetite. But we can also conceive of them as degrees of rationality, as Socrates does when he says that all three parts of the soul are governed by their beliefs (*ὁμοδοξῶσι*, 442d1). Again, we can also conceive of the three parts of the soul as species of spiritedness, such as when Socrates portrays them as waging war against one another (440b, 440e, 442d, 444b) or "biting and fighting against each other" (589a). Even if we conceive of the soul as a continuum, then, appetite, spiritedness, and rationality may be regarded as its primary features.

Why, however, does Socrates take the trouble to establish a tripartite soul if he does not believe it to be an adequate model—why does he allow Glaucon to veto his misgivings and reluctance? For one thing, Plato often formulates a clear, simple model to provide an easily grasped answer to a question, but then introduces complicating factors that enable us to see beyond the limitations of the first model and to notice a sub-

tlar but more adequate model that begins to emerge.¹² Nowhere is that more true than in the *Republic*, which begins with an apparent victory by Socrates over Thrasymachus that is subsequently disparaged by Socrates himself as meaningless at the end of Book 1; then follows it with the elaborate development of a political constitution to illuminate the nature of the soul, at the climax of which Socrates again denigrates the accuracy of his methods. At that very point, however, ignored by Glaucon, he added, "although there is another longer and fuller road that does lead to such an answer" (435c). Although Glaucon showed no curiosity about this longer, more satisfactory way, a different kind of dissatisfaction by Polemarchus and Adeimantus (450a) forces Socrates to take that longer road and develop the material that will become Books 5-7. From that standpoint we will be able to see why Socrates progressively transformed the tripartite model of the soul into a continuous model.

At the end of the longer road, Socrates speaks of our ability to leave the cave of the visible world and perceive the ultimate reality of the principle of goodness that is the source of all being. In Book 7 he says that for those who have attained that vision, through what Socrates calls a conversion (518c, 521c), weakness of will is no longer a danger because it is a vision that cannot be attained by a divided soul:

[T]he instrument by which everyone understands is like an eye that cannot be turned to light from darkness except together with the whole body. Thus that instrument must be turned from the realm of becoming together with *the whole soul* until it becomes able to contemplate that which is, and what is brightest of that which is; and we say that this is the good. (518c-d, emphasis added)

If the whole soul must turn to it together, then there can be no discord within the soul, no weakness of will, and no need for self-mastery. If we can achieve this knowledge, the relative poverty of other kinds of goods is so clear to us that we can no longer be tempted by them. Those who have had this experience and recall the goals that most other people pursue would "go through any sufferings rather than share their opinions and live as they do" (516d). The gradual erosion of the distinctions between the parts of the soul, and the emergence of a conception of soul as a continuum, is the bridge between the simplified "inaccurate" model of Book 4 and the more elusive but more adequate conception of Book 7.

12. As Grace Hadley Billings puts it, "one of Plato's favorite methods of developing a theme [is that a] partial or superficial view of the subject is first presented, only to be superseded or supplemented by further discussion. The *Republic* [is among the] notable examples of this method"; see *The Art of Transition in Plato* (Menasha, Wis.: George Banta, 1920), 21 and n. 64.

6. Indefeasible Knowledge

If we want to understand the position of the *Republic* as a whole, and not only that of Book 4, on the question of whether knowledge is sufficient for virtue, we must first see what the *Republic* means by knowledge, and for that we must turn to the Divided Line. The Line draws an absolute distinction between knowledge and opinion, not only in the expected way that knowledge is certain and opinion uncertain (477e), but in the stronger and more surprising sense that knowledge corresponds to the intelligible realm of being (the forms) while opinion corresponds to the visible realm of becoming (509d–10a). We have no knowledge of the visible world, only opinion. If we apply that to the context of Book 4, we can see that what may have been called knowledge there can never be more than opinion. The entire project of the dialogue between Books 2 and 4 is to construct a visible analogue of the soul, a city, to serve as a model for our understanding of the soul. Throughout that stage of the argument, and again in Books 8 and 9, after Socrates returns from the longer road to pick up from where he had left off at the end of Book 4, the soul is treated as if it is nothing more than a smaller version of the city (368e–69a), in other words, as if it could be assimilated to visible models. Similarly, since in Plato divisibility is always the mark of visible objects,¹³ a tripartite conception of the soul in which the parts are fully distinct from one another would assimilate the soul to visible things. Thus at the end of Book 10, with the intelligible world brought into the picture in a way that had not yet happened in Book 4, Socrates says,

We must see the soul as it is in truth, not maimed by communion with the body and other evils, as we now see it, but discern it adequately with reason when it has become pure. . . . Now, however, we told the truth about the soul as it appears at present. . . . But we must, Glaucon, look elsewhere, . . . [namely,] to its love of wisdom, and we must understand what it touches upon, and what kind of things it yearns to associate with, as being akin to the divine, and to the immortal, and to what always is . . . And then one might see whether in its true nature it has many forms or one form, or in what way it is and how. (611c–12a)

Books 2–4 never introduce the theory of forms, but remain entirely within the visible world of becoming. Although the rational guardians are educated to philosophy, the love of wisdom (410e, 411e–12a), and are called wise (428e), their wisdom consists only of good judgment (εὐβουλος, 428b) and has nothing to do with a love of the intelligible realm of forms, which becomes the hallmark of the true philosopher in

13. E.g., *Phaedo* 78c.

Book 5 (475e–80a). Their education comprises only physical training and fine arts (376e). From the point of view of the Divided Line, then, what is called knowledge in Book 4 turns out to be, once the forms are introduced, only a kind of opinion. In Book 4 weakness of will is possible and self-mastery is necessary, because the rational element of our souls has not yet attained knowledge in the fullest sense of the word. True knowledge, once we attain it, can no longer be defeated, as we saw in the passages quoted above from 518c–d and 516d.

The need of the account in Book 4 to be reevaluated as the dialogue progresses is a function not only of Plato's practice of using simple models as starting points for more nuanced ones, but also a function of his practice of illustrating doctrines before they are explicitly formulated.¹⁴ The *Republic* does this in the case of the Divided Line: by the time the doctrine is actually stated, the dialogue itself has passed through each of its stages. The war of words with Thrasymachus in Book 1 was an example of "image thinking" (*εἰκασία*), the lower kind of opinion; the construction of a city as a visible model of the soul represented "conviction" (*πίστις*), the higher kind of opinion; and Books 5–7 illustrate the two kinds of knowledge, inferential (*διάνοια*) and intellectual (*νόησις*).¹⁵ If it seems misleading of Socrates to present a doctrine in Book 4 that is only meant to be superseded later on, at least we cannot say he did not warn us. To say that the account of Book 4 has been superseded by that of Book 7 is not to say that the former has been rejected. For the vast majority of us who have not achieved the true knowledge described in Book 7, the model presented in Book 4 is where we must begin.

When we look at Book 4 in isolation from what follows, it seems indeed as if Plato has abandoned the intellectualism of dialogues like the *Protagoras*. But when we take seriously Socrates' expressed dissatisfaction with the approach of Book 4, and follow his longer road to the present account, we find that true knowledge really is sufficient for virtue. It eliminates the need for self-mastery and the danger of weakness of will—but only if it is knowledge in the fullest sense.

The paradoxical nature of Socrates' intellectualism in the other dialogues, where virtue is defined in terms of knowledge, can be seen to turn on a similar kind of systematic ambiguity in the way the term "knowledge" is used there (although the ambiguity is not as compartmentalized as in the *Republic*), vacillating between wisdom and conventional kinds of

14. In the *Phaedo*, e.g., Simmias is made to anticipate the Method of Hypothesis at 85c–d.

15. I have provided evidence for this in "The Divided Line and the Structure of Plato's *Republic*," *History of Philosophy Quarterly* 21 (2004): 1–20, and throughout *The Transformation of Plato's Republic* (Lanham, Md.: Lexington Books, 2006).

knowledge. In the *Laches*, Nicias proposes that virtue follows from being wise (σοφός) and Socrates agrees, but when Laches asks what kind of wisdom (σοφία), Socrates restates the question as, "What is this knowledge [ἐπιστήμη]?" (194d–e), and the discussion ends in an impasse. The *Meno* hypothesizes that virtue can be taught if (and only if) it is a kind of knowledge (ἐπιστήμη) (87c), but the actual proof is a demonstration that virtue is a kind of wisdom (φρόνησις).¹⁶ The shift occurs at 88b and remains in effect until 89c, when "knowledge" returns, to be replaced again by "wisdom" in Socrates' later recapitulation: "Didn't virtue seem to us to be teachable if it was wisdom . . . and if it was teachable it would be wisdom?" (98d). However, in a passage reminiscent of the paradoxes at the end of the *Protagoras*, Socrates argues that since knowledge and teachability are equivalent, the fact that there are no teachers of virtue implies that virtue is not a kind of knowledge after all (89d and following). This passage, like that of the *Protagoras*, suggests that the difference between knowledge and wisdom in these dialogues is that wisdom is a species of knowledge that cannot be taught, at least not in the straightforward way that other kinds of knowledge can be taught. It is not information or technique. Knowledge is virtue, and invulnerable to weakness of will, only insofar as the kind of knowledge meant is wisdom. A comparable ambiguity is at work in the *Protagoras*. After reminding Protagoras of their agreement that knowledge is indefeasible (357c), Socrates smoothly extends the claim to belief: "no one who knows [εἰδώς] or believes [εἰόμενος] that other things are better than the ones he is doing, and are possible, then does these if the better ones are available" (358b–c). This extension bears on the aporetic ending of the dialogue which, like the *Laches* and the *Meno*, concerns the teachability of virtue. For if virtue is a kind of knowledge, and "knowledge" is broad enough to include "belief," then since beliefs are teachable, virtue (including wisdom) would have to be teachable.

In Book 4 of the *Republic*, too, knowledge is not clearly distinguished from belief. When Socrates defines self-control as when appetite and spiritedness "share the belief" (ὁμοδόξωσι) of reason about who should rule (442d), "belief" refers to what he had just called "knowledge" (ἐπιστήμη, 442c).¹⁷ The change of terminology implies that in cases of this kind, appetite and spiritedness cannot have true knowledge but only opinion, but when we turn to the kind of knowledge that is firmly

16. Σοφία is used at 91a and at 99b, where it is used interchangeably with ἐπιστήμη.

17. Carone shows that the rest of the tripartite soul discussion is consistent with 442d; see "Akrasia in the *Republic*," esp. 117–24. Space does not permit consideration of the interpretive difficulties of these texts, but I have discussed them in "Wisdom, Virtue, and Knowledge," *Review of Metaphysics* 51 (1997): 313–43.

distinguished from opinion, in Books 5–7, appetite and spiritedness no longer merely share the belief of rationality about who should rule, but also its knowledge of the good (518c). What the *Republic* and these other dialogues have in common is a systematically ambiguous employment of the term “knowledge,” so that in some places it seems to mean what we ordinarily call knowledge, but in others it refers to the perfection of knowledge as wisdom. The aporetic conclusions of the shorter dialogues invite us to discover what went wrong, and thereby to reflect on the difference between the knowledge that requires self-mastery and the wisdom that does not. The *Republic* does the same thing in a nonaporetic way, in terms of Socrates’ warnings in Book 4, and the higher point of view developed in Books 5–7.

7. Strengthening the Will

To combat weakness of will we must attain at least self-mastery of the kind described in Book 4, or, ideally, the conversion described in Book 7. Since both conditions are achievements of the rational part of our nature, the strategy for preventing weakness of will is the same in either case: we must make every effort to strengthen our rationality. One way of achieving this, good upbringing (409a), is beyond our control, but there are two things we can do: we can weaken whatever undermines our rationality, and we can strengthen rationality itself. Since rationality is by nature impartial, operating according to what is right rather than according to self-interest, we need to avoid anything that promotes biased behavior or emotions like self-indulgence, rage, and self-pity. In addition to not behaving that way ourselves, since behavior shapes character, we must also avoid associating too closely with people who do behave that way, because of the power of peer pressure (492b–c). We should even employ the same kind of restraint with respect to our entertainment. When we watch a play (or now a television program or movie) or read a book, we experience vicariously (“imitate within ourselves”) the behavior of the protagonists. If they behave unjustly or exhibit self-indulgence or self-pity, their behavior echoes within us and we imagine ourselves behaving that way. The more we imagine ourselves doing something that we disapprove of, the less unthinkable it is for us to behave that way ourselves. Even arts as abstract as music without words can convey emotions like self-pity and can make them dangerously beautiful and seductive. Plato gives examples in the music of his day (397c–99c), and there is no lack of examples in our own music. Examples are even more abundant in the case of music *with* words, from opera to all forms of popular music. These are the sort of influences we must be careful of, to avoid weakening our will by strengthening

the opponents of rationality. But how do we strengthen rationality itself?

The most obvious way is by making it more effectively rational—to be rational rather than to rationalize. Although rational behavior shows strength of will, while rationalization shows weakness of will, it is not easy to distinguish rationality from rationalization since rationalization means disguising as rationality what actually conflicts with rationality. It is the attempt to convince ourselves that we are behaving rationally (correctly) when we have ignored the promptings of reason in favor of those of appetite or spiritedness. It can be hard to tell whether we are rationalizing or being rational because we can easily mistake a lesser good for a greater one if the lesser good offers more immediate gratification—the way a small thing seems bigger than a large one if it is closer to us (*Protagoras* 356c–57b). We may choose the smaller rewards of appetite over the larger ones of rationality, or the smaller rewards of injustice over the larger ones of justice, because their rewards are more immediate. Moreover, the very impurity of the bodily pleasures is a source of their appeal: because they are mingled with pains, they are intensified by the contrast between the pleasure and the pain, which can make them madly exciting (586b–c). We must learn to recognize that pleasures that depend on relief from pain are less truly pleasant and will make us less happy than pleasures that are independent of pain, even though they are more intense.

In the case of pleasures that seem bigger only because they are closer, and provide immediate gratification instead of delayed gratification, we must develop the ability to measure goodness independently of its nearness or farness. The educational system of the *Republic* is designed to give us the ability to make this kind of measurement, but it is not a precise calculation like that of utilitarianism. At the end of the Myth of Er, Socrates says that someone's ability to "distinguish the good from the bad life" involves

knowing how beauty, combined with poverty or wealth and with what kind of character of the soul produces good or evil, good birth and bad birth, private life and governing, strength and weakness, ease of learning and difficulty of learning, and all such things regarding the soul, both natural and acquired, so that from all these things—and looking at the nature of the soul—he will be able to choose rationally between the better and worse life, calling a life worse which leads him to become more unjust, better if it leads him to become more just, and disregarding all other considerations. . . . He would know how to always choose the mean among such lives, and avoid each of the extremes. (618b–19b)

If the good is a mean between extremes, we cannot calculate it because the extremes do not have a precise value. Unlike a mathematical mean,

where we know the extremes first and then calculate a midpoint, here we must know the mean in order to know the extremes. We only know that something is an extreme because we recognize that it goes too far or not far enough, and we can recognize those cases only if we can identify the point that is just right.

As long as rationality needs to continually assert itself against appetite and spiritedness by means of self-mastery, and our emotions push us through rationalization toward the extremes, it will be difficult for us to recognize the mean. Only if we can attain the complete understanding of goodness that removes all temptation (516d, 518c-d) can we reliably discern the mean in all situations. Plato recommends two paths that can bring us to that understanding—one that operates at the level of appetite and another at the level of rationality. The first is eros and the second dialectic. There is also a path that operates at the level of spiritedness, namely, the self-discipline of purification that is present throughout the *Republic* in the mission of spiritedness to make appetite obedient to rationality, but the privative nature of discipline must be combined with a positive goal furnished by eros or dialectic. The path of eros is explored especially in the *Symposium* and the *Phaedrus* and is only briefly indicated in the *Republic*,¹⁸ which, like the *Meno*, the *Phaedo*, the *Timaeus*, and the *Philebus*, is more concerned with dialectic as an extension of mathematics.

The higher education of the rulers (523a-33a) begins with arithmetic because arithmetic begins in the realm of appetite but leads to that of rationality: its employment is in the visible world but its principles are intelligible. We cannot perceive numbers themselves—as distinct from their written symbols—with our eyes but only with our mind. If the study of mathematics makes us realize that the visible world is governed by principles that are perceived only by the mind, we are on the first step of the road to a dialectic that concerns itself only with intelligible being rather than visible becoming, and which leads to the source of the intelligible principles in the Idea of the Good. This “turning” of our attention from the physical world of becoming to the intelligible world of being leads to Socrates’ designation of it as a conversion (518c-d). The more we attend to intelligible reality, the more effectively will our own

18. E.g., 499b-c. For a discussion of the relation between the *Republic's* references to eros and Plato's doctrine of eros generally, see Irwin, *Plato's Ethics*, chap. 18. This positive sense of eros in the *Republic* should be kept in mind together with the negative sense that has been noted by Leo Strauss, *The City and Man* (Chicago: Rand McNally, 1964), 111; Stanley Rosen, “The Role of Eros in Plato's *Republic*,” *Review of Metaphysics* 18 (1965): 452-75; and Jacob Howland, “The *Republic's* Third Wave and the Paradox of Political Philosophy,” *Review of Metaphysics* 51 (1998): 633-57, at 646-55.

rationality function within us; and the more we appreciate the relative unimportance of physical things, the less power will our appetites have over us.

However convincing we may find Socrates' account, his words and our intellectual acceptance of them can do no more than give us the morally fragile kind of knowledge described in Book 4 that is not immune to weakness of will and must be enforced by self-mastery. The complete knowledge that turns the whole soul toward the good can only come from inner experience:

You will no longer be able, my dear Glaucon, to follow me, although for my part I would not willingly omit anything. But you would no longer see an image of what we are saying, but the truth itself. . . . And the power of dialectic alone can reveal it to someone who is experienced in the things we just went through, and it is not possible in any other way. (533a)

Few of us will ever get that far; for the rest of us weakness of will continues to be a danger, and so at the beginning of Book 8 Socrates turns away from the whole preceding discussion of Books 5–7 as a digression (543c), and returns to the divided self of the tripartite soul, proceeding to illustrate in detail how, if we do not take the necessary steps to achieve self-mastery, what starts out as weakness of will may become in the end the unexamined life of those in whom what is best in them is enslaved by what is worst.

Conclusion

We praise strength of will but not willfulness, because only "will" in the first sense implies that our goal is worthy. Thus the strength of will that Plato calls self-mastery implies that it is our better self that leads; but to demonstrate that one of our motivations is better than another is no easy matter. The discussion of self-mastery in Book 4 simply assumes that our better self is (to update the terminology) not the pleasure principle, nor the will to power, but the love of truth. Not until Book 9 does Socrates attempt to redeem that assumption with arguments. But even that is not the whole story. The *Republic* is a book of levels, and beyond the level at which strength of will is necessary to resist the temptations and fears that threaten to undermine our morality is a kind of knowing that is inviolable so that supplemental considerations of strength or mastery become irrelevant. It is important to remember that just as the discussion of Book 1 was declared inadequate (354b–c) and superseded by that of Book 4, that of Book 4 was disparaged as well (435c–d) and is superseded by that of Book 7.