

NO ONE ERRS WILLINGLY:  
THE MEANING OF SOCRATIC  
INTELLECTUALISM

HEDA SEGVIC

ἐκὼν ἐκὼν ἤμαρτον, οὐκ ἀρνήσομαι.

(Willingly, willingly I erred; I won't deny it.)

[AESCHYLUS], *Prometheus Bound*, 266

Video meliora proboque, deteriora sequor.

(I see what is better and approve of it, but pursue what is worse.)

OVID, *Metamorphoses*, 7. 20

Concepts, just like individuals, have their history and are no more able than they to resist the dominion of time, but in and through it all they nevertheless harbour a kind of homesickness for the place of their birth.

SØREN KIERKEGAARD, *The Concept of Irony*, 13. 106

I

THE Western philosophical tradition is deeply indebted to the figure of Socrates. The question 'How should one live?' has rightly been called 'the Socratic question'. Socrates' method of cross-examining his interlocutors has often been seen as a paradigmatic form of philosophical enquiry, and his own life as an epitome of the philosophical life. What philosophers and non-philosophers alike have often found disappointing in Socrates is his intellectualism. A prominent complaint about Socratic intellectualism has been memorably recorded by Alexander Nehamas: 'And George Grote both expressed the consensus of the ages and set the stage for modern

© Heda Segvic 2000

I am grateful to Myles Burnyeat, David Furley, John McDowell, and Julius Moravcsik for their helpful comments on earlier versions of this paper. I also wish to thank the editor of *Oxford Studies in Ancient Philosophy* for his generous criticisms and corrections.

attitudes toward Socrates when he attributed to him “the error . . . of dwelling exclusively on the intellectual conditions of human conduct, and omitting to give proper attention to the emotional and volitional”.<sup>1</sup>

The complaints against Socratic intellectualism take two main forms. According to some, Socrates ignores or overlooks—or at least vastly underestimates the importance of—the emotional, desiderative, and volitional sides of human nature, being too preoccupied with the intellect. The error attributed to him by Grote belongs here. The second line of criticism does not charge Socrates with ignoring or marginalizing desires, emotions, and volitions, but rather with giving an inadequate, over-intellectualist, account of them. These two lines of criticism have sometimes been combined, and sometimes confused. What they have in common is the thought that the desiderative, the emotional, and the volitional are not given their due by Socrates.

I wish to challenge this understanding of Socrates. He holds that living a good life is a matter of living in accordance with a certain kind of knowledge. Since knowledge is an accomplishment of reason, his view is in some sense intellectualist or, perhaps more appropriately, rationalist. However, I argue that desiderative, emotional, and volitional propensities and attitudes are an integral part of the knowledge in which Socrates locates virtue. This is meant to undermine the more prevalent first line of criticism. Towards the end of the paper I address the second line of criticism and suggest a different overall understanding of Socratic intellectualism, one that centres on the view that every act of the human soul involves an act of reason. I work my way towards this understanding of Socratic intellectualism by looking into the role that volitions, emotions, and desires play in Socratic virtue.

A large part of this paper deals with two Socratic theses. The first, that no one errs willingly, has long been recognized as crucial to Socratic intellectualism; however, the precise meaning of this thesis has remained elusive. I argue that ‘willingly’ is used here in a highly specific sense. The text which in my view offers the clue to the proper understanding of the No One Errs Willingly thesis is a passage in the *Gorgias* that has been much slandered in the literature

<sup>1</sup> Alexander Nehamas, ‘Socratic Intellectualism’, in his *Virtues of Authenticity* (Princeton, 1999), 24–58 at 27; the reference is to George Grote, *Plato and the Other Companions of Sokrates*, i (London, 1865), 399–400.

on Socrates. The argument has often been thought confused, and the whole passage has sometimes been treated as a deliberate exaggeration on Socrates' part. I claim that the passage makes perfect sense, that Socrates intends it seriously, and that it plays a central role in the overall philosophical structure of the dialogue. I then turn to the second thesis, that *akrasia*—weakness of the will, as the Greek term is usually rendered—does not exist. I offer an interpretation of the denial of *akrasia* based on my analysis of the No One Errs Willingly thesis. The joint reading of the two theses leads to a perhaps surprising result. Certain kinds of wantings and volitional propensities are constituents of moral knowledge. The same can be shown for desiderative and emotional attitudes and propensities. Far from disregarding the volitional, desiderative, and emotional, Socrates attempts to build them into his account of virtue as knowledge. Furthermore, his remarks on wanting or willing, sketchy and conversational though they are, point—I argue—to a distinct notion of the will. If Socrates does have a concept of the will, this is the first appearance of such a concept in the Western philosophical tradition.<sup>2</sup>

This interpretation shows that it is wrong to assume (as people have done since Aristotle) that Socrates ignores or marginalizes the desiderative and the emotional side of human nature, focusing solely on the intellectual.

## II

Socrates claims that no one errs knowingly.<sup>3</sup> Why an intellectualist would make such a claim, we might think, is not so difficult to grasp. The intellectualist believes that when a person does what is

<sup>2</sup> We should not conclude from the fact that ancients discuss *akrasia*, which we label 'weakness of the will', that they have a concept of the will. The term *akrasia* indicates only some kind of weakness: the weakness of one who acts against his knowledge or better judgement of what is best. It is not uncommon to find the literature associating a notion of the will with this or that ancient figure, including Socrates. But it is not by dint of translation that we should come to think of the ancients as having a concept, or concepts, of the will, but by dint of interpretation and argument.

<sup>3</sup> See *Prot.* 352 c 2–7: 'Now, do you [Protagoras] too think that that is how things stand with it [sc. knowledge], or do you think that knowledge is fine and such as to rule the person, and if someone recognizes what is good and bad, he would not be overpowered by anything else so as to act otherwise than knowledge dictates, but wisdom is sufficient to help the person?' Protagoras promptly grants that knowledge has this power. See also 358 B 6–c 1.

morally wrong, that moral failure is due to an intellectual error. If only the person exercised his intellect well—if he knew better—he would not do what is wrong. Hence what we have to do in order to make people better, an intellectualist would have us think, is help them see how things really are; in particular, help them see what really is good or bad. I do not dispute that Socrates is a rationalist or intellectualist of some sort, or that a line of thought roughly corresponding to the one just sketched may be linked to his claim that no one errs knowingly. What I wish to emphasize is that in order to determine what *kind* of intellectualist he is, we must see how he conceives of the knowledge the absence of which he takes to be responsible for wrongdoing. I shall argue that Socrates' conception of moral knowledge makes many of the objections traditionally lodged against his intellectualism unwarranted.

In addition to claiming that no one errs knowingly, Socrates also claims that no one errs *willingly*. Why does he make this latter claim? An answer to this question does not leap to one's eye from the pages of Plato's dialogues. One would expect that, if anywhere, an answer is to be found in the *Protagoras*, where Socrates argues at length for the view that *akrasia* does not exist, and where he also briefly formulates, and appears to endorse, the claim that no one errs willingly (*Prot.* 345 C 4–E 6; cf. 352 A 1–358 D 4). But the *Protagoras* is silent on what precisely the dictum 'No one errs willingly' amounts to and how it is related to Socrates' denial of *akrasia*. In view of this silence, it is tempting to think that Socrates himself was in error. He must have thought, mistakenly, that 'No one errs knowingly' implies 'No one errs willingly'. Those who recall Aristotle's discussion of voluntary and involuntary action in the *Nicomachean* and *Eudemian Ethics* may be especially inclined to think that Socrates simply made an error in passing from 'knowingly' to 'willingly'.

Aristotle was the first Greek philosopher, as far as we know, to undertake a systematic analysis of voluntary and involuntary action, and to connect the voluntariness and involuntariness of actions with the agent's knowledge or ignorance. He tried to specify as precisely as he could the kinds of ignorance concerning the circumstances of an action that make it involuntary (see especially *NE* 3. 1 and 3. 4). He stressed that not every kind has this effect: for some sorts of ignorance people are neither forgiven nor pitied—as might be appropriate if their action were due to ignorance. Instead, they are blamed (*NE* 3. 1, 1110<sup>b</sup>28–1111<sup>a</sup>2). In Plato, however, we find no

comparable attempt at a careful philosophical analysis of voluntariness and involuntariness. So it would be plausible to think that Socrates perceived that knowledge of some kind is connected with voluntariness, but never looked into the thorny issue of voluntariness with proper care. That allowed him to overlook the blunder involved in passing from ‘No one errs knowingly’ to ‘No one errs willingly’.<sup>4</sup>

Tempting as this line of thought might be, we should resist it. We should not assume without examination that when Socrates describes someone as acting willingly, the action in question would be of the sort Aristotle classifies as ‘voluntary’. (Likewise, we must not assume that those who on Socrates’ diagnosis act unwillingly are not to be blamed for their actions.) The intended meaning of ‘No one errs willingly’ should be gleaned in the first place through careful reading of Plato’s dialogues. The relevant passages seem to me to reveal that Socrates was not the least bit confused when he said that no one errs willingly.<sup>5</sup> Rather, I shall argue, he proposed a coherent and interesting, albeit unusual, view.

### III

In Plato’s *Protagoras* Socrates introduces the thesis that no one errs willingly (at 345 C 4–E 6) while presenting an analysis of a poem by Simonides. That no human being errs willingly is something, Socrates contends, that Simonides as a wise and educated person would surely have known. He proceeds to use this thought to guide his interpretation of Simonides, but he offers no gloss on the thesis itself. Although the *Protagoras* provides us with indispensable material for understanding Socrates’ ethical outlook, and hence also for understanding the No One Errs Willingly thesis, a more direct clue to the meaning of this thesis comes from the *Gorgias*.

Our starting-point should be *Gorgias* 466 A 4–468 E 2. In his exchange with Polus Socrates declares that orators and tyrants do

<sup>4</sup> John McDowell takes this view in his unpublished piece ‘Irwin’s Socrates and an Alternative Reading’. The culprit, however, is ultimately Aristotle. See next note.

<sup>5</sup> It is not just our knowledge of the philosophical analysis of voluntary action provided later by Aristotle that might mislead us into thinking that Socrates was confused. The picture of him as confused about voluntariness probably originated with Aristotle himself. Evidence suggests that Aristotle saw Socrates as mistaken on two issues: first, the role of knowledge and ignorance in voluntary and involuntary action (see, in particular, *NE* 3. 1 on τὸ ἐκούσιον and τὸ ἀκούσιον), and second, the issue of the proper object of βούλησις—rational wish or wanting (*NE* 3. 4).

not do what they want to do (467 B 2, 466 D 8–E 1), and that they have the least power of any in the city. Startled by this, Polus asks if it is not the case that orators, just like tyrants, kill anyone they want (*ὄν ἂν βούλωνται*), and subject anyone they please (*ὄν ἂν δοκῆ αὐτοῖς*) to expropriation or exile (466 B 11–C 2). Socrates retorts that Polus has raised two questions rather than one (466 C 7, 466 D 5–6), and proceeds to draw a distinction between doing what one pleases, on the one hand, and doing what one wants, on the other (466 D 5–E 2). Applying this distinction, he now grants that orators and tyrants do ‘what they please’ (*ἃ δοκεῖ αὐτοῖς*, at 467 A 3 and 467 B 8) or ‘what they take to be best’ (*ἃ δοκεῖ αὐτοῖς βέλτιστα εἶναι*, at 467 B 3–4),<sup>6</sup> but denies that they do what they want to do (*ἃ βούλωνται*, 467 B 2, B 6, 467 A 10; cf. 466 D 8–E 1)—presumably when engaged in the actions mentioned: killing, expropriating, banishing. The passage makes it fairly clear why Socrates claims that orators and tyrants do not do what they want to do: what they do is not good, and one can only want those things that are good (see especially 468 C 2–7). But why should he construe ‘wanting’ in such a peculiar way? To answer this question, we should take a broader look at the matters discussed at 466–8.

Socrates’ claim that neither orators nor tyrants do what they want to do is meant to be startling. What in common opinion distinguishes a tyrant from others is precisely the enormous power he has. As Polus had observed at 466 B 11–C 2, the tyrant can put to death anyone he wants; he can dispossess or exile whomever he pleases. Thus he can visit what in common opinion are the worst of evils upon the head of anyone he wants. Another bit of common lore is that having power consists in being able to do what one wants. Power is so understood by Socrates’ interlocutors in the *Gorgias*, and Socrates raises no objection. What Gorgias and Polus add to the common view is the claim that orators are at least as powerful as tyrants, and probably more so (see especially 452 E 1–8). This, of course, is advertising on behalf of oratory by its practitioners or

<sup>6</sup> See also the variants: *ὄτι ἂν αὐτοῖς δόξῃ βέλτιστον εἶναι* (466 E 1–2) and *ἃ ἂν δοκῆ αὐτῷ βέλτιστα εἶναι* (466 E 9–10). The two expressions ‘what they please’ (*ἃ δοκεῖ αὐτοῖς*) and ‘what they take to be best’ (*ἃ δοκεῖ αὐτοῖς βέλτιστα εἶναι*) are treated by Socrates as equivalent throughout the passage under consideration (466 A 4–468 E 2). A reader with no Greek will observe that the two expressions are rendered quite differently in English, but actually they look very similar in Greek. To capture the similarity, one could translate respectively ‘what seems to them’ (meaning roughly: as they see fit, or as they please) and ‘what seems to them to be the best’.

sympathizers. The advertisement none the less correctly identifies some of the aspirations, and some of the accomplishments, of oratory in the ancient world. Faced with Gorgias' and Polus' claims on behalf of oratory, Socrates does not take the obvious course, to reject as an exaggeration the claim that orators are so powerful. Rather, he takes the entirely non-obvious course of saying, first, that neither orators nor tyrants do what they want to do when they engage in the actions mentioned, and second, that they consequently have no great power in the cities. In making the transition from the first claim to the second, he relies on the above-mentioned assumption about power: to have power is to be able to do what one wants to do; to have a lot of power is to be able to do much of what one wants to do.

There can be no doubt that Socrates wants to shock his interlocutor by his apparently bizarre claim about orators and tyrants. Polus reacts as intended: he describes the claim as 'outrageous' and 'monstrous' (σχέτλια, ὑπερφυῆ, at 467 B 10). It would be a serious error, however, for us to understand the claim as a piece of histrionics, or an exaggeration meant to bring into sharper relief some other views that Socrates seriously holds.<sup>7</sup> He means what he says: orators and tyrants do not do what they want to do. If Polus is shocked by this claim, the shock is meant to prepare him for a more general claim which Socrates wants to be taken quite as seriously.<sup>8</sup>

That doing what one pleases or what one sees fit (ἃ δοκεῖ αὐτῷ) amounts to acting in accordance with one's opinion (δόξα) is suggested in Greek by the very form of the words (δοκεῖν is a verbal counterpart to the noun δόξα). This suggestion is further supported by *Gorg.* 469 C 4–7. There Polus explains to Socrates who, on his understanding, a tyrant is. A tyrant, he says, is someone who is 'in

<sup>7</sup> Pace Roslyn Weiss, in 'Killing, Confiscating, and Banishing at *Gorgias* 466–468', *Ancient Philosophy*, 12 (1992), 299–315. Her contention that the argument of 466–8 'deliberately . . . exaggerates and distorts' views that the Socrates of Plato's early dialogues 'seriously holds' (p. 299) strikes me as a counsel of despair in the face of the fact that the argument has persistently resisted coherent and plausible interpretation.

<sup>8</sup> The way Socrates proceeds here is not unusual. Something similar goes on during his interpretation of Simonides' poem in the *Protagoras* (338 E 6–347 A 5). In the course of making a peculiar sort of display, he introduces views he seriously holds, including the No One Errs Willingly thesis. A further similarity between his exercise in literary criticism in the *Protagoras* and his handling of Polus in the section of the *Gorgias* we are discussing is that he seriously proposes his thesis while being mockingly playful. Later in the *Gorgias*, as I shall point out below, he subjects Callicles to similar treatment.

a position to do whatever he pleases [ $\delta\ \alpha\ \nu\ \delta\omicron\kappa\eta\ \alpha\upsilon\tau\omega$ ] in the city, whether it is killing a person or expelling him from the city, and doing everything [ $\pi\acute{\alpha}\nu\tau\alpha\ \pi\rho\acute{\alpha}\tau\tau\omicron\nu\tau\iota$ ] in accordance with his opinion [ $\kappa\alpha\tau\grave{\alpha}\ \tau\eta\ \nu\ \alpha\upsilon\tau\omicron\upsilon\ \delta\acute{\omicron}\xi\alpha\nu$ ]. Polus here treats doing ‘whatever he pleases’ and doing everything ‘in accordance with his opinion’ as equivalent. The phrase quoted,  $\pi\acute{\alpha}\nu\tau\alpha\ \pi\rho\acute{\alpha}\tau\tau\omicron\nu\tau\iota\ \kappa\alpha\tau\grave{\alpha}\ \tau\eta\ \nu\ \alpha\upsilon\tau\omicron\upsilon\ \delta\acute{\omicron}\xi\alpha\nu$ , which I have rendered ‘doing everything in accordance with his opinion’, could equally well have been rendered ‘doing everything as he pleases’.<sup>9</sup>

If doing what one pleases amounts to acting in accordance with one’s *doxa*, opinion or belief,<sup>10</sup> and there is, Socrates suggests, a sharp contrast between doing what one pleases and doing what one wants, it is not unreasonable to suppose that doing what one wants is linked with acting in accordance with one’s *epistēmē*, knowledge. I shall defend the view that this is indeed so. In fact, I shall propose that wanting, as understood by Socrates in the present context, is even more intimately connected with knowledge than the phrase ‘acting in accordance with knowledge’ might suggest. Before I do so, let me make some remarks about the appropriateness of bringing knowledge into the picture.

The contrast between *doxa*, opinion, and *epistēmē*, knowledge, is at the heart of the *Gorgias* as a whole. Socrates recoils from oratory, which he considers dangerous to the human soul. Oratory is dangerous because it enshrines mere *doxa*, opinion, and aims to convert it into *πίστις*, conviction, without regard for the truth of the opinion, hence *a fortiori* without regard for knowledge. *πίστις*, conviction, is what persuasion (*πειθώ*), if successful, leads to, and producing persuasion is the business of the orator. Following Gorgias’ descriptions, Socrates characterizes the orator as a *πειθοῦς δημιουργός*, ‘a manufacturer of persuasion’ (*Gorg.* 453 A 2). Socrates sees himself, by contrast, as concerned with knowledge, hence he keeps denouncing practices that systematically bypass this concern.

<sup>9</sup> In Greek, the difference between acting  $\kappa\alpha\tau\grave{\alpha}\ \tau\eta\ \nu\ \alpha\upsilon\tau\omicron\upsilon\ \delta\acute{\omicron}\xi\alpha\nu$  and doing  $\alpha\ \nu\ \delta\omicron\kappa\epsilon\iota\ \alpha\upsilon\tau\omega$  lies merely in choosing between a noun-based idiom and a verb-based one. The difference can be illustrated in English by a choice between, say, acting ‘as one wishes’ and acting ‘in accordance with one’s wish’.

<sup>10</sup> I use ‘opinion’ and ‘belief’ interchangeably. ‘Opinion’—a more common rendition of *δόξα* in Plato—may be too narrow for the passages of the *Protagoras* and *Gorgias* under consideration here. Roughly, one has a *δόξα* when one takes something to be the case, correctly or incorrectly. This corresponds to ‘belief’ fairly well, as well as to ‘opinion’ loosely understood.



The orator and the tyrant, each in his own way, stand accused by Socrates of being mired in such practices.

To say that doing as one pleases is to be understood as acting in accordance with one's opinion or belief invites the question: an opinion or belief about what? Likewise for acting in accordance with one's knowledge. As far as opinion or belief is concerned, the very fact that Socrates treats ἃ δοκεῖ αὐτοῖς, what pleases them (467 A 3, B 8), as interchangeable with ἃ δοκεῖ αὐτοῖς βέλτιστα εἶναι, what they think (believe, opine) is best (467 B 3–4), suggests an answer. The opinion is about what is best, or perhaps more generally about what is good, better, or best. Although I think that we can take our cue from the expressions Socrates uses, I do not mean to suggest that his understanding of these matters is determined by the peculiarities of certain Greek idioms. Socrates has philosophical reasons for seeing the matter this way—reasons which will emerge as we proceed. These reasons stand behind the form of words he uses.

My suggestion was that Socrates describes orators and tyrants as not doing what they want to do because in doing what they do they do not act in accordance with knowledge. But what does wanting have to do with knowledge? Why should only those who have knowledge, or perhaps those who have the relevant knowledge, be correctly described as doing what they *want* to do?

I propose the following, preliminary, characterization of the notion of wanting which Socrates relies on in the orators-and-tyrants passage: the agent wants to  $\phi$  just in case he desires to  $\phi$  taking  $\phi$ -ing to be the good or right thing to do (in the circumstances in question), and his  $\phi$ -ing (in those circumstances) is (or would be) good or right in the way he takes it to be. The point of glossing 'good' as 'right' is that wanting to do something, as wanting is understood here, does not merely involve a desire to  $\phi$  because  $\phi$ -ing is seen by the agent as having some goodness in it; the agent wants to  $\phi$  only if he desires to  $\phi$  seeing it as the right or correct thing to do.

Now this sort of wanting, which I shall call *Socratic wanting* or *willing*, is presumably still a desiderative state of some sort, in a broad sense of the word 'desiderative'. How can the ascription of a desiderative state to an agent possibly depend on the object of the desiderative state being in fact good? Whether an agent wants something, wishes for it, longs for it, and so on, depends on how he sees, or conceives of, the object of his wanting, wishing, or longing. Must we not leave open the possibility that the agent is wrong in

his conception of the object desired, whatever the modality of his desire?

That, I take it, is how many people think of desiderative states; clearly, it is how Polus thinks of them. Socrates, however, is putting forward a different proposal. The issue here is not whether, generally speaking, one can be mistaken about the object of one's desire. Of course Socrates would agree that one can be. The issue is whether every kind of desire or volition that can be ascribed to a person is independent of the correctness of the person's conception of the object desired or wanted. A parallel may be of help here.

In claiming that orators and tyrants do not do what they want to do, Socrates is inviting us to think of *wanting* as a volitional state that is in some ways like perceiving. I do not perceive an object if I have some images; I perceive it only if my sensory impressions derive from the object itself in the right kind of way. Socratic volition is likewise a receptivity of the soul to certain evaluative properties of the object of volition, the properties Socrates designates by the term 'good'. However, wanting is not sheer receptivity; it is mediated by a correct conception of the object of desire as the good or the right thing to do. Just as perception latches on to that aspect of reality that has an impact on our sensory apparatus, so Socratic volition latches on to a certain evaluative aspect of reality. Thus this kind of wanting can be correctly ascribed to the agent only if the object of his volition has the required evaluative properties and the agent recognizes, and responds to, these properties. We should call to mind again the relationship between belief and knowledge. Whereas having a belief consists in taking something to be true,<sup>11</sup> knowing on Socrates' view is the secure grasp of truth. Likewise, he seems to suggest, whereas desire involves believing that the object of desire is good,<sup>12</sup> wanting—the sort of wanting referred to in the *Gorgias* passage—implies knowing that the object of volition is good.

I can now offer a more precise characterization of Socratic wanting: I Socratically want to  $\phi$  just in case I want<sup>13</sup> to  $\phi$ , recognizing

<sup>11</sup> If we want to be fastidious, we can say that believing is taking something to be the case, which implies that something—some proposition or statement—is true.

<sup>12</sup> See *Meno* 76 B 6–78 B 2. I shall come to this passage below.

<sup>13</sup> In this occurrence, 'wanting' should be taken in its generic sense, not implying a correct conception of the goodness of the object of the want. I take bitter medicines because I want to be healthy; I try to preserve my health because I want to live well, and so on.

that my  $\phi$ -ing (in the given circumstances) is the good or right thing to do.<sup>14</sup> Thus I (Socratically) want to  $\phi$  only if my wanting to  $\phi$  is linked to my recognition of the goodness of  $\phi$ -ing; if it is a mere coincidence that I believe that  $\phi$ -ing is the right thing to do and that  $\phi$ -ing in fact is the right thing to do, my wanting to  $\phi$  is not Socratic wanting.

This characterization is meant to bring Socrates' notion closer to us, while staying reasonably close to his own idiom. Its drawback is that it unravels a unitary notion: Socratic wanting is meant to be, I think, *both* a volitional *and* a cognitive state. On the best reading, the wanting would be a volitional state *in virtue of* being a certain kind of cognitive state. Socrates has philosophical reasons for offering us this notion of wanting. Before turning to them, let me make a few remarks in defence of my interpretation of the orators-and-tyrants passage.

#### IV

I have already pointed to one line of thought that makes it difficult to understand why orators and tyrants do not do what they want to do. This is the idea that the claim is a deliberate exaggeration or a piece of histrionics. Another, more widely shared, line of thought is the following. To understand the orators-and-tyrants passage one first has to settle the question whether Socrates uses the verb 'to want' in a special sense. For, if he does not use it in a special sense, then it appears that his claim cannot possibly be true; but if he does use it in a special sense, then he and Polus are not speaking of the same thing; hence his disagreement with Polus, or with anyone who shares Polus' point of view, is not genuine.<sup>15</sup> The prevalent interpretation

<sup>14</sup> Compare with this Socrates' formulation of what people take *akrasia* to be: the many [who believe that there is such a thing as *akrasia*] say that 'a lot of people, recognizing what is best [*γινώσκοντας τὰ βέλτιστα*], do not want to do it [*οὐκ ἐθέλειν πράττειν*], when it is possible for them to do so, but do something else instead' (*Prot.* 352 D 6–7). The relevance of this comparison, which connects Socratic wanting to his denial of *akrasia*, will become clear below.

<sup>15</sup> Terry Penner's interpretation of the passage is driven by an attempt to avoid the second horn of the dilemma. He consequently aims to preserve the ordinary sense of 'wishing' (his rendition of *βούλεσθαι*). According to Penner, Socrates' position is this: orators and tyrants (like everyone else) do what they want to do only if they get what they want. Whatever they may think they want, it is their real happiness that they want in everything they do. The only action one ever wants to do (or desires to do: there is no difference, on this view, between desiring and wanting or wishing to do something) is the one that in fact leads to the ultimate end, one's own

of the passage seems to be that Socrates does introduce a special sense of ‘wanting’ in the passage under consideration, but that for this very reason his overall argument is marred by equivocation, and hence flawed.<sup>16</sup>

happiness, through the chains of means and ends that one has correctly envisaged (Terry Penner, ‘Desire and Power in Socrates: The Argument of *Gorgias* 466 A–468 E that Orators and Tyrants Have No Power in the City’, *Apeiron*, 24/3 (Sept. 1991), 147–202; see esp. pp. 170, 182–97). To use Penner’s own example, if the tyrant’s killing of his prime minister does not lead to the tyrant’s happiness, in the way envisaged by him, we have to conclude that he did not want to kill his prime minister. I cannot offer here a detailed analysis of Penner’s rather intricate interpretation. Anticipating the analysis I am about to provide, for Socrates there is indeed a legitimate sense in which we want to do things if the doing of them is good and otherwise we do not want them. This wanting is conditional not, as in Penner, upon what one *gets* from the action in the future, but upon the goodness of what is wanted through the action. The virtue of Penner’s interpretation (as of his previous work on Socrates) is that he takes Socrates at his word, refusing to settle for ‘charitable’ readings of his claims. Thus I think that Penner is right in insisting that for Socrates the object of wanting (in one sense) is what really is good, rather than what one believes to be good. However, Penner in this article attributes to Socrates—and apparently himself subscribes to—an implausible general theory of desire (what we want or desire when we desire to do anything is, without exception, the whole chain of means leading to one’s actual happiness, and if we don’t obtain happiness by means of an action, then we did not want to do what we did), and an unattractive theory of action (no action is ever undertaken for its own sake). As I am about to argue, the dilemma which motivates Penner’s interpretation is not, as he believes, inescapable. In the appendix below I offer an analysis of Socrates’ argument at 467 C 5–468 E 5, and show that an instrumentalist account of action cannot be the correct interpretation of this passage.

<sup>16</sup> For a statement of this view see Robin Waterfield’s note on *Gorg.* 468 D: ‘The problem with the argument is that “want” is ambiguous, in a subtle way. To use the familiar philosophical example, Oedipus wants to marry Jocasta, but he does not want to marry his own mother; one can want and not want the same thing under different descriptions’ (Plato’s *Gorgias*, translated with explanatory notes by R. Waterfield (Oxford, 1994), 142). Further down, Waterfield accuses Socrates of trading on the ambiguity between ‘good’ and ‘apparent good’. (‘Just as importantly, Socrates’ argument has not really dented Polus’ position because of the ambiguity within “good”, *ibid.* 143; for this, compare Aristotle’s discussion at *NE* 3. 4 on whether the good or the apparent good is the proper object of βούλησις.) Terence Irwin similarly claims that Socrates’ question, ‘Does A do what he wants?’, is misleading, since the answer may be Yes when the action is considered under one description the agent believes true of it, and No when it is considered under another description. Consequently, Irwin takes Socrates’ conclusion, that the orator and the tyrant lack power, as ‘unjustifiably strong’ (Irwin’s notes to his translation of the *Gorgias* (Oxford, 1979), 145–6). According to Kevin McTighe, Socrates confuses *de dicto* and *de re* analyses of the verb ‘to want’—see McTighe, ‘Socrates on the Desire for the Good and the Involuntariness of Wrongdoing’, *Phronesis*, 29 (1984), 193–236; repr. in Hugh H. Benson (ed.), *Essays on the Philosophy of Socrates* (Oxford, 1992), 263–97. McTighe provides a useful survey of the received interpretations of the orators-and-tyrants passage, all of which he sees as flawed (see esp. pp. 264–

As Socrates uses the verb ‘to want’ (*βούλεσθαι*) in the orators-and-tyrants passage, a sentence saying that someone wants something is false if what the person is said to want is not good. When *βούλεσθαι* is used in this way, the sentence in question has truth conditions that are different from those that the sentence would have if *βούλεσθαι* were used as Polus uses it, and as presumably most Greeks of this time would use it. So Socrates does use the verb ‘to want’ in a special way here. But from this it does not follow that he and Polus are speaking of different things, and hence cannot disagree. The notion of Socratic wanting is meant to express a truth about the underlying structure of human motivation. If we recognized this structure, Socrates appears to think, we would see that the notion is legitimate and useful. Not everyone would agree with his picture of human motivation, and he can disagree with those who reject it.

Socrates is aware that his construal of ‘wanting’ is not ordinary. When he introduces the distinction between doing what one wants, on the one hand, and doing what one pleases, on the other (*Gorg.* 466 c 9–467 c 4), he deliberately goes against Polus’ prior implicit identification of the two. He has quite a bit of explaining to do before it becomes clear what he means by his claim that Polus has raised two questions rather than one (466 c 7–e 2). None the less, he speaks as if Polus is in some way committed to the distinction, whether he realizes this or not. The very fact that Socrates proceeds to produce an argument, at 467 c 5–468 e 5, for the thesis that orators and tyrants do not do what they want to do indicates that he does not take himself to be merely stipulating a new sense for the verb ‘to want’. His argument starts from a more or less ordinary sense of ‘wanting’. He begins by making claims about wanting that appear acceptable to Polus, as a person with commonsensical views about such matters, but somehow, at the end of the argument, Polus finds himself obliged to agree to the claim he had a little earlier labelled

7). The interpretation that seems to me closest to the truth is that of E. R. Dodds. I cannot agree with him when he says, in his classic commentary on the *Gorgias*, that the concept of wanting employed in the orators-and-tyrants passage—which he construes as the concept of what one really wants as opposed to what one thinks one wants—is ‘perhaps only fully intelligible in the light of Plato’s later distinction between the “inner man” who is an immortal rational being and the empirical self which is distorted by earthly experience’ (Plato’s *Gorgias*, edited with a commentary by Dodds (Oxford, 1959), 236). The passage appears to me to be fully intelligible without any such distinction. None the less, Dodds seems to me to be quite right in taking the notion of wanting here as special, and in recognizing that this does not vitiate Socrates’ argument.

‘outrageous’ and ‘monstrous’.<sup>17</sup> So it seems that the not exactly ordinary construal of wanting which Socrates proposes to Polus is meant to be connected with what Polus and others normally understand by ‘wanting’.

At 468 B 1–4 Socrates formulates the following general claim about human motivation for action: ‘Therefore it is because we pursue what is good that we walk whenever we walk—thinking that it is better to walk—and, conversely, whenever we stand still it is for the sake of the same thing that we stand still, [namely, for the sake of] what is good.’ Although Socrates does not mention desire (other than wanting) in the *Gorgias* passage, he presumably would not deny that desires move us to act. However, looking at actions in terms of desire, the same principle holds—that we do whatever we do because we pursue what we take to be good—since Socrates believes that people always desire what they take to be good.

For this understanding of desire, we should look at *Meno* 77 B 6–78 B 2. The argument in this passage is meant to bring Meno round to the view that everyone desires good things. Socrates puts the following question to Meno: ‘Do you assume that there are people who desire bad things [*τῶν κακῶν ἐπιθυμοῦσιν*], and others who desire good things [*τῶν ἀγαθῶν*]?’ (77 B 7–C 2). Further below, the object of desire turns out to be what the person who desires takes to be good, not what as a matter of fact is good. As for those who at first appear to Meno to desire what is bad (77 C 2–3), Socrates argues that they desire what they do thinking (*οἰόμενοι*) that it is good, and not recognizing (*γινώσκοντες*) that it is bad (77 C 3–E 4). Those who appear to desire what is bad are also described by Socrates as being ignorant about the object of their desire (*ἀγνοοῦντες αὐτά* [sc. *τὰ κακά*], 77 E 1 and E 2).

The object of desire according to the *Meno* passage is what people take to be good, whether or not their belief is correct. We should think of this as holding of all desiderative and volitional states: no one desires or wants a thing unless he takes it to be good. The sort of wanting Socrates invokes when he says that orators and tyrants do not do what they want to do is no exception; it fits entirely into the general theory of desire outlined in *Meno* 76–8. One does not Socratically want something without taking it to be good. But the notion of Socratic wanting is stronger, because the agent who

<sup>17</sup> An analysis of Socrates’ argument at 467 C 5–468 E 5 is provided in the appendix.

Socratically wants to  $\phi$  does not merely take  $\phi$ -ing to be good; he recognizes  $\phi$ -ing to be good. Thus Socrates does not waver between two different accounts of desiderative and volitional states, unclear whether it is the good or the ‘apparent good’ (that is to say, what people take to be good) that is the object of such states,<sup>18</sup> as some have suggested. He has a unified view of desire that covers all its modalities, plus a special notion of a volitional or desiderative state that is also a cognitive state. Socrates does think that this sort of wanting in some way underlies all other desiderative and volitional states. This, however, is part of a substantive philosophical position, not the result of an elementary confusion. I shall address this position in Section VII below. In the two sections that follow, I wish to bring out the larger significance of the orators-and-tyrants passage.

## V

The ostensible conclusion of the discussion between Socrates and Polus at *Gorg.* 466 A 9–468 E 5 is simply that orators and tyrants—when engaged in killing, expropriating, and banishing—do not do what they want to do (468 E 3–5; see also 468 D 6–7). But Socrates’ concern is clearly with anyone who does τὰ κακά, what is bad or wrong. Much later in the dialogue, at 509 E 2–7, he expressly formulates the conclusion of the argument in these wider terms. Talking now to Callicles, he refers back to his discussion with Polus. He says:

Why don’t you answer at least this question, Callicles? Do you think that Polus and I were rightly forced to agree in our previous discussion [ἐν τοῖς ἔμπροσθεν λόγοις] that no one does what is unjust (or what is wrong)<sup>19</sup> wanting to [μηδένα βουλόμενον ἀδικεῖν], but that all who do what is unjust (wrong) do so unwillingly [ἄκοντας]? (*Gorg.* 509 E 2–7)

The conclusion of the discussion with Polus is now formulated as follows: no one who does what is wrong does so βουλόμενος, wanting to. βουλόμενος is directly contrasted with ἄκων, unwillingly, suggesting that we should construe βουλόμενος here as equivalent

<sup>18</sup> The term ‘apparent good’ is Aristotle’s. It is, however, used by interpreters of Plato, especially to refer to the confusion Socrates is alleged to suffer from. See n. 16 above.

<sup>19</sup> The term ἀδικεῖν, ‘to do what is unjust’, can be used more broadly to include doing anything that is wrong.

to *έκών*, willingly.<sup>20</sup> If so, the conclusion of the orators-and-tyrants passage turns out to be the claim that no one errs willingly. For a more familiar wording of this claim, see *Prot.* 345 E 1–2: *οὐδένα ἀνθρώπων έκόντα έξαμαρτάνειν*, no human being errs willingly. The *Protagoras* passage reads in full:

For [says Socrates] Simonides was not so uneducated [*ἀπαιδευτος*] as to say that he praised whoever did nothing bad willingly [*ὅς ἂν έκών μηδέν κακόν ποιῆ*], as if there were anyone who willingly did bad things [*κακά*]. I am pretty sure that none of the wise men thinks that any human being errs willingly [*οὐδένα ἀνθρώπων έκόντα έξαμαρτάνειν*], or willingly does anything shameful or bad [*αἰσχρά τε καὶ κακά έκόντα έργάζεσθαι*]. They know well that all who do what is shameful or bad [*πάντες οἱ τὰ αἰσχρὰ καὶ τὰ κακά ποιούντες*] do so unwillingly [*ἄκοντες ποιούσιν*]. (*Prot.* 345 D 6–E 4)

The Greek verb translated as ‘to err’, *έξαμαρτάνειν* or *άμαρτάνειν*, ranges over a wide territory. It covers both doing wrong, in a moral sense, and simply going wrong, in the sense of making an error. This suits Socrates’ purposes very well. We might try to capture the way in which *άμαρτάνειν* is suitable for his purposes by stating his position this way: no one commits injustice or does what is wrong willingly, but everyone who does wrong goes wrong. When wrongdoing is thought of as involving an error or mistake, it is easy to conclude that this is something one would not want to do. But however felicitous *άμαρτάνειν* may be for Socrates’ purposes, he does not rely too heavily on the properties of this particular word.<sup>21</sup> When he suggests that Simonides was not so uneducated as to imply that a human being errs willingly, he may well be ironic, and in more than one way. None the less, he associates a recognition that no one errs willingly with education and wisdom, thus treating it as something that requires insight.

<sup>20</sup> I am not suggesting that Socrates always uses *έκών* as equivalent to *βουλόμενος*, but only that he does so in this specific context. For that matter, he does not always use *βουλόμενος* in the sense of Socratic wanting either. He does not do so, for instance, later in the *Gorgias*, at 511 B 4.

<sup>21</sup> A lot has been written about *άμαρτία*, especially in connection with Greek tragedy and Aristotle’s *Poetics*. From the point of view of this paper, the most useful discussion is that of T. C. W. Stinton, ‘*Hamartia* in Aristotle and Greek Tragedy’, *Classical Quarterly*, NS 25 (1975), 221–54; repr. in Stinton, *Collected Papers on Greek Tragedy* (Oxford, 1990), 143–85.



## VI

At *Gorg.* 509 E 2–7 Socrates gets Callicles to agree that no one does what is unjust or wrong wanting to, but that all those who do so do it unwillingly. The larger immediately relevant passage starts at 509 C 6. Socrates has been focusing his and Callicles' attention on two evils—the evil of suffering injustice (*ἀδικεῖσθαι*) and the evil of doing it (*ἀδικεῖν*). Now he raises the question of what it would take for us to save ourselves from falling into each of the two evils. In each case, he asks, is it *δύναμις*, power, or *βούλησις*, wish—as *βούλησις* is customarily translated—that enables us to avoid the evil in question?

To avoid being treated unjustly, Socrates and Callicles quickly agree, one needs power (509 D 3–6). But what about doing what is unjust: is it *δύναμις* or *βουλήσις*—power or wish—that saves us from this evil? Socrates permits Callicles to say that one needs power in this case as well (510 A 3–5), even though just a moment ago he had secured Callicles' agreement to the conclusion of the previous discussion with Polus, that no one does what is unjust *βουλόμενος*, wanting to so do (509 E 5–7). He intends Callicles to make the required connection between *βούλησις* and *βουλόμενος*. Like Polus before him, Callicles does not quite get Socrates' point. But Callicles is not entirely wrong in his answer, and this may be the reason why Socrates lets him off as he does. *βούλησις*—as construed by Socrates—is sufficient for a person not to do what is unjust. But this *βούλησις*, of course, is not merely a *wish*, but rather *wanting* or *willing* in the highly specific sense that Socrates had introduced in his discussion with Polus, and reintroduced here in his discussion with Callicles. This kind of wanting or willing is (in a certain sense) power. Socrates' point is the following. To avoid becoming a victim of an unjust action, one needs power in the straightforward sense; indeed, the power often needed is brute force. To avoid committing injustice, on the other hand, what a person needs is that his will be in a certain condition. When one's will is in this condition, one has all the power one needs, and all the power one can have, not to do what is unjust.

In speaking here of one's will being in a certain condition, I am of course relying on some more current notion of the will. There has been a long-standing dispute over the question whether the ancients

had any notion of the will. Presumably, given the large number of widely different conceptions of the will that have emerged in Western philosophical thought since antiquity, the question is whether any of the ancient thinkers had a notion that is in some important way linked to one or more of these later notions. In his claim that orators and tyrants do not do what they want to do, as well as his claim that no one errs willingly, as I have interpreted these claims, Socrates introduces—apparently for the first time in Greek philosophical thought—a certain notion of the will, or something very much like such a concept. This notion of the will is in some ways peculiar. The *βούλησις* in question—the will, understood as I have suggested—prevents us from doing anything that is wrong. If so, this will—which is essentially the good will—cannot be weak. (This point is linked to Socrates' denial of *akrasia*, which I discuss below.)

We should not fail to notice the playfulness with which Socrates takes up the question whether it is *δύναμις* or *βούλησις* that can save us from the evils of suffering and of doing injustice (509 D 2–510 A 5). The playfulness in part depends on the usual meaning of *βούλησις*—that of a wish. Socrates asks about *δύναμις* and *βούλησις* in the course of renewing his argument for the view that the evil of suffering injustice is utterly trivial in comparison with the evil of doing injustice. To acquire things that are much prized by people, a great deal of power is usually needed. What makes the tyrant so enviable to many is the tremendous power he has—power so unrestricted that he can deprive people of what are considered to be their greatest goods: their life, their property, a place in their own city. It now turns out that the evil which it is incomparably more important to avoid—the evil of doing injustice—does not require the usual machinery of power. It would seem, in fact, that nothing could be easier than securing something by means of a wish. Neither the brute force the tyrant employs, nor the skilful manipulation of the soul by words that the orator relies on, is required here. What one needs, Socrates appears to suggest, is hardly anything at all: a mere wish, *βούλησις*. However, if one follows him to the end of his thought, it transpires that this thing—*βούλησις* in the sense of Socratic wanting or willing—is something that it is tremendously difficult to have.

Power was also the main ground on which earlier in the dialogue the great orator Gorgias had defended and praised oratory. In ar-

guing that orators are at least as powerful as tyrants, Gorgias had relied on the enormous and nearly universal appeal of power. Polus inherited his argument from Gorgias. Thus in discussing the tyrant's actions of killing, expropriation, and banishing with Polus, Socrates is still addressing Gorgias' defence of oratory. Socrates now in response leaves his three interlocutors, Gorgias, Polus, and Callicles, with the following dilemma: either the power that enables a person to inflict what people consider to be the greatest evils on others is not good, and hence not something to be in the least admired, coveted, or envied; or else if power as such is good, orators and tyrants have none of it.

The notion that power as such is something good—clearly a notion that all three of his interlocutors are eager to push—undergoes a peculiar, deliberate transformation at Socrates' hands. He in effect offers his interlocutors an option of choosing between two concepts of power. In both cases power is the ability to do as one wants. On the first concept, a person is powerful if he can do what he wants or desires, as the words 'wants' or 'desires' are usually understood. On the second concept, a person is powerful if he can do what he wants in the more special sense—in the sense of what I have called Socratic wanting. Socrates is not blind to the fact that this notion is a novelty to his interlocutors. What he wants is to recast the debate in a novel way. Gorgias, Polus, and Callicles may insist as much as they please that power, as they understand it, is good. They are simply wrong about this. Relying now on the second concept of power—the one that Socrates himself is pushing—virtue is power. To express his thought in a different way: a certain kind of knowledge, and a certain kind of will, are power.

## VII

Socrates seems to propose his special notion of wanting—that of Socratic wanting—not as a notion we already have at work in our language, but rather as a notion that we occasionally grope for, and a notion that we need. We need it because it enables us to express something that is of relevance to all the willing, wishing, and desiring that we ordinarily do and ordinarily speak of.

The notion of Socratic wanting announces a certain ideal. There is nothing arbitrary, however, about this ideal. Desires and wants of all varieties are, as we would put it, intentional phenomena. They

are directed towards something. In Socrates' view, they embody a certain direction of the soul: a striving of the soul for what is good, and a striving of the soul for its own good, or perhaps for the good proper to a human being. The ideal of wanting that he introduces in the orators-and-tyrants passage, and in its follow-up later in the *Gorgias* (509 c 6 ff.), is meant to embody the shape that this striving of the soul takes when the soul has got a grip on what the good that it is after in fact is.

The Socrates of Plato's early dialogues does not often invoke human nature. But here is what we find him saying about it in the *Protagoras*:

Now, no one goes willingly towards things that are bad [ἐπί γε τὰ κακὰ οὐδείς ἐκὼν ἔρχεται] or towards those one thinks are bad [οὐδὲ ἐπὶ ἃ οἴεται κακὰ εἶναι], nor is it in human nature [ἐν ἀνθρώπου φύσει], so it seems, to want to go towards what one thinks is bad instead of to what is good [ἐπὶ ἃ οἴεται κακὰ εἶναι ἐθέλειν ἵέναι ἀντὶ τῶν ἀγαθῶν]. And when one is forced to choose between two bad things, no one chooses the greater if he is able to choose the lesser. (*Prot.* 358 c 6–d 4)<sup>22</sup>

We, humans, are hardwired to seek our own good. What we want is, ultimately, to do well for ourselves. The striving for this condition of doing well, which Socrates calls 'the good', is something that every human soul comes equipped with. Striving after the good is as basic to the human soul as is its striving after the truth.

With regard to the considerations that impelled Socrates to introduce his special concept of wanting, it may be useful to quote a passage from outside what we consider Plato's Socratic writings, even if we do not, as we should not, treat it as evidence for the Socratic view:

And isn't this also clear? In the case of just and beautiful things [δίκαια . . . καὶ καλὰ], many would accept things that are believed (reputed) to be so [τὰ δοκοῦντα], even if they are in fact not so, and they do such things, acquire them, and get a reputation for doing and acquiring them [ὁμῶς ταῦτα πράττειν καὶ κекτῆσθαι καὶ δοκεῖν].<sup>23</sup> But when it comes to good things, no

<sup>22</sup> I return to this passage in sect. x below.

<sup>23</sup> Older English translations of this passage seem to me greatly preferable to the more recent ones. The first English translator, Spens (1763), is very much on the right track: 'But what, is it also not evident, that with reference to things just and beautiful, the multitude chuse the apparent, even though they be not really so, yet they act, and possess, *and are reputed of accordingly*; but the acquisition of goods . . .' (emphasis added). The best rendition, to my mind, is that of Davies and Vaughan (1852), who clearly take ταῦτα πράττειν καὶ κекτῆσθαι as the antecedent of

one is content to acquire things that are believed to be so [*ἀγαθὰ δὲ οὐδενὶ ἔτι ἀρκεῖ τὰ δοκοῦντα κτᾶσθαι*], but everyone seeks things that are in fact good [*ἀλλὰ τὰ ὄντα ζητοῦσιν*] and spurns mere belief [*τὴν δὲ δόξαν ἐνταῦθα ἤδη πᾶς ἀτιμάζει*] . . . This, then, [sc. the good] is what every soul pursues [*διώκει*] and for the sake of which it does everything it does . . . (*Rep.* 505 D 5–E 1)

Whatever special interpretation Plato might be putting in the *Republic* on the distinction between τὰ δοκοῦντα and τὰ ὄντα—things that are reputed (opined, believed) to be good, on the one hand, and things that are good, on the other—there can be no doubt that the Socrates of the early dialogues is interested in a similar distinction: a distinction between what *appears* to be good, and what *is* good. Towards the end of the *Protagoras*, Socrates announces that it is the power of appearance (*ἡ τοῦ φαινομένου δύναμις*) that makes us wander all over the place and regret our actions and choices (356 D 4–7). We mistakenly take for good things that in fact are not good, but merely appear to us to be so. If we had knowledge about what is good and bad, the appearing (*τὸ φάντασμα*) would lose its grip over us (become *ἄκυρον*, 356 D 8); consequently, we would achieve peace of mind (*ἡσυχία*, 356 E 1) and salvation in life (*σωτηρία τοῦ βίου*, 356 D 3; see also 356 E 2, E 6, E 8, and 357 A 6–7).

Furthermore, both the Socrates of the *Republic* and the Socrates of the *Protagoras* take goodness to be an evaluative property of a special sort. No other question is of more importance to the business of living than the question ‘Is this (what I am about to do, what I contemplate doing, what I am doing) really good?’ We might believe that the action we are considering is admirable or useful; or that we shall be envied for it; or perhaps that it is in keeping with our outlook, although we shall be despised for it. But the nagging question always remains whether the action under consideration is really good; whether in acting as we do, we do good for ourselves.<sup>24</sup> This concern is the driving force behind much

*δοκεῖν*: ‘Once more: is it not evident, that though many persons would be ready to do and seem to do, or to possess and seem to possess what seems just and beautiful, without really being so; yet when you come to things good . . .’ (quoted from the 4th edn. (1868); emphasis added). *καὶ δοκεῖν* is misconstrued by Lindsay (1935), Grube (1974), and Grube–Reeve (1992). Until they provide a parallel for their construal of *δοκεῖν*, the older translations must take precedence. The Grube–Reeve translation reads: ‘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’ (emphasis added).

<sup>24</sup> We should set aside the complaint that Socrates wavers between two different

ethical reflection. But it is a concern that is operative already at a pre-reflective level. What the nagging question brings out is that we aim—pre-reflectively no less than reflectively—not at what appears good, but at what is in fact good.

Thus the special, Socratic wanting is what wanting becomes when we have tracked down what we have been after all along. What we have been after all along—what our desiderative states are always tracking down—is where our well-being in the world lies.

### VIII

In saying that no one errs willingly Socrates has in mind, roughly, that no one does what is wrong recognizing it as wrong and wanting it as one wants things one recognizes to be good. We might find it helpful to put the thought this way: no one does what is wrong knowingly and willingly. But Socrates has no need to add ‘knowingly’ to ‘willingly’, since his claim that no one does what is wrong willingly implies that no one does it knowingly. If ‘willingly’ is understood as I have suggested, the claim is clearly not that wrongdoing is involuntary in Aristotle’s sense of the word (see *NE* 3. 1). If one thinks that Socrates takes wrongdoing to be involuntary in Aristotle’s sense of the word (or in something close enough to this sense), one will feel a need to explain how he came to embrace such a view. This, I think, is what gives rise to the mistaken belief that he infers that no one does what is wrong willingly from the idea that wrongdoing involves ignorance. He fails to realize—unlike Aristotle after him—that only certain kinds of ignorance concerning one’s action make that action involuntary (cf. Section II above). On the reading I have proposed, Socrates’ claim makes perfect sense; it does not reflect any such gross failure of judgement.

Special as the notion of Socratic wanting or willing is, it is part of a larger disagreement with many of us. Socrates believes, for instance, that all who do what is wrong do so simply because they go wrong. Wrongdoers do not aim at something they recognize as wrong or bad; rather, they are misguided and ignorant about the nature of their action and its goal. Further, the thesis that no one

questions—the question of what is good, and the question of what is good for the agent. The more basic question for him is: what is good? He does also think that everyone seeks his own good. However, since ‘what is good for the agent’ has little antecedent content, it is left open what the content of the ultimate good will turn out to be. The ultimate good need not be egoistic.

errs willingly, as will transpire shortly, implies that *akrasia* is not possible. This is certainly not what many of us today think about weakness of the will, or what many people thought about *akrasia* in Socrates' own time.<sup>25</sup>

We ought to start, however, with the position that Socrates takes himself to be denying when he rejects *akrasia*. At *Prot.* 352 D 4–7 Plato formulates with some care the position that Socrates rejects:

You [says Socrates to Protagoras] know that the many [οἱ πολλοὶ τῶν ἀνθρώπων] are not going to be persuaded by us. They say that a lot of people [πολλοὺς], recognizing what is best [γινώσκοντας τὰ βέλτιστα], do not want to do it [οὐκ ἐθέλειν πράττειν], when it is possible for them to do so [ἐξὸν αὐτοῖς], but do something else instead [ἀλλὰ ἄλλα πράττειν] . . .<sup>26</sup>

The view that Socrates rejects—imputed to and indeed put into the mouth of ‘the many’—is that a lot of people act against their recognition, that is to say, against their knowledge, of what is best. This I take to be Socrates' primary, or official, characterization of *akrasia*.

Nowadays weak-willed action is often characterized as action against one's better judgement—one's judgement of what, under the circumstances, is the better thing to do. When understood in this way, there is no reason why an akratic action could not in principle be a good thing to do, or at any rate better than the action which the agent (incorrectly) takes to be better. However, according to the characterization of *akrasia* which Socrates gives in the passage quoted, akratic action is by assumption wrong: the akratic agent does what is wrong knowing that it is wrong, considering or having considered a different course of action that is open to him,<sup>27</sup> which

<sup>25</sup> To be sure, we should be careful here. We are dealing with more than one party. We are not Socrates' interlocutors: when he discusses *akrasia*, his interlocutors are the ‘the many’ that he conjures up (see, in particular, *Prot.* 352 B 2–3 and 352 D 5). The notion of *akrasia* that Socrates rejects is somewhat different from the notion (or notions) that we nowadays have of weakness of the will. None the less, as I shall later argue, there is some reason to think that he would not only deny the existence of *akrasia* as (he thinks) his contemporaries conceive of it; he would also deny the existence of *akrasia* as many of us nowadays think of it.

<sup>26</sup> Having formulated what an akratic action is, Socrates goes on to state the cause, τὸ αἴτιον, that the many cite to explain such an action. I shall follow him in keeping the issue of the ‘cause’ separate.

<sup>27</sup> Davidson argues that Socrates—or, strictly speaking, G. Santas, whose interpretation of Socrates Davidson discusses—fails to realize that an alternative course of action need not in fact be open to the agent, because the agent's *belief* that the course of action is open to him is sufficient. (Donald Davidson, ‘How is Weak-

he knows to be better or best. It is because Socrates construes *akrasia* in this way, and not merely as action against one's better judgement, that his denial of *akrasia* follows from his No One Errs Willingly thesis.

One important aspect of the official characterization of *akrasia* at *Prot.* 352 D 4–7 has been generally overlooked. The many, Socrates says, claim that a lot of people, recognizing what is best, do not want to do it (*οὐκ ἐθέλουν πράττειν*), when it is possible for them to do it, but do something else instead. He invokes wanting here, and builds it into the characterization of *akrasia* offered by the many (see also *ἐθέλει* at 355 B 2, *ἐθέλουν* at 358 D 2, and *ἐθελήσει* at 358 E 3). Thus the thesis he intends to deny is not just that one can fail to do what one recognizes is best, but more fully that an agent may recognize what is best and yet not want, or not be willing, to do it, and consequently, not do it. By contrast, we have to assume, Socrates contends that a person who knows what the right thing to do is, does want to do it and, other things being equal, will do it. (The more neutral word for wanting, *ἐθέλουν*, that he uses here is appropriate since the position he is denying is that of the many, who would not put their own point in terms of Socrates' special notion of wanting or willing. To express his own position, Socrates could use either the more neutral *ἐθέλουν* or the more specific *βούλεσθαι*.)

If Socrates uses 'willingly' in a special way when he claims that no one errs willingly, to designate a volitional act that is also cognitive, does this not make his claim problematic? His concept of willing is not ours. What can we do with such a peculiar concept? In response, I shall match these questions with another one. Socrates' rejection

ness of the Will Possible', in *Essays on Actions and Events* (Oxford, 1980), 21–42 at 22 n. 1; the reference is to Santas's article 'Plato's *Protagoras* and Explanations of Weakness', *Philosophical Review*, 75 (1966), 3–33. Davidson's remark is appropriate given his perspective on weakness of the will. But Socrates is not in error here. His characterization reflects his approach to *akrasia*. Unlike us (and, to some extent, unlike Aristotle), Socrates approaches *akrasia* from outside in. He focuses on putative akratic actions that are in fact cases of wrongdoing, where an alternative course of action is available to the agent and is *recognized* by the agent as available to him. Socrates is interested in (putative) *akrasia* as an ethical problem—primarily an ethical problem of wrongdoing of some kind. One could, of course, restate his views on *akrasia* by taking in an inside-out approach, focusing on the beliefs that the agent has about the action he takes and about the alternative course of action available to him. I shall go on to address this approach (which Socrates to some extent adopts later on in the *Protagoras*: see 358 C 6–D 4), but only after I have looked into what I take to be the primary account of *akrasia*, given in a preliminary fashion at *Prot.* 352 B 5–8, and then, more carefully, in the passage just quoted, 352 D 4–7.



of *akrasia* amounts to the view that one cannot act against one's knowledge of what is best. Now the conception of knowledge that underlies this view should strike us as at least as peculiar as the concept of Socratic wanting. Here is what Socrates has to say about the relevant kind of knowledge:

Now, do you [Protagoras] too think that that is how things stand with it [sc. knowledge], or do you think that knowledge [*ἐπιστήμη*] is fine and such as to rule the person, and if someone recognizes what is good and bad [*ἐάνπερ γινώσκῃ τις τὰγαθὰ καὶ τὰ κακά*], he would not be overpowered by anything else so as to act otherwise than knowledge dictates, and wisdom [*τὴν φρόνησιν*] is sufficient to help the person? (*Prot.* 352 c 2–7)

We no more share with Socrates his conception of *knowledge* than we share with him his conception of *wanting* or *willing*. But if this is so, should we regard his claim that no one errs willingly as more suspect than his claim that no one errs knowingly? As I pointed out at the beginning of this paper, the wanting or willing that the expression 'willingly' refers to involves recognition of what is good or bad; it has now turned out that the knowledge of what is good and bad involves wanting that accords with the knowledge in question. Hence, one claim is as problematic or as unproblematic as the other; both claims stand or fall together. They should also be examined together.

## IX

Socrates does not want to deny that we have episodes which we incorrectly describe as akratic or weak-willed. We should now take a look at his characterization of the 'cause' of *akrasia*, which is kept separate from the characterization of *akrasia* itself. We have heard what the many believe: that a lot of people, recognizing what is best, do not want to do it, but do something else instead (352 D 4–7). When asked what they think the cause (*αἴτιον*) of this is (D 7–8), the many—according to Socrates—reply that 'those who act in this way do so being overcome [*ἡττωμένους*] by pleasure or pain, or being overpowered [*κρατουμένους*] by one of the things I [Socrates] referred to just now' (352 D 8–E 2). Socrates has in mind the things he referred to at 352 B 5–8, the passage in which he gives his first, informal characterization of *akrasia*. According to this characterization, the many believe that 'often, although knowledge is present

in a person, what rules him is not knowledge but something else: sometimes anger [*θυμόν*], sometimes pleasure [*ἡδονήν*], sometimes pain [*λύπην*], at other times love [*ἔρωτα*], often fear [*φόβον*] . . .'. So, on the account given by the many, people act akratically—i.e. against their knowledge of what is best—because they are overcome by pleasure or pain, by desire, or by any of a number of passions.

At *Prot.* 352 E 5–353 A 2 Socrates says that he and Protagoras should now attempt to persuade and teach the many what the *πάθος* is which the many describe as being overcome with pleasure, and which in their view is why they don't do what is best when they recognize what it is. He refers to the same thing as a *πάθημα* a few lines below, at 353 A 4–6: the many, he says, will demand an explanation from him and Protagoras as to what this *πάθημα* is, if it does not amount to being overcome by pleasure.<sup>28</sup> Thus Socrates grants that a certain *πάθος* or *πάθημα*—a particular way of being afflicted—is present. What is presumed to be missing is a correct characterization of this affliction. The two words, *πάθος* and *πάθημα*, which are here used interchangeably, refer, I believe, not to the experience associated with putative *akrasia* but to the affliction of the putative akratic—namely, what a person undergoes when he undergoes what the many think of as *akrasia*. The usual translation of *πάθος* or *πάθημα* as 'experience' does not seem to me to be accurate. When Socrates draws attention to what is happening with the presumed akratic agent, he may have in mind an experience that the agent has, but he need not. For instance, when he further down declares that the *πάθημα* in question is in fact ignorance (357 C 7), he is not saying that the *experience* characteristic of putative *akrasia* is ignorance, but rather that the condition of the agent's soul that is wrongly attributed to *akrasia* is in fact ignorance.

Keeping in mind Socrates' preliminary formulation of *akrasia* at 352 B 5–8, and bypassing the hedonistic assumptions from which the discussion of *akrasia* in the *Protagoras* proceeds,<sup>29</sup> the description 'being overcome by pleasure' should be taken as representative of a number of related descriptions that the many had offered to explain *akrasia*. The presumed akratic was described at 352 B 5–8 as

<sup>28</sup> Socrates comes back to this *πάθημα* at 357 C 7. He is now ready to provide his answer to the question pressed by the many. The *πάθημα* in question, he now claims, is ignorance (*ἀμαθία*). See 357 E 2–4.

<sup>29</sup> Here and in what follows I am interested in Socrates' general position on *akrasia*. I thus aim to reconstruct the considerations on which he based his rejection of *akrasia* in a way that should be of interest to hedonists and non-hedonists alike.

being overcome not only by pleasure, but also by pain, desire, fear, love, and so on. In speaking of the condition of being overcome by passion below, I use ‘passion’ broadly, to refer to any of these states.

It would be wrong to assume that Socrates has an easy task here. Once we strip the phenomenon commonly described as *akrasia* of all the descriptions Socrates would find incorrect, it is not quite clear what remains. This, I take it, sets him a task. We speak of our akratic episodes; we know what it feels like to be in the grip of one; we understand what others have in mind when they describe theirs. This presumably is not what Socrates wants to deny. But when the many say, for instance, that people are overcome (*ἡττώμενοι* or *κρατούμενοι*) by passion, they seem to have in mind a contest of two forces: one that wins and one that is defeated (being defeated is a usual meaning of *ἡττώμενος*). Socrates, as we shall see, rejects the picture of contest between two forces as a proper description of what happens in putative *akrasia*. The agent is not really acting against his knowledge; nor is he, as I am about to argue, acting against his better judgement. If this is so, Socrates should be able to tell us how to identify the putative phenomenon of *akrasia* in a way that is independent of all the incorrect descriptions usually given of it. In the *Protagoras* he does not endeavour to do this.

## X

Let me now turn briefly to a broader notion of *akrasia*, one that involves acting against one’s judgement or belief<sup>30</sup>—not necessarily against one’s knowledge—that some course of action is best. The characterization of *akrasia* in the *Protagoras* discussed so far has not included this kind of case. However, in the course of arguing against *akrasia* as officially characterized, Socrates makes observations that amount to grounds for rejecting *akrasia* in a broader sense, viz. *akrasia* thought of as action against one’s judgement as to what is best.

The main thought behind this denial can be expressed in the following way. *Akrasia* presupposes an awareness on the part of the agent of alternative courses of action available to him. What

<sup>30</sup> One might argue that making an evaluative judgement can fall short of holding an evaluative belief. However, I think that Socrates would not want to make a distinction between belief and judgement. To judge that something is good is to take it as good, and taking something to be such-and-such is on his view a *δόξα*, opinion or belief.

supposedly happens here is this: the agent considers two courses of action; he believes that one of them is correct; none the less, he does what he believes to be wrong. The main reason why Socrates thinks this is not possible is that our actions embody our evaluative beliefs, and that they embody them in a very strong sense. By going for one of the considered alternative courses of action rather than the other, the agent shows that he takes the preferred course of action to be better. Recall again Socrates' statement about human nature:

Now, no one goes willingly towards things that are bad [*ἐπί γε τὰ κακὰ οὐδείς ἐκὼν ἔρχεται*] or towards those one thinks are bad [*οὐδὲ ἐπὶ ἃ οἴεται κακὰ εἶναι*], nor is it in human nature [*ἐν ἀνθρώπου φύσει*], so it seems, to want to go towards what one thinks is bad [*ἐπὶ ἃ οἴεται κακὰ εἶναι ἐθέλειν ἰέναι*] instead of to what is good. And when one is forced to choose between two bad things, no one will choose the greater if he is able to choose the lesser. (*Prot.* 358 c 6–d 4)<sup>31</sup>

In saying that no one goes willingly towards bad things (*ἐπὶ . . . τὰ κακὰ*), Socrates has in mind that no one goes willingly towards things that are bad, when it is transparent to the person's mind that they are bad. (Similarly for choosing between two bad things.) One reason why one cannot act against one's knowledge of what is better is that by acting so one would show one has a belief that contradicts the knowledge in question. But Socrates' practice of cross-examining his interlocutors implies that he thinks that a person who has a body of knowledge cannot have a belief that contradicts it.

The main intuition behind Socrates' denial of *akrasia* in the broader sense—the intuition that evaluative beliefs are both embodied and displayed in our actions—seems sound. This intuition is presumably something that would be understandable and in some form acceptable to the many. Socrates, however, pushes this thought much further than the many. He presumably believes that taking something to be good and going for it are connected far more tightly than people tend to think.

Now when I  $\phi$ , where this  $\phi$ -ing is a presumed akratic action, and I take myself to be acting against my belief that my  $\phi$ -ing (here and now) is wrong, or worse than an alternative action open to me, is my belief that I have such a belief an illusion? Socrates' views on *doxa*, opinion or belief, seem to push him in two different direc-

<sup>31</sup> Compare also 358 e 2–6.

tions. On the one hand, having a belief on his view implies having a commitment. Evaluative beliefs in particular involve practical commitments. So he might well argue that the presumed akratic is not committed to his professed evaluative belief to the degree that would be necessary for the ascription of the belief to him to be correct. If he took this line, what Socrates would be telling the presumed akratic is this: you claim to believe that your  $\phi$ -ing (here and now) is wrong, but in fact you don't believe that. What you in fact believe, as your action shows, is that  $\phi$ -ing (here and now) is good or right.

However, Socrates often uses *doxa* in a considerably more relaxed way. For instance, each of his interlocutors is said to have an opinion or belief whenever he sincerely agrees with the view that Socrates proposes for consideration. When we read in our translations of the *Protagoras* that this or that interlocutor 'concurred with' Socrates or 'agreed with' him, the word not infrequently used is *συνδοκεῖν*. The very word indicates that the interlocutor shares Socrates' *doxa*; that he believes (opines: *δοκεῖν*) that things are as Socrates says they are. The interlocutors often have a poor grasp of the content of what they agree to, and this (among other things) leads them to contradict themselves. Socrates takes such a contradiction as an indication that the interlocutor does not have knowledge, not that he does not have the relevant opinion or belief.<sup>32</sup>

When *doxa* is understood in this relaxed way, Socrates should say, as before, that the agent believes that his  $\phi$ -ing (here and now) is right or good, since this belief is implied by his action. However, Socrates should now also grant that the presumed akratic agent believes that his  $\phi$ -ing (here and now) is wrong. Now if Socrates

<sup>32</sup> For *συνδοκεῖν* see 358 B 6, C 3, C 6, and D 4. The four occurrences of *συνεδόκει* are part of an important global dialectical move Socrates makes at 358 A 1–359 A 1, at the end of his case against *akrasia*. He secures the agreement here not only of Protagoras, but also of Hippias and Prodicus, to the claim that pleasure is the good (358 A 5–B 6); to his denial of *akrasia* and his diagnosis of what in fact happens in putative *akrasia* (B 6–C 3), along with his explanation of what *ἀμαθία* is (C 3–6), and to the claim that no one willingly goes towards what he thinks is bad (C 6–D 4). In each case, *συνεδόκει* punctuates the concurrence in belief among the four principal interlocutors. What is at issue at 358 A 1–359 A 1 is what Protagoras, Hippias, and Prodicus believe, or rather have come to believe, having been persuaded by Socrates (see the instances of *δοκεῖν* at 358 A 3 and A 4, D 7, and E 6). Socrates then moves on to secure the agreement of all three to his definition of fear (358 D 5–359 A 1). Having secured these admissions, Socrates immediately (starting at 359 A 2) moves to show that Protagoras' position on the unity of virtue is incompatible with the admissions he has just made.

takes this line, then on his view an agent can after all act against his belief that his  $\phi$ -ing (here and now) is wrong. Would this amount to a recognition of *akrasia* on Socrates' part? I have in mind here a recognition of *akrasia* understood in the broader sense, i.e. *akrasia* thought of as action against one's belief about what is better or best. Although it is true that on this analysis the agent acts against his belief, I am inclined to think that this is not what those who hold that *akrasia* exists for the most part have in mind. Being akratic does not consist merely in acting against a belief, in a weak sense of this word, that something is good.<sup>33</sup> Although of course there are many conceptions of what *akrasia* consists in, the agent is usually thought to be acting against something a bit stronger than this sort of belief. I shall come back to this question in a moment.

Let me return to *akrasia* as originally defined—namely, as action against one's knowledge of what is better or best—and look at the diagnosis Socrates would give of the presumed akratic's predicament. It seems reasonable to assume that Socrates sees the putative akratic as himself believing that he acts against his knowledge of what is best. Admittedly, in his official characterization of *akrasia* at *Prot.* 352 D 4–7 Socrates does not explicitly state that the many believe *of themselves* that they often know what is best and yet do something else; his claim is rather that the many (*οἱ πολλοί*) allege that akratic episodes happen to many (*πολλοῖ*). Although this invites us to think that the relevant ascriptions of knowledge involve self-ascriptions, the formulation itself does not settle the question whether such self-ascription of knowledge is constitutive of (what passes for) *akrasia*.

There is some reason to think that it is. What makes it so difficult to deny *akrasia* is precisely the repeatedly insistent first-person claim: 'but—whatever your theory—I knew full well that what I was going to do was bad; yet I did it.' When we find it difficult to go along with philosophical worries concerning the existence of *akrasia*, we do so not because we are confident about third-person ascriptions of knowledge to agents who happen to act against their knowledge. What makes it difficult to deny *akrasia* is rather the first-person experience of going against one's own firm conviction that something is bad, often because of some powerful desire or

<sup>33</sup> This, I presume, is why Aristotle comes to think of *akrasia* as action against the agent's *προαίρεσις*, choice. As he understands it, *προαίρεσις* involves a firm practical commitment.

impulse. Is the firmly held conviction taken by the akratic himself to be a case of knowledge? In everyday life people often describe their weak-willed episodes this way; they do it when they say, for instance: ‘But I *saw* clearly that this was bad; yet I did it.’ This conception of *akrasia* is vividly conveyed by Ovid’s memorable ‘video meliora proboque, deteriora sequor’, cited at the head of this paper. Except for being couched as a first-person statement, Ovid’s formulation is strikingly close to Socrates’ own (at *Prot.* 352 D 4–7): the akratic acts against what he *sees* is better—the Greek *γινώσκειν*, to recognize, has become the even more emphatic Latin *videre*, to see.<sup>34</sup>

If the akratic agent believes, Ovid-style, that he acts against his knowledge of what is better or best, Socrates’ case against *akrasia* implies that this belief is false. Socrates would diagnose the Ovid-style akratic as suffering from the affliction he believed it was his task to unmask, and if possible, eradicate (see the *Apology*): ignorance of one’s own ignorance. The akratic agent not only lacks knowledge of what is better or best; he also wrongly believes that he possesses this knowledge. When Socrates declares that the *πάθος* or *πάθημα* of being overcome by pleasure (or, in general, passion) is in fact *ἀμαθία*, ignorance<sup>35</sup>—indeed, *ἀμαθία ἢ μεγίστη*, the greatest ignorance—he might have in mind the ignorance of what is good or bad. However, every wrongdoer is on Socrates’ view ignorant of what is good or bad. What is specific to the central type of wrongdoing that the many incorrectly describe as akratic is the specific ignorance of one’s own ignorance that this type of wrongdoing involves.

Now what would be Socrates’ diagnosis of the putative akratic condition if *akrasia* is construed more broadly, as action against one’s better judgement? Here the diagnosis would have to await a more precise description of what *akrasia* is. There is a view ac-

<sup>34</sup> One interesting difference is that Socrates does not find it necessary to make separate mention of the approbative attitude that goes along with the recognition of what is better or best. The relevant knowledge or recognition, on his view, implies an approbative attitude.

<sup>35</sup> See *Prot.* 357 E 2–4: ‘So this is what being overpowered by pleasure [*τὸ ἡδονῆς ἤττω εἶναι*] is, the greatest ignorance [*ἀμαθία ἢ μεγίστη*], which is what Protagoras here and Prodicus and Hippias claim to cure.’ What Protagoras professed to teach earlier in the dialogue was (civic) virtue (318 E 5–319 A 2). It is because he claims to teach this (among other reasons) that Socrates, ironically, counts him as being on his side in the argument against the many. Protagoras had better know what putative *akrasia* is, since this issue is at the heart of his professed expertise.

according to which we can go for something without taking it in any way as good. Values are here seen as being at some remove from the desires or impulses on which we act. So on this view an akratic may, for instance, do something in spite of his judgement that what he is about to do is bad; he does so simply because he ‘feels like it’, not because he values it in some way. If this is what it means to act against one’s better judgement, Socrates would deny that such *akrasia* exists. He would do so because he would reject the view that the agent can act without taking anything to be good or bad. As the *Meno* passage referred to above indicates (as well as the statement in the *Protagoras* about human nature quoted above, 358c 6–d 4), when an agent acts on a desire, he acts in accordance with the value judgement involved in the desire. This judgement is the one that motivates his action.

But suppose the opponent grants a part of Socrates’ point, admitting that our actions are shot through with value judgements, and that value judgements are not motivationally inert; suppose he also agrees that it is not possible to go for something without considering it good in some way. The opponent might none the less think that it is possible to act against one’s reflectively considered scheme of values, and he might propose that the ‘better judgement’ against which the akratic acts be identified with such a reflectively considered judgement. Would Socrates deny this?

Socrates would not be the one to deny that reflective thought can generate values. He could also hardly deny that one’s impulses might go against reflectively generated values. However, *akrasia*, as usually understood, is not an affliction that consists merely in holding contradictory evaluative beliefs, and acting sometimes on one such belief and sometimes on another. *Akrasia* is more than confusion about values. A proponent of *akrasia* usually regards the so-called ‘better judgement’ as something more than a mere judgement that some course of action is better. The ‘better judgement’ is ‘better’ because it is reflectively endorsed; or because it has higher epistemic credentials; or because it is the judgement with which the person more fully and directly identifies.

For instance, having considered the evaluative point of view that pushes me into this action, I may form a judgement that the evaluative viewpoint in question is not one that I can ultimately embrace; a more carefully considered judgement, or a judgement that rests on a wider point of view, or a judgement that expresses more di-



rectly my will—these are the candidates for that ‘better judgement’ against which I act when I act akratically. But the more weight we thus put on the notion of better judgement, the less likely it is that Socrates would agree with us. It is likely that he would stick to his basic intuition that our actions reveal more about us and our values than any product of detached reflection might. As he would see it, the mere fact that a desire is a second-order one, or that it is endorsed by some second-order thinking on our part, is neither here nor there. The reflection he is interested in is practical reflection: one that changes preferences, and goes all the way down, to influence the very valuations on which we act. The more we add to the notion of better judgement in terms of one’s identification with it, the closer we get to the grounds on which Socrates refused to admit that one can ever act against one’s knowledge of what is better.

On his understanding of what knowledge is, in order to know that a course of action is good, it is not sufficient to believe that it is good, and to hold this belief for the right reasons. If one knows something, then on Socrates’ view, one cannot have a belief that contradicts that knowledge. Knowledge of what is good precludes false appearances of goodness. This suggests that knowledge—the sort of evaluative and practical knowledge that he has in mind when he speaks of the knowledge of good and bad—cannot be had in bits and pieces. To have the relevant sort of knowledge is to be in possession of a certain regulative and organizing principle that is in control of the overall condition of the soul. Socrates seems to think of knowledge as a condition in which none of one’s doxastic commitments ever goes unheeded. One does not concur with a certain opinion, and then proceed to concur with a contradictory opinion a little later; one does not say ‘Yes’, and fail to recognize what this ‘Yes’ implies. Hence knowledge could be ascribed only to someone who has thoroughly thought things out. Only someone who grasps what his beliefs imply and how his various beliefs hang together possesses knowledge.

Although the considerations that I have suggested might impel Socrates to reject *akrasia* in the broader sense (thought of as action against one’s better judgement) must remain speculative, the considerations he relies on in denying *akrasia* as action against one’s knowledge give the impression that he might not be easily persuaded into accepting the existence of *akrasia* by its latter-day

proponents. I do not mean to suggest here that he would be likely to reject the possibility of a weak-willed action on most, or even many, present-day conceptions of weakness of the will. For instance, there is no reason why he should reject the notion that a person may act in a way that stands in conflict with some of his second-order desires. In denying *akrasia*, Socrates is denying a certain picture of how human motivation operates. Thus he would be likely to deny weakness of the will as thought of by those who subscribe to the wrong picture of human motivation. In my next section I turn to the issue of what conception of the human soul and its workings he intends to reject when he rejects *akrasia*.

## XI

Socrates can deny *akrasia* without ever mentioning desire.<sup>36</sup> Citing the link between actions, on the one hand, and motivating beliefs or opinions, on the other, suffices to bring out the most general grounds on which he denies *akrasia*. However, his view of desires—and also more broadly of passions or feelings (*πάθη*)—is central to his rejection of what he presents as the cause, *ἄτιον*, of *akrasia*: the account the many give of how it comes about that one acts against one's knowledge (or belief) of what is better. Socrates' understanding of desires and passions is also central to his own full account of what actually goes on in putative cases of *akrasia*.

The common explanation of *akrasia* that he wants to reject has it that one acts akratically (weakly, as the Greek term indicates) because one is overcome by desire or passion. Recall again the view of the many: 'often, although knowledge is present in the human being, what rules him is not knowledge but something else: sometimes anger, sometimes pleasure, sometimes pain, at other times love, often fear; they [the many] think that his knowledge is dragged around by all of these just like a slave' (*Prot.* 352 B 5–C 1).

At *Prot.* 358 D 5–7 Socrates characterizes fear as *προσδοκία τις κακοῦ*, some kind of expectation of something bad. The word *προσδοκία*, translated usually as 'expectation', means something like 'anticipatory belief'—the component *-δοκία* is closely related to *δόξα*, belief or opinion.<sup>37</sup> By adding *τις* to *προσδοκία* (in Greek, *χ* *τις*

<sup>36</sup> Indeed, that is what he does. In the *Protagoras* Socrates does not mention desire until his argument against *akrasia* is completed.

<sup>37</sup> The definition of fear as *προσδοκία τις κακοῦ* is sandwiched between *συνεδόκει* at

means, roughly, some  $x$ ; a sort (kind) of  $x$ ;  $x$  of a sort), Socrates apparently wants to indicate that not every sort of expectation of, or anticipatory belief about, something bad qualifies as fear. What sort of expectation of something bad in fact qualifies as fear is at least in part connected with the sort of bad or evil (*κακόν*) that is the proper object of fear. Socrates intends this as a genuine (even if not fully spelt out) definition of fear. However he would want to spell it out, he appears to take fear as a certain, highly specific, case of taking something to be bad. Other passions would presumably be characterized too in terms of their specific way of taking something to be good or bad. I have already argued that Socrates thinks of desires as involving beliefs that something is good. If we can think of desires as passions of some kind, it would follow that desires too are ways of taking something to be good.

Socrates is sometimes said to reduce desires or feelings to mere beliefs. The assumption seems to be that, in doing so, he is leaving something out. Being in the grip of a passion can be a harrowing, wrenching, or delightful experience. How can having such an experience, people tend to ask, be a matter of holding a mere belief? I suggest that Socrates' characterization of fear need not be thought of as reductive. His view might well be the following: the very motion of the soul that constitutes the passion of fear is what it takes for us to believe that this or that thing is frightful. If this is his view, it need lose nothing from the phenomenological richness of our experiences of fear. What holding a belief amounts to depends on the sort of belief that is being held. Many evaluative beliefs have motivating force; some evaluative beliefs—those that on this view constitute desires and passions—are of such a sort that having them amounts to having experiences of a particular sort.

The belief that being *afraid* cannot consist in *taking* something to be of some sort can perhaps be traced to the view that *taking* or *considering* something is, as such, an act of intellect, and that intellect, again, is something from which the stormy movements of the soul are removed. Contrary to that line of thought, I would suggest that Socrates need not be seen as reducing desires and feelings to something else, with the richness of experience being lost in the process. He can be seen as offering an alternative analysis of what it is to desire something, or what it is to have a certain feeling.

D 4 and ἐδόκει at D 7, in a way that puts emphasis on the -δοκία part of the definition. See n. 32 above.

Similarly, in denying that a host of pleasures and pains, desires and emotions, can drag knowledge around ‘like a slave’, he need not be seen as denying the heterogeneity of states of the soul (mental states, as we might want to put it) that move us to act. That he does so is a fairly frequent misconception. What he rejects is a picture according to which passions or feelings are psychic states independent of reason. Against this, he believes that in every passion reason is in some way exercised. There is nothing in this view that would commit him to denying that the ways in which reason takes things to be good or bad are many, or even that some of these ways of taking things to be good or bad are irreducibly distinct from others.

What on Socrates’ view accounts for wrongdoing—akratic and otherwise—is not the condition of being vanquished by the forces of desire and passion. Rather, wrongdoing is in each case due to an improper functioning of reason. When passion leads us astray, what leads us astray is the incorrect valuation that our reason has adopted. It is perhaps easy to jump from this view to the position that passions as such are nothing but states in which reason has gone off track, and hence to the conclusion that we should get rid of them. (Likewise, it is easy to suppose that Socrates’ memorable rejection of the image of knowledge as a slave dragged about by a myriad of passions implies hostility to passions.) But according to the discussion of courage that follows upon Socrates’ definition of fear, in the last pages of the *Protagoras*, courage is not a state in which fear is extinguished. Far from it. Courage is a state of the soul which makes one fear those things that ought to be feared, that is to say, things that are genuinely bad. The courageous, as he puts it, do not ‘fear disgraceful fears’ (*Prot.* 360A 8–B 2). But they do fear. The courageous person’s fear—which is some kind of abhorrence of vice—would admittedly be very different from the sort of fear an ordinary soldier might feel in a battle; none the less, one should not be too quick and on account of this difference deny it the status of fear. Socrates does think that the knowledge that is virtue involves a certain peace of mind—*ἡσυχία* (*Prot.* 356E 1). We are given no ground, however, to take this kind of tranquillity to be freedom from passions—*ἀπάθεια*. On the contrary, Socrates’ discussion of courage in the *Protagoras* provides us with a picture of the virtuous person as prone to the right kind of fear.

Virtue is a condition in which one’s takings-to-be-good and takings-to-be-bad are not only correct, but are instances of know-

ledge. Those takings-to-be-good or takings-to-be-bad that constitute the passions of a virtuous person are also not just correct takings, but states of knowledge. The view here is not the more common one, that a virtuous person's passions are fully appropriate responses to the situations he encounters; rather the view is that virtue itself (in part) consists in such passions as are correct takings-to-be-good and takings-to-be-bad.

To put the same point differently: Socrates no doubt believes that someone who is not sensitive to the aspects of a situation that a virtuous person would be sensitive to does not know what there is to know about what is good and bad. However, he goes beyond this belief. He takes it that such sensitivities are themselves bits of the knowledge that is virtue. A comparison with Aristotle might make the point clearer. Socrates is often thought to differ from Aristotle in not including desiderative and emotional propensities—what Aristotle calls states of character—in virtue, making virtue instead into a mere excellence of the inert intellect that judges things correctly. On the interpretation I have offered, Socrates is precisely insisting that such propensities constitute virtue. I would locate the main difference between Socrates and Aristotle in the fact that excellent states of character for Socrates are at no remove from moral knowledge. The excellent states of character simply are states of knowledge. However close the two might lie for Aristotle (and this might be closer than some of his formulations suggest), he did want to make at the very least a notional distinction between the emotional and desiderative propensities that constitute virtue of character, on the one hand, and moral knowledge, on the other.

After Socrates gets Polus to agree to the conclusion that orators and tyrants do not do what they want to do, Polus in effect exclaims: 'As if you, Socrates, do not envy the tyrant!' (468 E 6–9). Olympiodorus, who in his commentary on the *Gorgias* judges Polus' intervention to be vulgar (*ιδιωτικόν*), entirely misses the point.<sup>38</sup> Polus' reaction is relevant, and revealing. Envy of the tyrant is not a minor lapse that can be overlooked if the person in question professes the correct beliefs. Someone who forcefully argues that doing injustice is one of the greatest evils, yet envies the tyrant, displays a soul that lacks knowledge and is very much in need of repair. One reason why so many of Socrates' interlocutors contradict themselves on the issue of virtue is that they so glaringly lack it. They think, for

<sup>38</sup> See Olymp. *In Plat. Gorg.* 95. 22 ff. Westerink, and Dodds ad loc.

instance, that what the tyrant does is disgraceful (*αἰσχρόν*), but also envy him.

Often, one of the first things Socrates wants his interlocutor to agree to is that virtue is something beautiful or fine (*καλόν*), and vice something ugly or disgraceful (*αἰσχρόν*). By concurring with this, the interlocutors commit themselves to more than they had perhaps imagined. They commit themselves, for instance, to not envying the tyrant, and to abhorring the things the tyrant does.

The virtuous person's actions express his evaluative knowledge. The evaluative judgements embodied in one's emotions and actions—the values one lives by—are of paramount importance to Socrates. A part of what in his view accounts for putative *akrasia* is precisely the fact that people are mistaken about what values they live by. If putative *akrasia* is so frequent, this is so in part because people are often mistaken about this. In addition to inconsistencies among a person's evaluative beliefs, which testify to a lack of knowledge of what is good and bad, the condition people describe as *akrasia* also involves a certain lack of self-understanding.

## XII

The many take it that sometimes, driven by a desire or emotion, we act entirely against what our reason tells us is good, better, or best. Against this, Socrates holds that our actions themselves embody judgements of value. Our reason speaks in the very passion that drives us, even if reason does not speak in a way that is consonant with our remaining opinions or judgements. We take ourselves to be fragmented where we are not. Socrates sees the human soul as one and undivided. In taking the human soul to be unitary and undivided, he is ruling out the possibility that there is an irrational or non-rational part of our souls that is capable of motivating us to act entirely on its own. But the unity of the soul he envisages has a further significance: it ties inextricably together the practical side of our nature—the desiderative, the emotional, and the volitional—with the supposedly non-practical side of us, namely the side that forms judgements and possesses knowledge.

On Socrates' view, it is an inadequate conception of reason that lies at the bottom of the belief that *akrasia* exists. An inadequate and impoverished conception of reason might also lie behind certain misunderstandings of his position. Socratic intellectualism is often

criticized as one-sided, on the ground that it does not do justice to the richness and complexity of our mental life. But on the account given here, the complexity and richness of our mental life, and of our nature, can remain untouched. Rather, Socrates' view might be that more of us goes into every state of our soul than we suspected; in some sense the whole power of the soul goes into every state of the soul. If our reason is at work in more places and in more ways than we might have thought, it should not be too surprising if it turned out to malfunction more often than expected. Specific malfunctionings of reason are also at the bottom of what people call *akrasia*.

One would expect that an intellectualist would propose an intellectual cure for an intellectual ailment. So, for instance, if virtue is knowledge, as Socrates appears to think, it might seem that all we need to do in order to instil virtue in those who lack it is instruct them about what virtue requires. But he never recommends such simple instruction; on the contrary, he insists that becoming virtuous involves much care and therapy of the soul. Reason is quite vulnerable. Susceptible to more maladies than we might have expected, it also requires more extensive and complicated care than expected. If we do not stick to the characterization of *akrasia* given in the *Protagoras*, we could concede that on Socrates' view humans are prone to a condition that might deserve to be labelled *akra-sia*. The Greek word simply indicates weakness, and Socrates does take it that weakness of reason is displayed in the episodes usually considered akratic. What he presents as powerful—as not dragged about 'like a slave'—is not reason as such, but knowledge, which is a stable overall condition of a well-functioning reason.

When Socrates describes virtue as knowledge, it is not just any kind of knowledge that he has in mind. Certain desires and feelings are part of the knowledge that is virtue. In addition, Socratic volition as discussed above is part of moral knowledge. This volition is an aspiration; it is part of an ideal of the good life. The virtuous person alone on Socrates' view does entirely what he wants to do. The virtuous person can do what he wants to do because the taking-to-be-good that his willing amounts to is itself a state of knowledge: it is an accurate grasp of what is in fact good. Being instructed on what one ought to want typically does not produce the desired wanting; this holds good for Socratic volition as much as it holds for volition as usually understood. Socrates would certainly

agree with those who think that becoming good requires that one's whole soul be turned around. What he might disagree with is what happens in the process of turning the soul around. On his view, any change in the desiderative, volitional, or emotional condition of the soul is itself a change in the condition of reason.

*University of Pittsburgh*

#### APPENDIX

##### *Gorgias 467 C 5–468 E 5*

Through his argument at *Gorg.* 467 C 5–468 E 5, Socrates gets Polus to agree, even if reluctantly, that orators and tyrants do not do what they want to do. I have argued above that Socrates' position is coherent and does not involve confusions of the sort interpreters have attributed to him. I wish now to show that the argument he uses at 467 C 5–468 E 5 to support his claim that orators and tyrants do not do what they want to do is likewise not flawed. In addition, the argument is worth looking into in its own right for at least two reasons. First, Socrates introduces here important concepts concerning human action, and second, his treatment of Polus is a paradigm of a kind of irony that he often displays in Plato's Socratic dialogues. In fact, one cannot properly assess the philosophical content of the argument at 467 C 5–468 E 5 without paying attention to the way Socrates treats Polus.

Socrates asks Polus if people who take medicines prescribed by their doctors want what they are doing (*βούλεσθαι ὅπερ ποιοῦσιν*), namely, taking the medicine and being in pain, or that for the sake of which they do this, namely, being healthy (467 C 7–10). Polus agrees that they want to be healthy. Similarly, seafarers do not want sea voyages with all the danger and trouble that these involve; what they want is wealth (467 D 1–5). Socrates then secures Polus' agreement to the claim that this is so in all cases (*περὶ πάντων*)—when a person does something for the sake of something, he does not want what he is doing (*οὐ τοῦτο βούλεται ὃ πράττει*), but the thing for the sake of which (*ἐκείνο ὃ ἕνεκα*) he is doing it (467 D 6–E 1).

The troubling clause is 'he does not want what he is doing'. The form of words chosen here might suggest that Socrates intends to assimilate all action to merely instrumental activities like taking bitter medicines. On the strict instrumentalist view ostensibly proposed here, no one ever wants what he does; the agent merely wants the beneficial result which he expects his action to have. Hence, like everyone else, orators and tyrants do not want what they are doing in killing, expropriating, and banishing. Now we know that the conclusion Socrates wants to reach is that orators and



tyrants do not do what they want to do. How is the transition made from orators' and tyrants' not wanting what they do to their not doing what they want? The instrumentalist account of action seemingly embraced at the beginning of the argument does not provide means for this transition. But in that case, why did Socrates start by drawing our attention to actions such as the taking of bitter medicine? I suggest that he is not in fact proposing to assimilate all action to merely instrumental activities like this. To see what he is up to, we have to look a bit ahead in the argument. Let me quote an important claim he makes further on:

*P*<sub>1</sub> Therefore it is because we pursue what is good [τὸ ἀγαθὸν ἄρα δι-  
ώκοντες] that we walk whenever we walk—thinking that it is better  
to walk [οἰόμενοι βέλτιον εἶναι]—and, conversely, whenever we stand  
still it is for the sake of the same thing [τοῦ αὐτοῦ ἕνεκα] that we stand  
still, what is good [τοῦ ἀγαθοῦ] . . . And don't we also kill a person,  
if we do, expel him from the city, or confiscate his property thinking  
that doing so is better for us than not doing it [οἰόμενοι ἄμεινον εἶναι  
ἡμῶν ταῦτα ποιεῖν]? . . . Hence it is for the sake of the good [ἕνεκα . . .  
τοῦ ἀγαθοῦ] that those who do all these things do them. (*Gorg.* 468 B  
1–8)

I propose to view everything that precedes the quoted passage, starting from 467 c 5, not as endorsing an instrumental account of action, but as an attempt to bring Polus to acknowledge the general picture of motivation offered in the lines just quoted. Socrates attempts to do this in two stages.

The first stage is the opening passage at 467 c 5–e 1, where the drinking of bitter medicines and navigation are discussed. Socrates introduces here the notion of an end or goal of an action—that for the sake of which (ἐκείνο οὗ ἕνεκα) an action is performed—and connects this notion with the notion of wanting (βούλεσθαι). What we want in whatever we do is that for the sake of which we do it. This point holds regardless of how the goal is related to the activity in question—whether it is separate from it, includes it, or is identical with it. The claim is simply that the proper object of wanting is the goal, whatever this goal might happen to be.

The second stage of his attempt to get Polus to subscribe to the picture of motivation painted in the lines quoted above is found at 467 e 1–468 b 1, the passage immediately following the section on bitter medicines (and immediately preceding the lines quoted). In this passage, Socrates links the goal of an action with something that is good. He gets Polus to accept a division of things into good, bad, and those that are neither good nor bad; he also gets Polus to agree that we do indifferent things for the sake of good things. That we do bad things for the sake of good things is not explicitly stated, but Polus is meant to have in mind examples like the taking of bitter

medicines—something that in itself is bad—which he had already agreed we do for the sake of something good, namely health.

Having thus provided the basic concepts of the theory of action he is trying to get Polus to subscribe to, Socrates is now ready to state a central principle of that theory. This is expressed in  $P_1$  above. What we are after—or what we pursue (*διώκειν*)—in everything we do is the goal (that-for-the-sake-of-which), and this in each case is something we take to be good. Our ‘pursuing what is good’ (at 468 B 1) is specifically glossed as thinking (note *οἰόμενοι* at B 2 and B 6) that doing what we do is ‘better’ (note *βέλτιον* at B 2 and *ἄμεινον* at B 6).

It is at this point (at 468 B 8) that Socrates invokes what he takes was agreed upon at the very outset. He says: ‘Now didn’t we agree that we want not those things that we do for the sake of something, but that for the sake of which we do them?’ (468 B 8–C 1; the reference back is to 467 C 5–E 1, especially to 467 C 5–7 and 467 D 6–E 1). I have interpreted the passage which he refers to here as establishing a conceptual connection between wanting and the goal: the goal is the proper object of wanting. If this is the point of 467 C 5–E 1, isn’t it distinctly misleading for him to insist that people do *not* want, say, to make troublesome and dangerous journeys at sea? According to my reading of the opening passage, the point Socrates makes there is in a sense trivial: *provided* that making troublesome and dangerous journeys at sea is not the goal, it is not wanted. That making such journeys is not the goal is simply built into his example. (This fits well the statement with which he ends the opening passage: ‘If someone does something for the sake of something, he does not want that which he is doing, but that for the sake of which he is doing it’, 467 D 6–E 1.) What he has in mind is clear from what he says immediately after he has reminded Polus of their previous point of agreement. Having received a positive answer (at 468 C 1) to the question ‘Now didn’t we agree that we want not those things that we do for the sake of something, but that for the sake of which we do them?’ (468 B 8–C 1), he goes on to say:

$P_2$  Hence [*ἄρα*], we do not simply [*ἀπλῶς*] want [*βουλόμεθα*] to slaughter people, expel them from the cities, or confiscate their property, just like that [*οὕτως*]; we want to do these things if they are beneficial [*ἀλλ’ ἐὰν μὲν ὠφέλιμα ἢ ταῦτα, βουλόμεθα πράττειν αὐτά*], but if they are harmful, we don’t [*βλαβερὰ δὲ ὄντα οὐ βουλόμεθα*]. (468 C 2–5)

The lines just quoted contain the second central principle of the theory of action Socrates advocates here. The claim now is that we want to do such things as killing, banishing, and confiscating, if they are beneficial; but if they are harmful, then not. So, on the view now expressed, we precisely *do* want to make dangerous journeys at sea—provided, that is, that we shall gain wealth by them. But Socrates cannot have forgotten here the point he

had made when bringing up such examples as taking bitter medicines and making dangerous journeys at sea, since at 468 B 8–C 1 he invokes what he takes to be the conclusion of that discussion, and does so precisely in order to establish the point he is now making (in  $P_2$ ). His current point is that we want to do those things the doing of which is beneficial, and we do not want to do those things the doing of which is harmful. This is a new point, quite different from what we find in  $P_1$ , and made expressly for the first time here in the lengthy argument he is presenting to Polus in support of the conclusion that orators and tyrants do not do what they want to do. And it is this new claim that he needs if he is to reach that conclusion.

So what was Socrates up to when he reminded Polus (just before stating  $P_2$ ) of the opening part of the argument, namely, their agreement concerning actions such as the taking of bitter medicines and the making of troublesome sea voyages? I suggest that he uses these examples to lead up to  $P_2$ . Even the things that people generally would not want to do—drink a bitter potion, say, or risk one’s life at sea—they will want to do, and will do, if the doing of them is beneficial. The attraction some find in adventures at sea is not a counter-example to the point he is trying to make, just as someone’s liking of bitter potions is not such a counter-example. The instances he cites are simply of things that are commonly regarded as highly undesirable. One generally does not want to do them. If this is so, the instrumental reading of the opening passage, 467 C 5–E 1, turns out to be beside the point.

Does Socrates manage to establish his claim that orators and tyrants do not do what they want to do by the argument we find at 467 C 5–468 E 5? The considerations adduced in support of his claim amount to a certain theory of action. Does the argument make it clear what this theory is and how it might be defended? Socrates, I contend, at the very least makes a move in the right direction. Admittedly, he does not say enough within the scope of his exchange with Polus fully to support  $P_2$ . In claiming that we want to do things if they are beneficial and otherwise not, he has moved from speaking of things that are merely *taken to be* good (recall *οἰόμενοι* at 468 B 2 and B 6; see also *οἰόμενος* at 468 D 3, quoted immediately below) as proper objects of wanting to speaking of things that (as a matter of fact) *are* good as proper objects of wanting. Is his transition from  $P_1$  to  $P_2$  surreptitious? The answer to this depends on whether we should take ‘surreptitious’ to mean attained by fraud or attained by stealth. Some stealth perhaps, but certainly no fraud, is involved.

Let me observe first that he makes the transition deliberately. He is not himself in a muddle. That the two passages quoted above impose different conditions on wanting is noted clearly, even if implicitly, by Socrates himself at the very end of the argument. He says:

Since we are in agreement about that, then, if someone, whether tyrant

or orator, kills someone or expels him from the city or confiscates his property because he thinks that [doing] this is better for himself [*οἰόμενος ἄμεινον εἶναι αὐτῷ*] when as a matter of fact it is worse [*τυγχάνειν δὲ ὄν κάκιον*], this person presumably does what pleases him [*οὗτος δὴπου ποιεῖ ἃ δοκεῖ αὐτῷ*], doesn't he? . . . And is he also doing what he wants [*ἃ βούλεται*] if these things are in fact bad [*εἴπερ τυγχάνει ταῦτα κακὰ ὄντα*]? Why don't you answer? [POLUS:] All right, I think that he isn't doing what he wants. (468 D 1–7).

Socrates here deliberately reserves the term 'wanting' for an attitude that hits upon what is in fact good, knowing of course that most people do not use the term in this way. Is he warranted in thus construing the term? I would like to argue that, dialectically speaking, he is. At 467 E 1–468 B 1 he got Polus to adopt the objective point of view: there Polus accepted the division of things into those that are good, those that are bad, and those that are neither good nor bad. The talk there was of things that actually are good, bad, or neutral, not merely of things that are considered to be so. Polus further agreed to the claim that people act for the sake of (actually) good things. It is this agreement that Socrates exploits further down, at 468 c 5–7, while moving towards *P*<sub>2</sub>. He says: 'For we want things that are good [*τὰ γὰρ ἀγαθὰ βουλόμεθα*]—as you say [*ὡς φῆς σύ*]—whereas we don't want those that are neither good nor bad, nor those that are bad.'

Polus is free to object here; he could interject that he had previously only meant to say that we go for—and want—things we take to be good, not things that are in fact good. But he does not object. Not only is he free to protest; Socrates repeatedly nudges and prods him to do so. When Socrates says 'For we want the things that are good, as you say . . .', he is clearly warning Polus that the conclusion has been reached using something Polus himself had previously said or agreed to. This is undoubtedly an invitation to Polus to reflect on what he had previously agreed to and why. What is especially intended to serve as a prod is the palpable irony in that 'as you say'.<sup>39</sup> When Socrates then pointedly asks whether Polus believes that what Socrates is saying is true, Polus, again, does not stir. Since he makes

<sup>39</sup> Polus has not expressly said this. Since he has gone along with Socrates, Socrates can count him as having said so, but in that case Polus must be saying something he does not mean. He clearly does not understand the full force of the view he assents to, otherwise he would not be granting what in a moment will force him to embrace what he had not too long ago described as 'outrageous' and 'monstrous'. There is a hint of irony in the question Socrates will ask immediately after making the assertion (at 468 c 5–7) that we want things that are good: 'Do you [Polus] think that what I am saying is true [*ἀληθῆ σοι δοκῶ λέγειν*]?' (468 c 8). Since 'what I am saying' refers to Socrates' allegation about what Polus is 'saying' ('For we want the things that are good, as you [Polus] say, *ὡς φῆς σύ* . . .', 468 c 5–6), Socrates is in effect confronting Polus. He is questioning whether Polus has any idea not only what Socrates is saying and believing, but also what he, Polus, is saying and believing.

no objection, his prior agreements now commit him to saying that we want things that as a matter of fact are good.

Is Socrates deliberately confusing Polus? Given that Polus does not quite grasp what he is up to, should he move so nonchalantly from  $P_1$  to  $P_2$ ? It seems to me that by examining Polus as he does, he precisely intends to bring Polus' confusion into the open. He is suggesting to Polus that he has to rethink the whole issue since he does not understand what he is saying. In thus underscoring the need to think more carefully about the transition from  $P_1$  to  $P_2$ , Socrates can hardly be muddling the issue. We must also bear in mind, as we wonder whether he might be confusing Polus, that he really believes what he describes Polus as saying. He decidedly believes that we 'want good things', namely, that we want things that are [as a matter of fact] good. It is therefore reasonable to think that when he says so, he is making a bona fide suggestion to Polus as to what he, Polus, ought to be saying. Finally, the shocking nature of the conclusion, that orators and tyrants do not do what they want to do—the conclusion which, as I have argued, Socrates embraces in all seriousness—is in itself a deterrent against muddles.

Socrates has philosophical reasons for maintaining that we want—in some legitimate sense of this word—what as a matter of fact is good, even if he has not fully set these reasons out in his exchange with Polus at 467 c 5–468 e 5. If in this argument he slips into the guise of the orator and tyrant, forcing what he takes to be true upon the unwilling Polus, we should perhaps be led to suspect that Polus lacks the sort of grasp of the matters discussed that would enable him to follow Socrates willingly.

OXFORD STUDIES  
IN ANCIENT  
PHILOSOPHY

EDITOR: DAVID SEDLEY

VOLUME XIX

WINTER 2000

OXFORD  
UNIVERSITY PRESS