

for a volume on Plato and the Divided Self,  
ed. Rachel Barney, Tad Brennan, and Charles Brittain  
word count: 9693 (9000)

## **The Unity of the Soul in Plato's Republic**

Eric Brown

Department of Philosophy  
Washington University in St. Louis  
eabrown@wustl.edu

### **1. Introduction**

In Book Four of Plato's Republic, Socrates divides the soul.<sup>1</sup> He argues that because one thing cannot do or undergo opposites in the same respect, in the same relation, and at the same time, some cases of psychological conflict can only be explained by supposing that the soul is actually more than one thing. But Socrates does not seem to deny that the soul is a unity. Rather, he seems to hold that the soul, like so many entities, is both a many and a one. He refers to the divisions as parts of the soul,<sup>2</sup> and even at the end of Book Four, when he ascribes actions and attitudes to a part of the soul, he also ascribes them to the whole soul (esp. 439a-b, 439d). More generally, throughout the Republic, he treats each agent he discusses as a single locus of responsibility, including even the most psychologically wretched agent, who suffers from tyranny within (572b-580a, esp. 577d13-

---

I thank the organizers of the conference in Toronto and the editors for including me. I also thank Jennifer Whiting for the paper on which I commented in Toronto and Devin Henry for the opportunity to present a successor to those comments at a conference at the University of Western Ontario in October 2008. Finally, I thank those who offered helpful comments on that version, especially Verity Harte and John Mouracade in Ontario, and Clerk Shaw, Rachel Singpurwalla, Eric Wiland, and the editors by correspondence.

<sup>1</sup> For the Republic, I cite the text of Slings 2003, and translations are mine, though I have borrowed heavily from Grube's translation, as revised by Reeve, which appears in Cooper 1997.

<sup>2</sup> Following Adam 1905, ad 435b, we may note that when Socrates is dividing the soul, the divisions are called εἶδη, γένη, or μέρη: εἶδη in 435c1, 435c5, 435e1, 439e1, γένη in 441c6, 443d3, μέρη in 442b10, 442c4, 444b3. Socrates might not be fully committed to divisions in 435c1, 435c5, and 435e1, but after he has articulated the principle of non-opposition (436b9-10), he uses all three words to refer to the same divisions. He also refers to the whole soul in 436b2, strongly implies the idea of parts in 439b-d, and returns to talking of μέρη at 577d4, 581a6, 583a1, and 586e5. See also 611b5-7: the embodied soul is a composite (σύνθετον).

e1). It seems, then, that although few souls enjoy the unity that virtue earns,<sup>3</sup> every human soul possesses an unearned unity.<sup>4</sup>

How should we understand this unearned unity of the soul? That is, given that the psychological divisions Socrates argues for in Republic IV are parts, what explains how they are parts of a single whole? This question has received little attention.<sup>5</sup> My first aim here is to establish that it is an important puzzle, and my second is to venture an answer.

## 2. The importance of the soul's unity

My question is in part mereological. Plato seems to assume that my rational, spirited, and appetitive elements constitute a complex unity, and that my rational, spirited, and appetitive elements together with your spirited element do not constitute a complex unity. What explains which psychological elements can compose a soul and which cannot?

But the interest of my question is not exclusively mereological. First, if the human soul is not a complex unity but is three distinct entities, it would seem difficult to explain what it feels like to be human. As Descartes puts it, "When I consider the mind, or myself insofar as I am merely a thinking thing, I am unable to distinguish any parts within myself; I understand myself to be something quite single and complete."<sup>6</sup> What accounts for this felt unity of consciousness if the soul is really three distinct entities?<sup>7</sup>

Second, psychological states and activities are just some of the vital activities for which the soul is supposed to be responsible. Again, if the soul is not a complex unity but is three distinct entities, what explains the apparently unified vitality in a human being? Plato offers no hint that the psychological functions of the soul are entirely separate from the other vital functions, nor should he. In Book One, Socrates argues that a good soul is

<sup>3</sup> For the unity that virtue brings, see especially 443d-e. That this condition is rare follows from the following three points: virtue requires knowledge (442c), knowledge requires philosophy (474b-480a), and philosophy is rare (491a-b).

<sup>4</sup> There is one prominent exception (554d), and I consider it in the last section below.

<sup>5</sup> Aristotle raises this question as an objection (De an. I 5 411b5-7), but modern scholars say strikingly little about it, though they sometimes note the problem in passing (e.g., Archer-Hind 1882, 124, and Lorenz 2006, 38-40). Scholars have been more interested in the earned unity that virtue brings (e.g., Hsu 2007) or the apparent tension between the tripartite soul of the Republic, Phaedrus, and Timaeus and the simple soul of the Phaedo and other dialogues (e.g., Archer-Hind 1882). Some scholars offer a principled excuse: they maintain that the soul is a simple unity with three different kinds of psychological states and activities that are mere conceptual parts (e.g., Shields 2001). I argue against this reading in §3 below.

<sup>6</sup> Meditation Six, in Cottingham, Stoothoff, and Murdoch 1984, 2:59, rendering the text in Adam and Tannery 1964-1976, 7:86.

<sup>7</sup> See Bobonich 2002, 254.

simply one that performs its function well, that is, that lives well (352d-354a), and this unifies the vital activities, moral-psychological and otherwise, into one function.<sup>8</sup> How can there be a unified function without a unified subject to perform it?

Third, we also treat each other as single agents when we assign praise and blame, rewards and punishments. People do not, at least typically, exonerate the Eric Brown who pursues learning when they criticize the Eric Brown who overeats. Rather, they hold one and the same agent responsible whether he pursues wisdom, reacts angrily, or seeks a yummy dinner. But, again, what justifies this practice if each of us is really three distinct agents?

The Republic does not much consider the unity of consciousness or vital activities. Socrates does presuppose that some information passes from sensory organs to more than one part of the soul and that each part of the soul has some awareness of what the others want; so he presupposes some shared mental content among the parts of the soul. But he leaves considerable room for speculation about how best to make sense of these presuppositions.

The Republic says much more about moral responsibility, from Cephalus' initial musings to Er's final story. These remarks make it difficult to construe the person—the single locus of moral responsibility—as anything other than the composite soul. They disqualify the complex of soul and body; according to the Myth of Er, the disembodied soul in the afterlife is responsible for what it did while embodied (614b-616b).<sup>9</sup> They also make it difficult to see how we might hold the rational part alone responsible. This is a tempting approach, since Socrates suggests that only the rational part of the soul survives separation from the body (611b-612a) and Er's myth plainly says that the soul in the afterlife bears responsibility for the person's past embodied existence (614b-616b) and for the next (617d-e). Although Socrates is not fully committed to the claim that the post-mortem soul is just

---

<sup>8</sup> Some scholars have thought that Socrates fails to unify the function. Shorey (1903, 41), for example, finds two functions here, one to provide life and the other to lead the moral life. But I see no reason to attribute this confusion to Socrates. The single function of the soul is to live, but living consists in a complicated set of heterogeneous activities, including moral-psychological ones, and the soul performs these activities (that is, lives) either well or poorly. Compare the cobbler: his function is to make shoes, but making shoes comprises heterogeneous activities, and the cobbler performs them (that is, makes shoes) either well or poorly.

<sup>9</sup> There are questions about the unity of the soul-body complex, as well. Socrates discusses an earned unity for this complex (462c-e), but not unearned unity. Presumably, the same materials are needed to address this and the unity of vital activities. A rich account of sense-perception might go some way to performing these tasks and to explaining the unity of consciousness. I leave these questions for the most part aside.

the rational part (612a3-6), he evidently considers it plausible,<sup>10</sup> and so he cannot think it inappropriate to locate full responsibility in the rational part. Unfortunately, what he says about the separate parts of an embodied soul do make it seem inappropriate. Many actions for which an embodied human being is responsible are motivated by appetitive or spirited desire, and holding the rational part responsible for such actions seems to blame an innocent bystander or, worse, a victim.<sup>11</sup>

What the Republic says about moral responsibility would make more sense if the whole complex soul were the locus of moral responsibility. This assumption can even accommodate the possibility that only the rational part of the soul survives death. The whole complex soul would be the locus of moral responsibility while it is embodied, and the whole remaining (incomposite) soul could be the locus of moral responsibility in the afterlife. This would be appropriate in general because these are the same soul: in connection with the body, the soul develops parts that it lacks without a body, but it is nonetheless the same entity with or without the body. This would also be appropriate in detail if (as I will suggest) the account that explains the complex unity of the embodied soul also explains how each of the soul's parts is at least indirectly responsible for everything that any part of the soul does.

So for several reasons Plato needs an account of the embodied soul's unearned unity.<sup>12</sup> (From this point on, unless otherwise noted, I am concerned with the unity of the embodied soul, and I, like Socrates in the Republic, leave unresolved the question of the soul's postmortem constitution.) In the next section, I argue that this requires an account of the soul's complex unity, because Socrates divides the soul into three distinct sources of psychological states and activities, and I show how Socrates' argument explicitly leaves an opening for the soul to be a complex unity. Then I consider possible solutions to the puzzle of unearned unity. I argue that Plato has an account sufficient to explain why the whole soul is a locus of moral responsibility. This account only begins to explain the unity of consciousness and the organization of vital activities—one needs to go beyond the Republic

<sup>10</sup> The suggestion of 611b-612a fits neatly with the "Affinity Argument" of Phaedo 78b-84b (see Shields 2001, 141-144) and with the Timaeus (see Archer-Hind 1882). See also Rep. 518d9-519a1. But the Phaedrus suggests, to the contrary, that the disembodied soul is complex (pace Archer-Hind 1882, 127).

<sup>11</sup> Perhaps one could defend the propriety of locating moral responsibility in the rational element in a soul that lacks complex unity. My point here is that some defense is needed and that without it, there is a significant cost to ascribing to the Republic the view that all psychic unity is earned.

<sup>12</sup> If he could make sense of all the coordinated vital activities without a unified soul, Plato could instead of accounting for the soul's unearned unity offer an error theory to explain why we systematically attribute praise and blame as if humans were unified persons when they are in fact not, and to explain why we experience an illusory unity of consciousness. But he shows no interest in such a theory.

to offer more—but it offers at least some sensitivity to the need to explain these phenomena.

### 3. The problem of the soul's unity

The puzzle about how Plato might account for the unearned unity of the soul stems from Socrates' argument to divide it. Indeed, it might seem that Socrates' argument commits him to denying unearned unity of the soul.<sup>13</sup>

Socrates introduces the argument by suggesting that there might be three different things in us "by which" we do three different kinds of things. He says,

But this is hard, [to determine] whether we do each of these things by this same thing or by three things, one thing by one and another by another. Do we learn by one thing, get angry by another of the things in us, and again by a third desire the pleasures concerning nutrition and procreation and all the things akin to these, or do we act in each of these cases by the whole soul, whenever we should have an impulse to act? This is hard to determine in a way worthy of our argument.<sup>14,15</sup>

Socrates and Glaucon then agree to "try to determine the things in us, whether they are same as each other or different."<sup>16</sup>

This opening raises the problem of the soul's unity without taking a stand. It raises the problem because it plainly introduces the possibility that there are three distinct sources of psychological states and activities in us. Socrates is not here merely categorizing the

<sup>13</sup> See Grote 1888, 2:386n ("three souls"), and Adam 1905, *ad* 435a, qualified by Adam 1905, *ad* 439b. Bobonich (2002, 219-257) has recently argued forcefully for this conclusion, though he also allows himself to talk about the composite soul without explaining the composition or suggesting that the composite allows for the unity of the person. His cogent denial (234 with 531n27) that there is any ultimate subject above and beyond the three parts of the soul does not address the question of how the three parts might be unified.

<sup>14</sup> 436a8-b4: Τόδε δὲ ἤδη χαλεπὸν, εἰ τῷ αὐτῷ τούτῳ ἕκαστα πράττομεν ἢ τρισὶν οὖσιν ἄλλο ἄλλω: μανθάνομεν μὲν ἑτέρῳ, θυμούμεθα δὲ ἄλλῳ τῶν ἐν ἡμῖν, ἐπιθυμοῦμεν δ' αὖ τρίτῳ τινὶ τῶν περὶ τὴν τροφήν τε καὶ γέννησιν ἡδονῶν καὶ ὅσα τούτων ἀδελφά, ἢ ὅλη τῇ ψυχῇ καθ' ἕκαστον αὐτῶν πράττομεν, ὅταν ὀρμήσωμεν. ταῦτ' ἔσται τὰ χαλεπὰ διορίσασθαι ἀξίως λόγου.

<sup>15</sup> Socrates says "in a way worthy of our argument" not to say that their standards are especially high but to remind his interlocutors that they are not as high as they might be: they are not traveling the "longer road" that leads to a fuller answer by means of the Good (435c9-d3 with 504b1-7). Socrates continues to remind his interlocutors of these lower standards when he notes that the principle of non-opposition is merely a hypothesis by their current argument (437a3-8). I take these to be related because I assume that the Good is supposed to be the unhypothetical first principle (see 533a-534c with 505a, 506d-509d, and 509d-511d). Hence, I take him to be saying that the longer argument would be required to render their conclusion in Book Four unhypothetical. If that is right, Socrates does not suggest that the longer road would necessarily lead to a substantially different conclusion.

<sup>16</sup> 436b6-7: Ὡςδε τοίνυν ἐπιχειρῶμεν αὐτὰ ὀρίσασθαι, εἴτε τὰ αὐτὰ ἀλλήλοις εἴτε ἕτερά ἐστι.

kinds of psychological states and activities; he has already done that easy task (435d-436a). He now raises the hard question of whether we should attribute these different kinds to different causal sources in us.<sup>17</sup> But this question does not take a stand on the soul's unity. Socrates and Glaucon agree to determine whether learning, anger, and desires for bodily pleasures stem from the same source or three different sources. But they have not considered the ramifications of their being from different sources. For all they say, the different sources might be parts of a single soul, as they are all sources "in us," or they might fail to be parts of a single soul, despite their being "in us."<sup>18</sup>

To answer his question, Socrates introduces the principle of non-opposition.<sup>19</sup> He first states it in very general terms that make it applicable to any entity whatsoever: "It is clear that the same thing will not be willing to do or undergo opposites [1] in the same respect and [2] in relation to the same thing [3] at the same time."<sup>20</sup> But he immediately draws out the consequence of the principle for the "things in us" that he has mentioned: "The result is that if we should somehow discover these things [viz., opposites] happening in them [viz., the things in us by which we learn, feel anger, and desire the pleasures of nutrition and generation and such], we will know that they were not the same thing, but many."<sup>21</sup> In line with the agreement he has reached with Glaucon, Socrates seems to presuppose that either we act by one thing or we act by many things. So either he is continuing to set aside the ramifications of acting from multiple psychological sources—perhaps the multiple psychological sources are parts of a single soul and perhaps they are

<sup>17</sup> This tells against the reading of those who, like Shields (2001), construe the parts of the soul as merely conceptual parts. See also n. 00 below.

<sup>18</sup> One might take Socrates' question to favor the former possibility, since he mentions the possibility that we act "with the whole soul [ὅλη τῆ ψυχῇ]" (436b2). I am not inclined to put much weight on this, since his contrast between doing one of these actions with the whole soul and doing it with just one of three psychological sources actually sets aside the possibility of complex unity that he ends up articulating, according to which our whole soul does the action insofar as just one of three psychological sources does it.

<sup>19</sup> Many other labels have been used. I think this one best hints at the close connection between Plato's principle and what is most commonly (although not unanimously) called the "principle of non-contradiction" (for which see Aristotle, Metaph. G3 1005b19-20).

<sup>20</sup> 436b9-10 (cf. 436e7-437a1): Δῆλον ὅτι ταῦτόν τάναντία ποιεῖν ἢ πάσχειν [1] κατὰ ταῦτόν γε καὶ [2] πρὸς ταῦτόν οὐκ ἐθέλησει [3] ἅμα...

<sup>21</sup> 436b10-c2: ...ὥστε ἂν που εὐρίσκωμεν ἐν αὐτοῖς ταῦτα γιγνόμενα, εἰσόμεθα ὅτι οὐ ταῦτόν ἦν ἀλλὰ πλείω. My supplement takes αὐτοῖς to refer to the same things that the αὐτὰ of 436b6-7 refer to, namely the things in us discussed in 436a8-b4. Unfortunately, the Grube-Reeve translation (in Cooper 1997) is misleadingly careless here. It reads, "So, if we ever find this happening in the soul, we'll know that we aren't dealing with one thing but many." The careless "in the soul" for ἐν αὐτοῖς apparently leads to the complete fabrication of "we aren't dealing with one thing" (since it would be nonsensical to say "if we ever find this happening in the soul, we'll know that it is not the same thing but many").

not—or he is assuming without argument that the only way the soul can be unified is for its apparently multiple sources of activity to be the same. He has explicitly said nothing to shut the door on complex unity.

But what has he implied? The principle of non-opposition allows that a single subject can undergo opposites (call them phi and psi) only when at least one of three conditions is met. A single subject can phi at one time and psi at another; a single subject can phi in relation to A and psi in relation to B; and a single subject can phi in one respect and psi in another. But can a single subject phi in one part and psi in another? Or has Socrates already ruled this out?

According to what I will call the standard reading, a single subject can phi in one part and psi in another, because to phi in one part and psi in another is just to phi in one respect and psi in another.<sup>22</sup> On this reading, the principle of non-opposition allows the soul to be a whole that comprises parts that are distinct sources of psychological states and activities.

Unfortunately, it is not hard to make trouble for the standard reading.<sup>23</sup> Socrates immediately clarifies the principle of non-opposition by considering two potential counter-examples to it. First, he says,

If someone were to say of a human being who is standing still but moving his hands and head that one and the same person is standing still and moving at the same time, we would not, I think, think that we should say this, but rather that one thing of him is standing still and another [thing of him] is moving.<sup>24</sup>

Socrates does not say that the single human being stands still in one respect (by some part of him) and move in another (by another part of him). He insists, instead, that one thing stands still and another thing moves.<sup>25</sup> If we assume that the one thing standing still and the other thing moving are parts of the human being, then we can generalize Socrates' point as follows: if two parts of a single thing are subjects of opposites, then we should not

---

<sup>22</sup> See, for examples, Irwin 1995, 204; Price 1995, 40-41; and Shields 2001, 145. The Grube-Reeve translation (in Cooper 1997) unfortunately forces this reading on its readers. They render the principle of non-opposition's *κατὰ ταὐτόν* (436b9: "in the same respect") as "in the same part of itself," and they use the same rendering for *κατὰ τὸ αὐτό* at 436c6.

<sup>23</sup> I draw on Bobonich 2002 and Lorenz 2006 for the "trouble-making" reading. Both believe that Socrates establishes three distinct "ultimate" or "proper" subjects of psychological states and activities, though they disagree about the nature of the subjects, their oppositions, and the implications of the division.

<sup>24</sup> 436c10-d1: εἰ γὰρ τις λέγοι ἄνθρωπον ἑστηκότα, κινουῖντα δὲ τὰς χεῖρας τε καὶ τὴν κεφαλὴν, ὅτι ὁ αὐτὸς ἑστηκὲ τε καὶ κινεῖται ἅμα, οὐκ ἄν, οἶμαι, ἀξιοῖμεν οὕτω λέγειν δεῖν, ἀλλ' ὅτι τὸ μὲν τι αὐτοῦ ἑστηκε, τὸ δὲ κινεῖται.

<sup>25</sup> It is possible to take the *τὸ μὲν* and *τὸ δέ* to be accusatives of respect and not nominatives and to translate this sentence differently; see n. 00 below.

say that the single thing undergoes opposites in different respects; rather, we should say, instead, that a many undergoes opposites.

The trouble-making reading leaves unclear what it would mean to say that something phis in one respect and psis in another. If different "respects" do not cover different parts, then what do they cover? The trouble-makers take Socrates' treatment of the second potential counter-example to answer this question:

And if the person who says these things [about the man standing still and moving] became even more amusing and was sophisticated enough to say that whole spinning tops stand still and move at the same time when they have fixed their center in the same spot and revolve, or that anything else that stays in one spot while moving in a circle does the same thing, we would not agree, because such things belong to things that stand still and move not in the same respect of themselves. Rather, we would say that they have an axis and a circumference in them and that with respect to the axis they stand still, since they do not incline to either side, while with respect to the circumference they move in a circle. And whenever one does incline their axis to the left or the right, or to the front or back, while it is spinning, we would say that it is in no way standing still.<sup>26</sup>

This "sophisticated" counterexample is supposed to prevent Socrates from saying that some parts of the top move while others stand still. The whole top is supposed to be moving and at rest (436d5). Presumably, the whole top is moving insofar as it is revolving and at rest insofar as it does not incline now in this direction and now in that but continues to occupy the same volume of space (since the axis remains fixed). Socrates accordingly explains how this top does not violate the principle of non-opposition by insisting that the top moves in one respect and stays at rest in another.<sup>27</sup> So understood, difference in respects does not cover difference in parts, and difference in parts is not one of the excusing conditions of the principle of non-opposition.

Although this makes trouble for the standard reading, it does not establish that Socrates has ruled out the possibility that the soul is a complex unity. In fact, he goes on to make room for something very close to the standard reading.

<sup>26</sup> 436d4-e5: Οὐκοῦν καὶ εἰ ἔτι μᾶλλον χαριεντίζοιτο ὁ ταῦτα λέγων, κομψευόμενος ὡς οἱ γε στρόβιλοι ὅλοι ἐστᾶσι τε ἅμα καὶ κινοῦνται, ὅταν ἐν τῷ αὐτῷ πήξαντες τὸ κέντρον περιφέρονται, ἢ καὶ ἄλλο τι κύκλῳ περιῶν ἐν τῇ αὐτῇ ἔδρᾳ τοῦτο δοῦν, οὐκ ἂν ἀποδεχοίμεθα, ὡς οὐ κατὰ ταῦτα ἑαυτῶν τὰ τοιαῦτα τότε μενότων τε καὶ φερομένων, ἀλλὰ φαίμεν ἂν ἔχειν αὐτὰ εὐθύ τε καὶ περιφερῆς ἐν αὐτοῖς, καὶ κατὰ μὲν τὸ εὐθὺ ἐστάναι, οὐδαμῆ γὰρ ἀποκλίνειν, κατὰ δὲ τὸ περιφερῆς κύκλῳ κινεῖσθαι, ὅταν δὲ τὴν εὐθυωρίαν ἢ εἰς δεξιὰν ἢ εἰς ἀριστερὰν ἢ εἰς τὸ πρόσθεν ἢ εἰς τὸ ὀπίσθεν ἐγκλίνη ἅμα περιφερόμενον, τότε οὐδαμῆ ἐστάναι.

<sup>27</sup> On the top, see Bobonich 2002, 228-231, and Lorenz 2006, 23-24. Many scholars are tempted to think that some part(s) of the top move while others are at rest, but this reading fails to take the case as a more difficult challenge than the first potential counter-example (see esp. 436d5).

Before he divides reason from appetite, Socrates offers a second specification of the general principle of non-opposition, tailored for the case of psychological conflict:

Therefore, if sometime something draws the soul back when it is thirsting, wouldn't that be something in it that differs from the thing that thirsts and drives it like a beast to drink? For, as we say, the same thing would not do opposites [1] by the same thing of itself [2] about the same thing [3] at the same time.<sup>28</sup>

This restatement differs from the original in three ways. The first is trivial: this specification leaves out the possibility of "undergoing [πάσχειν]" opposites. The second is potentially significant, but intelligible: "in relation to the same thing [πρὸς ταὐτόν]" has become "about the same thing [περὶ τὸ αὐτό]," perhaps because Socrates assumes that mental experiences are related to other things specifically because they are about other things. (Mental experiences, we might say, are characterized by intentionality.) Third, in place of "in the same respect [κατὰ ταὐτόν]" Socrates has now said "by the same thing of itself [τῷ αὐτῷ ἑαυτοῦ]." This is of some consequence.

Recall the human who stands still while moving his arms and head: "one thing of him is standing still and another [thing of him] is moving" (436d1). The trouble-makers point out that because Socrates does not say that the human moves in one respect and stands still in another, the case does not satisfy any of the three excusing conditions of the principle of non-opposition. It is not an allowable case of one thing being a subject of opposites but is instead a case of a plural subject. But the way Socrates characterizes this case does satisfy one of the three excusing conditions of the restated principle of non-opposition. As the principle is restated, this is an allowable case of one subject of opposites: the human moves "by one thing of him" and stands still "by another thing of him."

Why has Socrates restated the principle to such effect? There are two attractive explanations. First, we might say that the restated principle is just a more specific version of the original principle, tailored to the case of psychological conflict. On this view, something close to the standard reading was right all along: to phi with one part of oneself while psi-ing with another is one way of phi-ing in one respect and psi-ing in another. (It is not the only way because the case of the top highlights another, though it might be the only way relevant to psychological conflict.) On this view, the trouble-makers misunderstood Socrates' characterization of the human who stands still and moves. They took him to be

<sup>28</sup> 439b3-6: οὐκοῦν εἴ ποτέ τι αὐτὴν ἀνθέλλκει διψῶσαν, ἕτερον ἂν τι ἐν αὐτῇ εἴη αὐτοῦ τοῦ διψῶντος καὶ ἄγοντος ὥσπερ θηρίου ἐπὶ τὸ πιεῖν; οὐ γὰρ δὴ, φάμεν, τό γε αὐτὸ [1] τῷ αὐτῷ ἑαυτοῦ [2] περὶ τὸ αὐτὸ [3] ἀμα τάναντία πράττοι.

saying only that one thing of him stands still and another moves, whereas he is also saying that the human moves in one respect and stands still in another.<sup>29</sup>

Alternatively, we might think that although the trouble-makers were right about the original principle of non-opposition and the human who stands still and moves, the restated principle of non-opposition changes matters considerably. On this view, the original principle emphasizes the plurality required of an apparent subject of simultaneous opposites: we do not say that the human moves in one respect and stands still in another. But no one should infer from this that the human body is not a whole comprising parts some of which can move while others are at rest. Although the principle of non-opposition shows that there are multiple subjects of opposing actions, it does not show that those multiple subjects cannot be parts of a complex whole. The restated principle makes this allowance more explicit. With the restated principle, we can say that the human moves by one part and stands still by another. By putting the principle in both ways, Socrates emphasizes both the plurality and the unity of the human body and, by implication, the human soul.<sup>30</sup>

Thus, whether the revised standard reading of the original principle of non-opposition is correct or not, Socrates explicitly allows the soul to be a single thing comprising multiple parts and not merely a many. With his restated principle of non-opposition—whether it is merely a specification as the revised standard reading has it or is a more significant shift as the trouble-makers must think—he has made room for his original formulation of the problem: does a person, a soul, do diverse activities by diverse elements of itself, or does

---

<sup>29</sup> In support of this, one might make two points. First, one might point to the question Socrates asks just before he raises the first potential counter-example (436c6-7): "Can the same thing stand still and move at the same time in the same respect?" ἐστάναι, εἶπον, καὶ κινεῖσθαι τὸ αὐτὸ ἅμα κατὰ τὸ αὐτὸ ἄρα δυνατόν; Perhaps this question is supposed to focus Glaucon's attention on ways in which a thing can simultaneously stand still and move in different respects, and then the two potential counter-examples highlight two different ways in which a thing can stand still in one respect and move in another. Second, one might question the trouble-makers' construal of 436d1. It is possible to take the τὸ μὲν and τὸ δέ to be accusatives of respect and not nominatives, in which case Socrates reports that "we should say... rather that he stands still in one respect and moves in another." To reject this reading, Lorenz (2006, 23n9) appeals to the parallel case of the archer at 439b8-c1, where Socrates says that one hand moves one way and the other another, and not that archer moves one way with respect to one hand and another way with respect to the other. But the archer enters after Socrates has restated the principle of non-opposition and traded the more general "in the same respect" clause for the more specific "by the same thing of itself" clause. That would explain the general talk of respects for the human standing still and moving and the more specific talk of parts for archer.

<sup>30</sup> The trouble-makers do not always take full stock of how the restatement changes the principle of non-opposition. Adam (1905, ad 439b) says only that by the restatement "Plato betrays a sense of the unity of the soul." Bobonich (2002, 232-233 with 530n22) denies that anything significant has changed and appeals unpersuasively to the position of γε in 439b5. But Lorenz (2006, 27) gets it right: "It is exactly this kind of reformulation that is needed to allow Socrates to say that one and the same soul can (and all too frequently does) do opposites in the same respect, in relation to the same thing, and at the same time—just not with the same part of itself."

she do each by the whole soul? Socrates plainly opts for the first possibility, but he stresses that the diverse sources of an agent's actions are elements in or of the agent's single soul. He says that sometimes when "there are thirsty people who are unwilling to drink... there is something in their soul bidding them to drink, and there is something in [their soul], forbidding them to do so, something that is different and that overrides the thing that bids."<sup>31</sup> He quickly labels "that with which the soul calculates the rational thing of the soul and that with which it lusts, hungers, thirsts, and gets excited by other appetites the irrational and appetitive thing, companion of certain indulgences and pleasures."<sup>32</sup> Socrates takes himself to have shown not merely that there are multiple psychological sources within us but that "these two kinds are distinguished in the soul."<sup>33</sup>

This leaves our problem.<sup>34</sup> Although Socrates has left room for the soul to be a complex whole comprising multiple parts, he has also made it clear that the parts are separate sources of psychological activity. The rest of the Republic only makes this clearer, as each of the separate sources of psychological activity is characterized as agent-like.<sup>35</sup>

<sup>31</sup> 439c3-8: Πότερον δὴ φώμεν τινὰς ἔστιν ὅτε διψῶντας οὐκ ἐθέλειν πιεῖν; Καὶ μάλα γ', ἔφη, πολλοὺς καὶ πολλαίς. Τί οὖν, ἔφην ἐγώ, φαίη τις ἄν τοῦτων πέρι; οὐκ ἐνεῖναι μὲν ἐν τῇ ψυχῇ αὐτῶν τὸ κελεῦσθαι, ἐνεῖναι δὲ τὸ κωλύειν πιεῖν, ἄλλο δὲν καὶ κρατοῦν τοῦ κελεύοντος;

<sup>32</sup> 439d4-8, translating d5-8: Οὐ δὴ ἀλόγως, ἦν δ' ἐγώ, ἀξιόσομεν αὐτὰ διττὰ τε καὶ ἕτερα ἀλλήλων εἶναι, τὸ μὲν ὧς λογίζεται λογιστικὸν προσαγορεύοντες τῆς ψυχῆς, τὸ δὲ ὧς ἐρεῖ τε καὶ πεινῆ καὶ διψῆ καὶ περὶ τὰς ἄλλας ἐπιθυμίας ἐπιόηται ἀλόγιστόν τε καὶ ἐπιθυμητικόν, πληρώσεών τινων καὶ ἡδονῶν ἐταῖρον.

<sup>33</sup> 439e1-2: Ταῦτα μὲν τοίνυν, ἦν δ' ἐγώ, δύο ἡμῖν ὠρισθῶ εἶδη ἐν ψυχῇ ἐνόητα.

<sup>34</sup> Among others. I set aside here the problems concerning how many parts Socrates' argument establishes. One of these problems concerns whether he is entitled to think that there is just one appetitive part: the heterogeneity of its desires suggests that there can be conflicting appetitive desires. (Compare the strategies offered by Reeve 1988, 124-131; Irwin 1995, 203-217; Price 1995, 46-48; and Lorenz 2006, 13-52.) Another concerns whether he is entitled to think that there is a spirited part in addition to the appetitive and rational parts. (See Cooper 1984.) A third concerns whether there might not be some part(s) beyond reason, spirit, and appetite. (See 443e and Kamtekar 2007.) My problem concerns how the parts constitute a whole soul, and it applies however many parts there are and however they might be subdivided.

<sup>35</sup> This has been widely recognized, at least since Moline 1978, although there are debates about exactly what characteristics should be imputed to the different parts, and especially appetite (cf. Bobonich 2002 and Lorenz 2006). Anyone tempted by the revised standard reading of the partitioning argument and the claim that the parts of the soul are merely conceptual must somehow explain away Socrates' descriptions of the soul-parts as distinct agent-like sources of psychological states and activities. But there is no good reason not to take these descriptions at face value. Shields (2001) claims otherwise, arguing that if soul-parts were more than conceptual, then the soul could not be essentially simple and thus could not be immortal. But in Republic X (611b-612a), Plato entertains the possibility of a complex soul being immortal and the possibility of a simple disembodied soul being necessarily complex when embodied. Shields (2001, 147-151) offers two additional arguments. First, because the principle of non-opposition can be used to establish a distinction among conceptual parts (or at least abstract parts, which he mistakenly assumes must be conceptual parts), it cannot be used to generate organic or aggregative parts. But this does not follow. The principle can be and is used to establish multiplicity without characterizing the multiple entities or their relation to each other. Second,

But what, then, explains how these separate sources are parts of a single soul? Again, the question is partly mereological: why are my rational, spirited, and appetitive elements parts of one soul whereas my rational and spirited elements and your appetitive element are not? But Plato also needs to explain how the whole soul could be the locus of moral responsibility, in a way that leaves open the possibility that it could also be responsible for all vital activity and for unified consciousness.

### 3. Failed explanations of the soul's unity

One might be tempted to say that there is nothing psychological that explains the unity of a soul but only the brute fact of embodiment that renders my rational, spirited, and appetitive elements parts of me and leaves your psychological elements out. There are two reasons one might offer for this suggestion. First, the soul is supposed to explain the distinction between a living and dead body (353d9-10). What is to be explained are the vital activities of a living body. What does the explaining is the soul of that body. With this pattern of explanation, one might think that because we are dealing with just one living body, so we are dealing with just one soul to explain that life. Second, one might think that the soul by itself is supposed to be a simple unity whose embodied complexity is due to embodiment (611a-612a with 518d-e). Then, since the disembodied soul needs nothing to unify it because it is not a many, one might think that the embodied soul's complex unity is best explained not by the soul itself (which explains only the simple unity of a disembodied soul) but by precisely what makes it complex, the body.

These reasons are not convincing. The first gives at best a reason why we should want an account of the soul's unity; it does not actually give that account. The second is perhaps unduly speculative, since Socrates in Book Ten pulls up short of saying that the disembodied soul is partless (612a3-6).<sup>36</sup> But more importantly, the basic claim that the

---

Shields argues that Socrates does not establish that the soul is essentially a complex unity since the true nature of the soul might not be multiform (611b-612a) and since the Book Four argument in any case uses a contingent a posteriori premise ("Our souls sometimes both have an appetite to drink and refuse to drink" [439c1-2]). But Socrates does not need to show that the soul is essentially multiform; he can and does allow that the disembodied soul might be simple. He does mean to show that the embodied soul is necessarily multiform—notice the definitions of the virtues in terms of the distinct parts of the soul, though virtues are found in people free of internal conflict—but the a posteriori premise can and should be construed as the existence proof for the modal premise Socrates needs, which is that every embodied soul can experience conflict (cf. Bobonich 2002, 235).

<sup>36</sup> Before and after 611a-612a in Book Ten, he gives no support to the idea: his argument for immortality assumes nothing about the simplicity of the soul, unlike the so-called "Affinity Argument" in *Phaedo* 78b-84b (see Brown 1997, and contrast Shields 2001, 143-144) and some of what Er says is more readily intelligible if the disembodied souls retain spirited and appetitive

soul's unity is due to the body flouts two basic and widely attested Platonic commitments. First, it makes the soul's unearned unity a relational property, dependent upon something external to the soul. But when Plato's dialogues discuss unity, they invariably describe it as an intrinsic property.<sup>37</sup> In the Republic, Socrates says that the just person

puts himself in order, is his own friend, and harmonizes the three parts of himself like three limiting notes in a musical scale—high, low, and middle. He binds together those parts and any others there may be in between, and becomes entirely one thing out of many, moderate and harmonious. (443d5-e2)

This unity is internal harmony. In calling the condition moderate, Socrates also means that it, just like the unity of the city (cf. 462b-e), is defined by internal agreement (442c9-d2, cf. 432a6-b1). This is the earned unity of the virtuous few, and not the unity we are looking for, which is the unearned unity of every human soul. But this specification nevertheless casts doubt on taking unearned unity to be a relational property. Second, the claim that the soul's unity is due to the body flouts the Platonic dictum that the soul rules the body.<sup>38</sup>

At this point, one might try a different tack. If the human soul contains three distinct sources of psychological states and activities but is nevertheless a morally responsible agent, then perhaps it is quite literally a "plural subject" or corporate agent. On this view, the psychological unity that belongs to a human being is nothing more and nothing less than what belongs to a group such as a Parent-Teacher Organization, and the practice of holding a human being morally responsible is akin to recognizing collective responsibility.

This approach needs an account of collective agency. What makes a collective a plural subject and thus a single agent? One popular answer is that multiple agents become a plural subject by sharing an intention.<sup>39</sup> This is promising. The approach can work for all sorts of plural subjects and that it can even explain how a plural subject becomes a kind of unity.<sup>40</sup> Moreover, it would seem to fit well the Republic's insistence that earned unity is a matter of agreement among the parts of the soul. But there is some reason for doubt.

elements in them. Moreover, for what it is worth, the Timaeus and Phaedrus suggest divergent answers to the question of whether the disembodied soul is complex or simple. So no matter how suggestive 611a-612a might be, it is difficult to be confident that Plato held a settled view on the question.

<sup>37</sup> I here concentrate on the Republic and in the next section draw help from the Phaedrus and Gorgias. For evidence in the Theaetetus, Sophist, and Philebus, see Harte 2002.

<sup>38</sup> See 353d with Phaedo 80a and 94b-e, Alcibiades 130a, and Philebus 35d. Cf. Aristotle, De an. I 5 411b5-9.

<sup>39</sup> See, e.g., Gilbert 1989 and 2000.

<sup>40</sup> See, e.g., Velleman 1997.

Given the different aims that Socrates attributes to the three parts of the soul, it seems plausible that some vicious people will have psychological elements that fail to share intentions. What would explain the unity of their souls, and thus how their souls are the locus of moral responsibility? The approach needs to make explicit that a group can be a plural subject by sharing an intention that its members do not all recognize, if the relationships within the group give all its members intentions whether they recognize them or not.<sup>41</sup> Perhaps the psychological elements in even a deeply conflicted soul share intentions in this "pre-reflective" way.

There is, I shall show, something right about this approach. I will argue that the unearned unity of the soul is, like earned unity, a function of agreement among its parts, that every soul is disposed toward agreement in fundamental aims and achieves some minimal agreement. So the moral responsibility of the whole soul is akin to the collective responsibility of a corporate agent. But Plato's account of the soul's unity cannot take the soul to be merely a corporate agent. For one thing, at least some corporate agents are formed and dissolved voluntarily, by convention, whereas the human soul is a natural composite (cf. 588d5-6). There must be natural relations among the soul's parts that underwrite the tendency to agreement, and these need to be made plain. For another thing, Plato needs more than a plural subject if his account is supposed even potentially to explain the coordinated source of all vital activities and especially the unity of consciousness. He needs a fresh approach, and we can find one if we take a hint from the Phaedrus.

#### **4. A fresh approach to unity**

In the Phaedrus,<sup>42</sup> Socrates tries to lead young Phaedrus to take up the genuine art of oratory and to reject what ordinarily passes as the art of oratory. Among other things, he claims that ordinary teachers of oratory give insufficient attention to how an expert speaker fashions his speech into a whole (269c1-5). To motivate this claim, he appeals to the difficulty of crafting a tragedy, and Phaedrus readily sees that a tragedy must be a "system" whose elements fit with one another and with the whole (268d3-5). On Socrates' view, "every speech, since it has a body of its own, must be composed like a living being, so

---

<sup>41</sup> See esp. May 1987 and 1992.

<sup>42</sup> I cite the text in Burnet 1901, and quotations are mine.

that it is neither headless nor footless but instead has things in the middle and at the extremities which are written to fit each other and the whole."<sup>43</sup>

There are two points to emphasize here. First, something is made a whole by the special arrangement of its parts. (One might want to insist, more precisely, that something is made a whole by the arrangement of things that are, because of the arrangement, its parts. I forgo such precision here.) Second, this property of wholeness or unity that is conferred by the arrangement of parts belongs not just to speeches and tragedies but also to living bodies. In fact, the Gorgias seems to encourage still further generalization.<sup>44</sup> There Socrates maintains that every craft, and not just tragedy and oratory, puts its products into a special organization (503d-504e, cf. 506c-508a), and he suggests that the same kind of order or organization is manifest in the natural world (507e-508a). In the Gorgias, he does not talk of unity or wholeness exactly, but the language he does use (of order [κόσμος] and organization [τάξις]) perfectly fits what he says in the Phaedrus about what makes elements into parts of a whole.<sup>45</sup> This suggests a perfectly general account of complex unities (wholes). On a first approximation, diverse elements constitute a complex unity, a whole, if and only if they are suitably organized in relation to each other and in relation to the whole that they constitute.

But what makes diverse elements suitably organized in relation to each other and in relation to the whole that they constitute, as opposed to being organized in some unsuitable way? In the Phaedrus, Socrates maintains that if the elements of a speech are not organized well, then the speech not only fails to constitute a whole, its speaker fails "to speak these elements persuasively" (269c2-3). That is to say, since oratory is the art of leading souls through speech (261a7-8, cf. 271c10), a speech without good organization fails to perform its function well. Presumably, Socrates supposes that the elements of a speech have the effects they do not because each would have its effect no matter its context but because of the way in which they all interact with each other. So the speech as

---

<sup>43</sup> 264c2-5: ἀλλὰ τόδε γε οἶμαι σε φάναι ἄν, δεῖν πάντα λόγον ὥσπερ ζῶον συνεστάναι σῶμά τι ἔχοντα αὐτὸν αὐτοῦ, ὥστε μήτε ἀκέφαλον εἶναι μήτε ἄπουν, ἀλλὰ μέσα τε ἔχειν καὶ ἄκρα, πρέποντα ἀλλήλοις καὶ τῷ ὅλῳ γεγραμμένα. Most translators appear to construe the phrase σῶμά τι ἔχοντα αὐτὸν αὐτοῦ as if it expressed a purpose or object of the composition ("...so that it has a body of its own") instead of taking it as a (causal) circumstance. But the sentence already has a clause that expresses the intended result of the composition (ὥστε...). So I translate as though Plato is being playful: since speeches have bodies, they'd best be composed as if they were living bodies. But this might rest on a sense of 'body [σῶμα]' that comes later (as at Aristotle, Rhet. 1354a15). Nothing much hangs on this.

<sup>44</sup> Text: Dodds 1959.

<sup>45</sup> Dodds (1959, ad 504a1) sees the relevance.

a whole has the effect it does only because of the way in which its elements have been organized to have the effects they do.

This explanation generalizes well to other artifacts and living bodies. A shipbuilder's product, if it is not a whole of suitably ordered parts, will fail to be seaworthy. A living body, if it is not a whole of suitably ordered parts, will fail to be able to perform its vital activities. Plainly, there are degrees of organization and degrees of success here. A speech need not be perfectly organized to be persuasive, but the best speeches will display greater organization than lesser ones. A ship need not be perfectly well organized to be seaworthy, but the best ships will display greater organization than lesser ones. Similarly, a living body need not be perfectly organized, but it will perform more vital activities and perform them better if its parts are better or more organized. Still, in every case, the range of permissible organization has a lower limit. Below this limit, there is nothing performing the function of the relevant kind, and so nothing of the relevant kind at all: sometimes we have a heap of potential ship-parts but not a ship, a string of words but not a speech, or an assortment of limbs but not a living body.

So the diverse elements are parts of a whole if and only if they are organized in relation to each other and to the whole in such a way that the whole is able to perform its function. Of course, the function of the whole is related to the functions of the parts. The ship as a whole is seaworthy because its parts do what they are supposed to do, in relation to each other and in relation to the whole ship. No complex whole can be a greater unity than the one in which each of its parts is doing what it is supposed to do as a part (that is, in relation to the other parts and the whole).

So understood, the Phaedrus, with help from the Gorgias, suggests that a complex whole is a unity because of the organization of its parts, and it suggests that this is so because a complex unity's ability to do what it does depends upon the orderly relations of its parts.<sup>46</sup>

---

<sup>46</sup> Harte (2002) finds in Plato's Sophist, Timaeus, and (especially) Philebus a broadly similar account of wholes in terms of the structure or harmony of parts. The account she finds, however, does not welcome degrees of structure and wholeness or unity. If the Republic offered this same account, earned unity would be the only kind of unity that souls could enjoy, and in fact, any talk of the whole soul would be merely an unjustified manner of speaking except when we were speaking of the virtuous. It seems to me that Plato has good reasons for wanting the soul to have unearned unity (see §2 above) and that Socrates speaks sincerely as though it does (see §1 and §3). If my account is right, then Harte's is wrong, or Plato changed his mind after writing the Republic.

## 5. Explaining the soul's unity

By now, my hypothesis should be obvious: I suggest that Plato treats the human soul like a speech, a ship, or a living body. The soul's unity is a function of the order or harmony of its parts. The parts' causal relations with each other produce this order or harmony: the more the parts do what they are supposed to do, as parts, the more the soul enjoys unity. On this view, Socrates can and does say that the soul's unity takes the specific form of agreement, since agreement is the harmony or order of minded things. And on this view, the soul can perform vital activities—that is, its function of living—only if it has some minimal agreement and order, although it requires perfect agreement and order only for excellent or virtuous activities. That is the distinction between unearned and earned unity.

In the *Republic*, Socrates does not make this account fully explicit. He does explicitly characterize the maximal unity that the just person earns as a kind of harmony and agreement (462d5-e2, quoted and discussed above), and with his account of justice, he makes clear that the parts produce this harmony by doing what they are supposed to do. But he only hints at how to extend this account of unity to less than perfectly just souls. First, he explains to Glaucon,

For the sake of a model we were seeking both what sort of thing justice itself is and what sort of man the completely just man would be if he came into being, and likewise injustice and the most unjust man, in order that, by looking at them and how they seemed to us concerning happiness and its opposite, we would have to agree about ourselves as well, that the one who was most like them would have a portion of happiness most like theirs.<sup>47</sup>

Here Socrates makes two important suggestions. There is a correlation between psychological order and harmony (that is, justice) and psychological functioning (that is, living), and each of these features comes in degrees. That is why the imperfect soul lives more or less well (has a greater or lesser "portion of happiness") according to how closely its order and harmony approximates the just soul's. Socrates clarifies and elaborates on these suggestions in his first "proof" that it is always better to be just than unjust in Books Eight and Nine. He argues that the perfectly just soul lives best (is most happy) and other

---

<sup>47</sup> [Rep. 472c4-d1](#): Παραδείγματος ἄρα ἕνεκα, ἣν δ' ἐγώ, ἐζητοῦμεν αὐτό τε δικαιοσύνην οἷόν ἐστι, καὶ ἄνδρα τὸν τελέως δίκαιον εἰ γένοιτο, καὶ οἷω ἂν εἶη γενόμενος, καὶ ἀδικίαν αὐτὴν καὶ τὸν ἀδικώτατον, ἵνα εἰς ἐκείνους ἀποβλέποντες, οἷοι ἂν ἡμῖν φαίνωνται εὐδαιμονίας τε πέρι καὶ τοῦ ἐναντίου, ἀναγκαζώμεθα καὶ περὶ ἡμῶν αὐτῶν ὁμολογεῖν, ὅς ἂν ἐκεῖνοις ὅτι ὁμοίωτος ἦ, τὴν ἐκείνης μοῖραν ὁμοιοτάτην ἔξῃεν.

souls live more or less well (are more or less happy) according to how closely their order and harmony approximates the just soul's (580a-c).

That hint does not commit Socrates to the thought that imperfect souls enjoy unearned unity. But he also suggests in Book One that a soul needs some minimal justice if it is to perform its function. After arguing that a city or a band of thieves would be unable to achieve an unjust purpose if its members were unjust toward each other (351d), he generalizes the point thus: because injustice sows internal disharmony, it threatens to render "a city, a family, an army, or anything else," including a soul, "unable to do anything with itself."<sup>48</sup>

One might resist at least part of this account if one thought that unity, for Plato, required perfect harmony and agreement. On this competing view, a soul might need some minimal order to function, but the minimal order does not make the soul a unity. Even the souls that most resemble the model of justice lack unity; each fails to be a single soul.<sup>49</sup> Thus far, I have suggested three general reasons to think otherwise. First, Socrates talks of moral responsibility and seems to assume the coordination of vital activities and the unity of consciousness, and all of these are best explained if every soul is a unity. Second, in the Republic, Socrates refers to the divisions in the soul as parts and his argument for partitioning characterizes the soul as a whole comprising parts. Since he talks this way of souls that experience conflict, he must be inclined to think of even an imperfectly unified soul as a single, whole soul. Third, the Phaedrus suggests a general account of parts and wholes that calls for degrees of unity, and the Republic's hints are right in line with that general account.

But there is further reason, because the Republic offers resources to explain how even the most fractured souls maintain some agreement and unity. Socrates offers two causal principles about the interaction of the soul's parts that explain how my rational, spirited, and appetitive elements are parts of my soul but your spirited element is not. So the causal principles illuminate the structural order that makes the three psychological elements parts of a whole, and in particular, they illuminate how the soul is a natural composite and not merely a corporate agent. Moreover, these causal principles help to

---

<sup>48</sup> 351e10-352a3: Οὐκοῦν, τοιάνδε τινὰ φαίνεται ἔχουσα τὴν δύναμιν, οἶαν, ᾧ ἂν ἐγγένηται, εἴτε πόλει τινὶ εἴτε γένει εἴτε στρατοπέδῳ εἴτε ἄλλῳ ὄρωσιν, πρῶτον μὲν ἀδύνατον αὐτὸ ποιεῖν πράττειν μεθ' αὐτοῦ διὰ τὸ στασιάζειν καὶ διαφέρεσθαι, ἔτι δ' ἐχθρὸν εἶναι ἑαυτῷ τε καὶ τῷ ἐναντίῳ παντὶ καὶ τῷ δικαίῳ; 'Including a soul' after 'anything else' (ἄλλῳ ὄρωσιν), is secured, though it does not need to be, by 352a6-7 and following. The Grube-Reeve translation (in Cooper 1997) says, "...incapable of achieving anything as a unit," which might misleadingly suggest that the city or soul could function as something other than a unit.

<sup>49</sup> I take it that this would be what Harte would say were she inclined to extend her account of Plato on parts and wholes back from the later dialogues to the Republic (see n. 00 above).

explain the connection between unearned and earned unity, because what gives the soul unearned unity also makes earned unity possible for it.

The first causal principle can be called Plato's hydraulic principle of psychology. Socrates says, "When someone's desires incline strongly for one thing, they are thereby weakened for others, just like a stream that has been partly diverted into another channel" (485d6-8; cf. 328d3-5). This is plainly meant to apply across different soul-parts within the same soul. If my spirited desires grow stronger, my rational and appetitive ones grow weaker. Just as plainly, the principle cannot apply to psychological elements in different people: if your spirited desires grow, nothing automatically follows for my rational and appetitive desires. (Of course, you might influence me, and if my spirited desires were to strengthen, too, then my rational and appetitive desires would weaken. But this is a rather different causal story.)

The second I will call Plato's principle of psychological hegemony. In Books Eight and Nine, Socrates maintains that different kinds of people are ruled by different soul-parts. As he explains this, to be ruled by a soul-part is to take the ends of that soul-part to be one's ends, generally.<sup>50</sup> If I am ruled by spirited desire, I take my good to be honor or victory, the ends of the spirited part. If I am ruled by appetitive desire, I take my good to be bodily pleasure or, perhaps, if I calculate a bit, the best means to achieve bodily pleasure, which is money. This pattern requires certain causal interactions within the soul. Because the spirited or appetitive part of the soul gets more powerful, the rational part, which alone calculates what is good for the whole soul (441e3-4, 442c4-7), begins to calculate that it would be good to achieve what the powerful part wants. When a young man is influenced by his father to pursue wisdom and by everyone else to pursue money and publicly contested honors, he calculates that his good lies in the middle way, pursuing the spirited part's end of honor (549c-550b). When another young man sees his honor-loving father killed or reduced to poverty, he calculates that it would be safer to pursue money, the end of the appetitive part (553a-d).

This principle explains what it is for a soul to be ruled by some part of the soul by explaining what it is for the rational part of the soul to be ruled. Corollaries to the principle explain what it is for the spirited or appetitive part to be ruled. When spirit is ruled by either reason or appetite, the spirited part finds honor in the ends of the ruling part. When

---

<sup>50</sup> The evidence for this principle is diffuse; it emerges as the best explanation of what Socrates says in Books Eight and Nine about what it is for one part of the soul to rule. At Republic 580d-581e, Socrates insists only that people call most pleasant the life that is ruled by the same part of the soul that rules them. His description of the timocratic, oligarchic, democratic, and tyrannical souls earlier in Books Eight and Nine makes it clear that people also call good the ends that are pursued by the part of the soul that rules them (see below for some discussion).

appetite is successfully ruled by either reason or spirit, it limits its pleasures to those deemed good by ruling reason or honorable by ruling spirit (and thus good by ruled reason).

Again, the principle of psychological hegemony and its corollaries explain direct connections in me without imputing direct connections between your spirited element and me. That your spirited element has grown so strong as to dominate your soul does not force my rational part to determine that honor or victory are my good. Again, you might influence me. If my spirited desires were to strengthen, too, then I would calculate that honor or victory is good for me. (This is how the mother, servants, and others influence the young man with the philosophical father (549c-550b).) But that is not the direct causal influence that my soul-parts enjoy over each other.

It is perhaps obvious that the two causal principles are related. For instance, a rational part can come to accept that honor or victory is what is good only if it is too weak to grasp what really is good, but it will be too weak if the spirit has taken much of the soul's "hydraulic power" away from reason.

It should also be obvious that these are not the only causal principles Plato would invoke to explain psychological behavior. Among the most important other principles, Plato clearly believes that the appetitive part has its own internal tendency to grow stronger, which requires that it be checked at every turn (442a). Socrates paints this principle into his imaginative portrait of the soul at the end of Book Nine, when he says that the just person must continually prevent the many-headed beast of appetite from sprouting savage heads (589a-b; cf. 571b-572b). He also assumes it when he asserts that gold and silver need to be kept away from guardians (416e-417a) and when he predicts that a person with a timocratic soul will become progressively more money-loving over the course of his lifetime (549a-b). So there are causal tendencies in the soul that encourage it to become less orderly and unified.

Still, the hydraulic principle of psychology and the principle of psychological hegemony explain how the soul tends to organize and unify itself, even when its appetite grows. Even in these circumstances, there is a tendency for the rational part to agree with appetite about what is good and a tendency for the spirited part to find some honor in what appetite pursues. This tendency pertains to my soul-parts without extending to yours, and it explains the sense in which my soul-parts are parts of one psychological system, one soul.

These causal principles also help to explain why all three soul-parts share responsibility. First, they foster a tendency to agree which is a tendency to have shared aims, whether the parts fully recognize these shared aims or not. Second, when one part of the soul does something, the others are causally implicated; at the least, each of the others

could have been a more powerful obstacle. So it would not be inappropriate to hold the rational part alone responsible in the afterlife for what the other parts had done during embodied life. The rational part at least tended to endorse the aims that motivated those actions and at least failed to prevent the actions so motivated.

Finally, these causal principles not only explain the unearned unity of every human soul but they also help to explain how it is possible for any human soul to earn a more perfect unity. The hydraulic principle allows for the cultivation of spirit to minimize the threat that appetite poses, and the corollaries to the principle of hegemony explain how the spirit and appetite could lash their pursuits to what reason recognizes as good for the whole person.<sup>51</sup>

## 6. Sustaining this explanation of the soul's unity

This account of the soul's unearned unity is incomplete. Most obviously, it lacks a precise notion of internal agreement. This calls for some reckoning of the mental content of the various soul-parts, to see how there could be shared content. The problem is connected, then, to the problem of the unity of consciousness. But these might be viewed as further issues to be worked out within the framework provided by a basic account of the soul's unearned unity. Here I consider objections to two claims essential to the basic account, that every human soul is a unity, and that the unearned unity of every human soul is explained in terms of the orderly arrangement of its parts.

To develop the latter claim, I have argued that the interactions of the soul-parts, regulated by two causal principles, tend to give the soul agreement and orderliness. But one might object that I have made this too easy by assuming that every soul is perfectly ruled by some part.<sup>52</sup> Socrates disavows this assumption in the *Republic* (544c-d, cf. 445c). If some human beings fail to be ruled by just one part of their soul, then their psychology will have to be characterized as a "mixed constitution." How is such a mixed constitution a unity?

---

<sup>51</sup> I talk more about this, and especially the role of the hydraulic principle, in Brown 2004.

<sup>52</sup> One might also object that I have ignored the differences among souls that are ruled by appetite. Some are ruled by necessary appetitive desires (oligarchically), some by unnecessary but lawful appetitive desires (democratically), and some by lawless appetitive desires (tyrannically). Socrates plainly believes that these lives differ in their internal agreement and in their functioning (580a-c), and this needs to be explained. Presumably, Socrates believes that necessary appetitive desire cohere more regularly than unnecessary ones do, perhaps because they are constrained by facts about human nature (that is, facts about what is naturally necessary for a human), and that the democratic soul's agreement that every lawful desire is to be treated equally constitutes more agreement than the tyrannical soul's lawless lusts can muster.

I do not have a convincing answer to this question, in part because Socrates says so little about mixed constitutions in the Republic. But I do think that the question need not undermine the account of unearned unity I am attributing to Plato. After all, even in the mixed constitutions, the causal principles I have highlighted apply. Even here, there is a tendency toward internal agreement among the parts of the soul, and this tendency explains what unifies these parts together without including those elements in the mix, it explains how these parts share responsibility, and it explains how the soul is capable of greater unity.

I suspect that dissatisfaction with this response rests on a deeper worry that some souls fail to have unearned unity. So let me address three grounds for the deeper worry. First, one might appeal to the city-soul analogy. Socrates tells Adeimantus,

You are happily innocent if you think that anything other than the kind of city we are founding deserves to be called a city... We need a bigger category to address the others because each of them is a great many cities, not a city... At any rate, each of them consists of two cities at war with one another, that of the poor and that of the rich, and each of these contains a great many.<sup>53</sup>

Now, if none but the ideal city is really a city, because every other claimant fails to be a unity, then the analogy of city and soul suggests that none but the ideal soul is really a soul, because every other claimant fails to be a unity.<sup>54</sup>

But the analogy of city and soul can only suggest; it cannot prove. There is an open question about which features of cities have analogues in souls and vice versa, and so it is easy enough to suppose that this particular feature of cities does not have an analogue in souls. Moreover, we can justify this supposition. First, cities do not seem to have the same causal regularities governing the interaction of their parts that souls do. In particular, the principle of psychological hegemony has no analogue for the city: that the oligarchs rule

<sup>53</sup> Rep. 422e1-423a2: Εὐδαίμων εἶ, ἦν δ' ἐγώ, χρῆ ὅτι οἶει ἄξιον εἶναι ἄλλην τινὰ προσεπειεῖν πόλιν ἢ τὴν τοιαύτην οἶαν ἡμεῖς κατασκευάζομεν... Μειζόνως, ἦν δ' ἐγώ, χρῆ προσαγορεύειν τὰς ἄλλας: ἐκάστη γὰρ αὐτῶν πόλεις εἰσὶ πάμπολλαι ἀλλ' οὐ πόλις... δύο μὲν, κὰν ὅτι οὖν ἦ, πολεμία ἀλλήλαις, ἢ μὲν πενήτων, ἢ δὲ πλουσίων: τούτων δ' ἐν ἑκατέρῃ πάνυ πολλαί. To remove potential distractions, I have elided Adeimantus' expression of confusion and a reference to some proverbial saying related to some game, for which see Adam 1905, ad 422e.

<sup>54</sup> One might doubt that Socrates is entirely serious here. In Book One, he urges that even a band of thieves must have some justice to achieve its purpose (351c-d), which reinforces the idea that justice comes in degrees and encourages the thought that even an imperfectly functioning community has enough justice to be a community. I am nevertheless inclined to think that Socrates is seriously reserving 'polis' for the most successful communities (cf. Pol. 293e). Perhaps the ensuing discussion above helps to explain why he does this: he might be worried that attributing goodness to every community undercuts the goodness that he wants to predicate of only the best communities. Alternatively, his reasons might run deeper: perhaps the causal relations among human beings, unlike those among the parts of the soul, do not underwrite unearned unity. I thank Anton Ford and Clerk Shaw for discussion of this point.

does not tend to make the democrats accept the oligarchs' rule. This is not to say that there is no tendency toward shared values in a city: Plato plainly recognizes the power of a shared culture to shape the citizens' values.<sup>55</sup> But shared values do not necessarily produce agreement about who should rule in the city as it does in the soul. Agreement in values just is agreement about who should rule in the soul. In the city, by contrast, the oligarchs and democrats can agree broadly in their values but disagree about whether the rich alone should rule. Second, competing factions in a city do not share responsibility for "what the city does" as parts of a divided soul do. That is why Socrates insists that when Kallipolis is at war with another Greek city, the Kallipolitans will not ravage the entire enemy city but will target the real enemies in the city who are responsible for the war (471a-b).<sup>56</sup> These differences between cities and souls explain why it is reasonable to say that a divided city is not a whole city but not to say that a divided soul is not a whole soul.

But this might seem to prove too much. In Book Eight, Socrates returns to his point about divided cities, and he extends his analysis to divided souls. Having said an oligarchic city is "not one but two, one of the poor and one of the rich" (551d5), he claims of an oligarchic soul, "Such a person would turn out not to be free of civil strife in himself, and he would be not one but in a way twofold, but for the most part he would have better desires in control of worse."<sup>57</sup> This is the strongest evidence that Socrates is willing to deny unearned unity to a human soul. But it is striking how limited his willingness is. First, in the case of divided cities he says forthrightly that they are two (δύο, 422e6 and 551d5), but of the soul he says only that it is "in a way twofold" (διπλοῦς τις, 551e1). Second, this threat to the soul's unearned unity arises exceptionally, in special cases. The "in a way twofold" soul does not arise in everyone who seeks money (and the satisfaction of necessary appetitive desires) above all else. Rather, only those who have these aims and lack education experience unnecessary appetitive desires that they can restrain only by force and the fear of punishment (554b-d). Those who are educated well by their cities can lack the desire for wisdom and nevertheless have a share of virtue (619c6-d1). Third, even the "in a way twofold" soul is causally organized in a way that tends toward agreement.

<sup>55</sup> See Lear 1992 and Burnyeat 1999.

<sup>56</sup> Cf. Thucydides' presentation of the Mytilenian debate. The Athenians come to understand that the democrats in Mytilene were not responsible for the revolt, because that revolt was led by the oligarchic faction and stopped by the democrats themselves (esp. 3.47.3). It is true that the Athenians come to see this belatedly (3.36) and only by a slight majority (3.49.1). But this is not because the point is reasonably contested: at least, Thucydides suggests that passionate rage has blinded those who want to punish the Mytilenian democrats (3.36.2, 3.44.4).

<sup>57</sup> 554d9-e2: Οὐκ ἄρ' ἂν εἴη ἀστασίαστος ὁ τοιοῦτος ἐν ἑαυτῷ, οὐδὲ εἷς ἀλλὰ διπλοῦς τις, ἐπιθυμίας δὲ ἐπιθυμιῶν ὡς τὸ πολλὸν κρατούσας ἂν ἔχοι βελτίους χειρόνων.

Socrates says explicitly that such a soul "for the most part would have better desires in control of worse," and presumably, if the better desires were to lose control of the worse, then the soul would tend away from oligarchy and towards democracy by jettisoning the distinction between better and worse appetitive desires and treating all equally. That is to say, the hydraulic principle and the principle of psychic hegemony find expression even in the "in a way twofold" soul. In sum, though such a soul lacks minimal agreement and thus unity, it is nevertheless potentially one, and this explains why Socrates treats it as not one but not quite two either. The way in which this exception is carefully limited proves the existence of the general rule: souls have unearned unity.

But there is a third source of the worry that souls lack unearned unity. In the Republic, Socrates seems to identify the Good with unity. He characterizes the good soul in terms of unity (443c-e), and he says that there is no greater good for a city than that which makes it a unity (462b1-2). These remarks are readily intelligible if unity is what makes a thing good. So, too, are the importance of mathematics to the ascent to the Good (through mathematics an account of the unit is learned), the superiority of the Good over the other Forms (the Good is the unity or coherence of them, and not another alongside them), the goodness of the other Forms (they are good by being part of the unified or coherent order), and the intelligibility of the other Forms (they are fully known only by coherence).<sup>58</sup> None of this is surprising to those who take seriously the report that Plato's Lecture on the Good identified the Good as Unity.<sup>59</sup> But if Unity and Goodness are the same thing and if every human soul is a unity, then every human soul is good. And this is plainly false: Socrates plainly thinks that some, indeed most, human souls are bad.

This objection is plainly right about the Republic's understanding of goodness.<sup>60</sup> But we have to distinguish between unearned and earned unity. Every soul enjoys unearned unity and to that extent is good, but most souls lack earned unity and to that extent are bad. In ethics, the latter point rightly receives emphasis. But we can hold a person responsible as a bad soul only because he or she has a soul that enjoys unearned unity. Every soul, because it is to some minimal extent unified, is to some extent good.

---

<sup>58</sup> See also Cooper 1977, 144; Hitchcock 1985; Fine 1990, 97-98; Irwin 1995, 272-273; Annas 1999, 108; and Burnyeat 2000.

<sup>59</sup> See, e.g., Aristoxenus, Elementa Harmonica II 1.

<sup>60</sup> Santas 2001 argues that Plato offers two accounts of goodness in the Republic: according to Book One, something is good by possessing the excellence of its kind and thus performing its kind's function well, and in later Books, something is good by being unified. My explanation of how Plato understands unity links these two accounts: the orderly interactions of parts unify complex entities and make them perform their functions well. See also Singpurwalla 2006.

This is only superficially puzzling. Although his account of goodness as unity commits him to the view that every complex material entity is good, Plato is independently committed to this view by his acceptance of cosmological providence: on his view, every material thing is made in a way that it is good for it to be.<sup>61</sup> This does not require him to believe that everything is a good member of its kind. It is easy to distinguish between the unearned unity that all members of a kind share and the earned unity that only some enjoy and between the goodness that belongs to every member of a kind and the goodness that belongs only to good members of a kind. Nor do these distinctions require Plato to say that there are two different forms of goodness or of unity.<sup>62</sup> The same unity makes every existing soul a complex unity and makes some souls virtuous, and the same goodness is in all existing souls to a limited extent and in virtuous souls to a much greater extent. Generally, this unity and this goodness are the order and harmony of the parts. Such order and harmony allow a thing to do what it does and thus be the thing that it is. In the soul, the order that brings unity and goodness manifests itself as agreement: without some minimal agreement, the soul cannot be a functioning agent,<sup>63</sup> and without perfect agreement, the soul cannot live well.

### Sources Cited

- Adam, C., and P. Tannery, eds. 1964-1976. Oeuvres de Descartes. 12 vols. rev ed. Paris: Vrin.
- Annas, Julia. 1999. Platonic Ethics, Old and New. The Townsend Lectures. Ithaca: Cornell University Press.
- Archer-Hind, R.D. "On Some Difficulties in the Platonic Psychology." Journal of Philology 10 (1882): 120-131.
- Bobonich, Christopher. 2002. Plato's Utopia Recast. Oxford: Clarendon.

---

<sup>61</sup> This view is developed most elaborately in the Timaeus, of course, but it is hardly limited to that dialogue. See, e.g., Phaedo 97c and Gorgias 507e-508a.

<sup>62</sup> It is tempting to think otherwise, and to bolster this thought, one might note that earned unity confers goodness picked out by 'good' as a "(logically) attributive adjective" (in the sense provided by Geach 1956, according to which an adjective a is attributive if 'X is an a F' does not entail that 'X is a'), whereas unearned unity confers goodness referred to by 'good' as a "(logically) predicative adjective" (for which the entailment holds). In Toronto, I gave in to this temptation, and distinguished between "ethical" and "metaphysical" unity. But I now think that this misrepresents Plato's thought. He takes both earned and unearned unities to be unities in the same way: unity is order or harmonious organization, which in the soul is a kind of agreement.

<sup>63</sup> Cf. 608c-611a with Brown 1997: disorder in the soul, vice, cannot fully undermine the soul's order and destroy it, which is why Socrates supposes that nothing can destroy the soul.

- Brown, Eric. 1997. "A Defense of Plato's Argument for the Immortality of the Soul at Republic X.608c-611a." Apeiron 30: 211-238. Reprinted in Essays on Plato's Psychology, ed. Ellen Wagner (Lanham, Maryland: Lexington Books, 2001), 297-322.
- . 2004. "Minding the Gap in Plato's Republic." Philosophical Studies 117: 275-302.
- Burnet, J., ed. 1901. Platonis Opera. vol. 2. Oxford Classical Texts.
- Burnyeat, M.F. 1999. "Culture and Society in Plato's Republic." in The Tanner Lectures on Human Values, vol. 20, ed. Grethe B. Peterson (Salt Lake City: University of Utah Press), 215-324.
- . 2000. "Plato on Why Mathematics is Good for the Soul." Proceedings of the British Academy 103: 1-81.
- Cooper, J.M. 1977. "The Psychology of Justice in Plato." American Philosophical Quarterly 14: 151-177. Reprinted in his Reason and Emotion: Essays in Ancient Moral Psychology and Ethical Theory (Princeton: Princeton University Press, 1999), 138-149, and in Essays on Plato's Psychology, ed. Ellen Wagner (Lanham, MD: Lexington Books, 2001), 91-114.
- . 1984. "Plato's Theory of Human Motivation." History of Philosophy Quarterly 1: 3-21. Reprinted in his Reason and Emotion: Essays in Ancient Moral Psychology and Ethical Theory (Princeton: Princeton University Press, 1999), 118-137.
- Cooper, J.M., ed. 1997. Plato: Complete Works. Indianapolis: Hackett.
- Cottingham, J., R. Stoothoff, and D. Murdoch, trans. 1984. The Philosophical Writings of Descartes. 2 vols. Cambridge: Cambridge University Press.
- Dodds, E.R., ed. 1959. Plato: Gorgias, A Revised Text with Introduction and Commentary. Oxford: Clarendon.
- Fine, Gail. 1990. "Knowledge and Belief in Republic 5-7." in Companions to Ancient Thought-1: Epistemology, ed. Stephen Everson (Cambridge: Cambridge University Press), 85-115.
- Geach, P.T. 1956. "Good and Evil." Analysis 17: 33-42.
- Gilbert, Margaret. 1989. On Social Facts. London: Routledge.
- . 2000. Sociality and Responsibility. Lanham, MD: Rowman and Littlefield.
- Grote, G. 1888. Plato and the Other Companions of Sokrates. 4 vols. London: John Murray.
- Harte, Verity. 2002. Plato on Parts and Wholes. Oxford: Clarendon Press.
- Hitchcock, David. 1985. "The Good in Plato's Republic." Apeiron 19: 65-92.
- Hsu, H-Y. 2007. "The Harmony of the Soul." Humanitas Taiwanica 67: 139-159.
- Irwin, T. 1995. Plato's Ethics. Oxford: Oxford University Press.

- Kamtekar, Rachana. 2007. "The Powers of Plato's Tripartite Psychology." Boston Area Colloquium in Ancient Philosophy, delivered at Brown University, 29 November.
- Lear, Jonathan. 1992. "Inside and Outside the Republic." Phronesis 37: 184-215. Reprinted in Essays on Plato's Psychology, ed. Ellen Wagner (Lanham, MD: Lexington Books, 2001), 169-202.
- Lorenz, Hendrik. 2006. The Brute Within: Appetitive Desire in Plato and Aristotle. Oxford: Clarendon Press.
- May, Larry. 1987. The Morality of Groups. Notre Dame, IN: University of Notre Dame Press.
- . 1992. Sharing Responsibility. Chicago: University of Chicago Press.
- Moline, J. 1978. "Plato on the Complexity of the Psyche." Archiv für Geschichte der Philosophie 60: 1-26. Revised version in his Plato's Theory of Understanding (Madison: University of Wisconsin Press, 1981), 52-78.
- Price, A.W. 1995. Mental Conflict. London: Routledge.
- Reeve, C.D.C. 1988. Philosopher-Kings: The Argument of Plato's Republic. Princeton: Princeton University Press.
- Santas, Gerasimos. 2001. Goodness and Justice: Plato, Aristotle, and the Moderns. Oxford: Blackwell.
- Shields, C. 2001. "Simple Souls." in Essays on Plato's Psychology, ed. E. Wagner (Lanham, MD: Lexington Books), 137-156.
- Shorey, Paul. 1903. The Unity of Plato's Thought. Chicago: University of Chicago Press.
- Singpurwalla, R. 2006. "Are there Two Theories of Goodness in Plato's Republic?" Apeiron.
- Slings, S.R., ed. 2003. Platonis Rempublicam. Oxford Classical Texts.
- Velleman, J. David. 1997. "How to Share an Intention." Philosophy and Phenomenological Research 57: 29-50.