

# On Socrates' Project of Philosophical Conversion

Jacob Stump

*Northeastern University*

© 2020 Jacob Stump

*This work is licensed under a Creative Commons  
Attribution-NonCommercial-NoDerivatives 3.0 License.  
<[www.philosophersimprint.org/020032/](http://www.philosophersimprint.org/020032/)>*

It is a startling fact that, in Plato's corpus, there is not one unambiguous instance of Socrates' reforming his interlocutor's way of life by the end of a dialogue. In fact, Socrates seems often to fail at improving his interlocutors at all, let alone changing their values. This uninspiring track record has understandably led many scholars to conclude that Socrates is a failure. Alexander Nehamas, e.g., writes that "Plato's works do not at all show that Socrates' dialogue with his fellows has ... beneficial effects." He then asks rhetorically, "How could Socrates claim success for himself in light of such a record?"<sup>1</sup> John Beversluis is more emphatic: "... if the early dialogues show anything, they show Socrates' monumental failure."<sup>2</sup> Some scholars consider such failure to be required by the moral psychology of Plato's middle dialogues, in particular the tenet that strong desires can control a person's evaluative beliefs.<sup>3</sup> Dominic Scott thinks that this is responsible for a Platonic pessimism concerning the role of argument in moral education,<sup>4</sup> and Raphael Woolf writes that it causes a "crisis for Socratic method."<sup>5</sup> Other scholars go further, claiming even that Plato charges Socrates with causing harm.<sup>6</sup> There is thus a wide consensus among scholars that Socrates is wrong to trust in reason and argument as capable of converting people — that, for various reasons, his strategy is deeply and irreparably flawed.

In this paper, I argue for the opposite. On my view, Socrates' project of using reason and argument to try to persuade his interlocutors to value wisdom the most — what I call his project of *philosophical conversion* — is not at all misguided, nor is it depicted to be a failure, nor is

1. Nehamas 1998: 66. See also Nehamas 1999: 70–71.
2. Beversluis 2000: 34. Cf. Blondell 2002: 125.
3. I use the term 'middle dialogues' out of convenience. Nothing in my argument depends on points of chronology, though, at times, I concede a standard ordering of the dialogues (as, e.g., in Vlastos 1991: 46–47) to facilitate dialogue with scholars partial to developmentalism.
4. D. Scott 1999: 28–32.
5. Woolf 2000: 1, fn. 1.
6. Nussbaum 1980: 88, Vlastos 1988: 100, Nehamas 1999: 60–61.

there any reason to suppose that Plato conceived of it as such.<sup>7</sup> What has misled scholars is the expectation that Socrates would be trying to *argue his interlocutors into* changing their values — that is, that he would be aiming to convert them by providing them with reasons why wisdom is more valuable than anything else, intending for their conversion to happen as a result of appreciating the force of those reasons. In contrast, I show that Socrates employs a more sophisticated strategy. Its key component is the use of philosophical argument not to lead an interlocutor to *rationally conclude* that he must change his way of life but rather to cause him to have a certain *affective experience*, one that can be effective at changing his beliefs about how best to live.

It is worth noting that behind these interpretive issues concerning Socrates and Plato lies a general philosophical problem. Arguments can prove conclusions, but they can also persuade. They can persuade a person to adopt new means to her avowed ends, for example. Can they also persuade a person to adopt some new ultimate end, to come to value something more than anything that she currently values? Can they persuade a person to take up an entirely new way of life? If so, how?

In the context of Plato's dialogues, I show that Socrates' project of conversion is best understood as a two-stage process. In the first stage, Socrates convinces his interlocutor of the instrumental value of wisdom, thus motivating him to do philosophy as a means to achieving some non-philosophical goal, e.g., advancing his career or boosting his reputation. In the second stage, this instrumental pursuit is meant to cause the interlocutor to experience the pleasure of philosophical inquiry, an experience that can reconfigure his beliefs about the good. The first stage thus involves an appeal to reason, but it is an appeal

7. The language of *conversion* is not out of place in scholarship on Plato. See, e.g., Kahn 1996: "For Plato, and for Socrates as Plato represents him, the commitment to philosophy is conceived as something comparable to a religious conversion .... This involves a radical restructuring of the personality in its values and priorities" (273). The language of *transformation* is also found, as, e.g., in G. Scott 2000: Socrates tries for "radical transformation in [his interlocutors'] previous goals, desires, and preferences" (102).

only to instrumental reason. The second stage still involves reason and argument, but their point here is not to convince the interlocutor of anything: it is rather to cause the experience of philosophical pleasure, an experience that can lead to his valuing wisdom the most.

Now, some scholars have recognized that, in the first place, Socrates' goal is often to motivate his interlocutors just to do more philosophy. For Thomas Brickhouse and Nicholas Smith, for example, Socratic refutation is meant to cause a person to feel shame, and that shame is then meant to motivate him to do philosophy as a means to rectify it.<sup>8</sup> Other scholars view Socrates as trying to motivate further philosophical inquiry by causing the experience of *aporia* or the awareness of ignorance.<sup>9</sup> What is missing from such accounts, however, is a satisfying answer to why doing more philosophy would at all be effective at causing a person to change what he values the most. As I show, the answer cannot be as simple as that it would expose one to reasons why something else is more valuable. Like some others, then, I view Socrates as aiming in the first place to motivate further philosophical inquiry. But what my account provides, and what other accounts lack, is an explanation of the process of philosophical conversion as a whole.

It is an explanation that we should find familiar and intuitive. Often a student enrolls in a philosophy course solely for an instrumental reason, e.g., to prepare for a successful career in law, but, by the end of the course, and for reasons that she cannot fully articulate, she finds herself attracted to the pursuit of wisdom, just for itself. This is exactly the sort of experience that is at the core of Socrates' project of philosophical conversion. The pre-law student today is the talented, ambitious, up-and-coming politician of ancient Athens, whose priorities Socrates intends to transform by first convincing him to do philosophy just for its instrumental value.

8. See Brickhouse and Smith 1994: 25 and 2000: 58–59.

9. See, e.g., Robinson 1953: 11, Mackenzie 1988: 337, Benson 1995: 89, Slings 1999: 140–141, Politis 2006: 104, and D. Scott 2006: 72.

I proceed as follows. I begin by reviewing the three main arguments that scholars typically advance in support of a Platonic pessimism concerning Socrates' project of philosophical conversion. I argue that none are convincing, and thus that we should consider Socrates' project anew. I go about that in the rest of the paper, aiming to show that philosophical conversion as Plato conceives of it is meant to happen not so much by arguing a person into accepting a new set of values as by causing her to have a certain affective experience. One result is that Socrates' conversion strategy is not only consistent with but deeply informed by Plato's moral psychology.

### 1. Platonic pessimism?

In the *Apology*, Socrates insists that his goal is to change his interlocutors' values by means of reason and argument:

... I shall not cease to practice philosophy, to exhort you and in my usual way to point out to any one of you whom I happen to meet: "Good Sir ... are you not ashamed to be caring about [ἐπιμελούμενος] how to get as much money, reputation, and honors as possible, while as for getting as much wisdom and truth as possible, and getting your soul into the best condition, that you do not care about and do not give any thought to [οὐκ ἐπιμελῆ οὐδὲ φροντίζεις]?" Then, if one of you disputes this and says he does care [ἐπιμελεῖσθαι], I shall not let him go at once or leave him, but I shall question him, examine him, and test him [ἐρήσομαι αὐτὸν καὶ ἐξετάσω καὶ ἐλέγξω], and if I do not think he has attained the excellence that he says he has, I shall reproach him because he attaches little importance to the most important things and greater importance to inferior things. ... For I go around doing nothing other than persuading both young and old among you not to care about [ἐπιμελεῖσθαι] your body or your wealth in

preference to or as strongly as getting your soul into the best possible condition (29d–30b).<sup>10</sup>

In what way does Socrates want to change what his interlocutors care about or value (*epimeleisthai*)? He says that he wants them to care *more* about *the best state of their souls*. I follow many interpreters in understanding the excellence of one's soul to be the condition of being wise and, further, Socrates to want his interlocutors not to begin valuing wisdom only a *bit* more but *more than anything else*.<sup>11</sup> Additionally, Socrates tries to cause this change by the use of reason and arguments.

10. Translations of Plato are from John M. Cooper and D.S. Hutchinson, ed., *Plato: Complete Works* (Indianapolis: Hackett, 1997), with occasional revision.
11. In fact, in the *Apology*, Socrates lists wisdom (*sophia*, *phronēsis*), truth, the best condition of one's soul, excellence (*aretē*), and the city itself all as things that he wants people to care about more. Scholars typically consolidate the list by understanding the relevant excellence to be the excellence of one's soul, and that to consist in wisdom, that is, grasping the truth about "the most important things" (τὰ μέγιστα, *Ap.* 22d7). See Guthrie 1971: 149–150, Burnet 1974: 123, de Strycker and Slings 1994: 332 and 187, and Rowe 2007: 75. This consolidation leaves out the city itself, but it is plausible that, if one values wisdom, then one values also benefitting the city; see de Strycker and Slings 1994: 368. Thus, valuing wisdom is most important because valuing the other things either reduces to it or is implied by it. For my purposes, no distinction between *sophia* and *phronēsis* is necessary (in fact, Plato may not distinguish between them; see Burnet 1974: 123), and the content of the relevant wisdom can be left open: it can range, for example, from a "human wisdom" that consists of not believing that one has knowledge of the most important things (cf. *Ap.* 23a), to "moral knowledge" (Destrée 2017: 223; cf. 219 n.7), to a more robust wisdom consisting of knowledge of the forms. As for valuing wisdom *the most*, note that Socrates identifies wisdom as one of "the most important things" (τὰ πλείστου ἄξια, *Ap.* 30a1–2), with the implication that one should value something in accordance with its importance. Further, he insists that people ought to value the excellence of their souls (i.e., wisdom) more than their reputations (29e1), honor (29e1), money (29d9, 30b1, 41e4), bodies (30a9), and possessions (36c6). But since such things are what people tend to value the most, insisting that wisdom ought to be valued *more* than them is, in effect, to insist that it should be valued more than anything else. See also *Ap.* 41e2–a1, where Socrates states that his sons ought to value virtue (i.e., wisdom) the most, a claim best understood not as specific to his sons but rather as an instance of the general principle that everyone ought to value wisdom the most – which, I hasten to add, need not mean living *just like Socrates* (see Doyle 2012: 42–64), but conceivably could find proper expression in a number of different activities.

He notes that he questions (*eresthai*), examines (*exetazein*), and tests (*elenchein*) his interlocutors, and, in many dialogues, Plato depicts him using these tactics and also others (e.g., humbling, impersonation, exhortation, and protreptic) in service to converting them.<sup>12</sup> Socrates thus uses reason and arguments to try to change what his interlocutors value the most. How did Plato think he could achieve that?

Many scholars believe that Plato thought nothing of the sort. In this section, I review the three most common arguments that Plato is deeply pessimistic about the viability of Socrates' project of philosophical conversion. I show that these arguments are either inconclusive or rely on misinterpretations of the text.

The first argument, the Argument from Socrates' Track Record, begins from the observation that, as mentioned, there is not one unambiguous instance in Plato's corpus of an interlocutor's responding to Socrates' arguments with a decision radically to reform his way of life.<sup>13</sup> In fact, Socrates often falls far short of achieving that result. His

12. It is widely held that several of Plato's dialogues depict Socrates trying to accomplish the project that he outlines in the *Apology*. See Irwin 1995: 7; Benson 2000: 24, 26, 32; Weiss 2006: 243–53 at 243; and Rowe 2007: 89. For an opposing view, see Doyle 2012: 64–71. For an overview of Socrates' argumentative methods, see the collected papers in G. Scott, ed., *Does Socrates Have a Method?: Rethinking the Elenchus in Plato's Dialogues and Beyond* (University Park: Penn State University Press, 2002). In this paper, I am concerned with Socrates' effects only on those interlocutors whose values are askew and thus stand in need of conversion, as Socrates outlines it in the *Apology*, and more specifically on those whom Socrates in some way tries to convert in the dialogues. There is room for disagreement about who falls into this group, and my overall argument does not depend on any one particular set of members. Nevertheless, I tend to consider it as including the following interlocutors: Euthyphro, Alcibiades, Clinias, Charmides, Critias, Laches, Nicias, Hippias, Ion, Hippothales, Lysis, Euthydemus, Dionysodorus, Meno, Protagoras, Phaedrus, Gorgias, Polus, Callicles, and Thrasymachus.

13. One might suppose the *Phaedrus* to be an exception, but, as Yunis 2011: 4 rightly notes, even by the end of the dialogue, Phaedrus remains foremost committed to gaining rhetorical fame and expertise. Belfiore 2012 thinks that, by the end of the *Alcibiades*, Socrates has convinced Alcibiades that his true desire is to gain "the power to rule correctly in the city, by imparting excellence to the citizens" (51), but this overlooks that, even then, Alcibiades' primary motivation is still to acquire fame and power, and becoming excellent, for him, is valuable just as a vehicle to those. Clinias also might be thought

interlocutors tend to become angry and impatient, frequently refusing to continue the discussion or attempting in some way to sabotage or shorten it. On the basis of this evidence, some scholars conclude that Plato thinks that Socrates' project is a failure.<sup>14</sup>

Suppose, however, that Plato believed that convincing a person to dedicate himself to the pursuit of wisdom takes significant time, especially in an age, not unlike our own, when success for most people means money, power, and fame. If so, then he would not be inclined to portray Socrates converting an interlocutor to the life of philosophy by the end of any dialogue, since he would not believe that such change realistically could result from just one conversation. Is there any evidence for attributing this belief to Plato?

Yes. Consider the spot in the corpus where Plato most directly describes a case of value transformation, the ascent in *Republic* 7. The passage is notoriously difficult to interpret, but here I rely only on two points.<sup>15</sup> The first is that ascending changes the prisoner's values. At the start, he is deeply confused about what is of ultimate importance, believing that "the truth is nothing other than the shadows" that flicker across the cave wall (515c). By the end, he sees for the first time the sun, which represents the form of the good, and as a result his values are

---

to be an exception, but, as Collins 2015 argues, Clinias' vow at 282d2–3 is ambiguous. The πάνυ μὲν οὖν may indicate less of a wholehearted agreement to pursue philosophy and more of an ambivalence: "a witness would not be able to determine exactly what it is that Clinias is affirming and how far his commitment goes" (97). Additionally, Socrates' first protreptic goes no further than presenting wisdom as of the highest instrumental value, valuable for the sake of knowing how to use things rightly, and thus, even if Clinias' agreement were wholehearted, it is not obvious that it would indicate that he now values wisdom the *most*. As for his behavior later in the dialogue, see Crito's disbelief that it was actually Clinias who spoke so impressively, and Socrates' confession that he may in fact be misremembering (290d–291a; cf. Chance 1992: 123–124).

14. For scholars who infer that Socrates' project is a failure from the fact that Socrates is never shown to convert an interlocutor outright, see Beversluis 2000: 34, Nehamas 1998: 66 and 1999: 70–71, G. Scott 2000: 174, and Blondell 2002: 125.

15. For general discussion, see, e.g., Annas 1981: 242–271, Wilberding 2004, and Barney 2008.

transformed: he regards as unimportant the honors and prizes that he once admired (516c–d), and now he understands what is really “the cause of all that is correct and beautiful in everything” (517b).

The second point is that the ascent occurs not all at once but rather in stages: first the prisoner is unshackled; then he is made to stand up, turn around, walk, look towards the light, and begin ascending. Finally he sees the sun (515c–d). In fact, what is most salient about the image of an ascent — what plausibly led Plato to choose it rather than some other image to illustrate the process of value transformation — is that it consists of levels that are progressed through over time. It would thus appear that Plato does consider radical change to a person’s values not to happen all at once but rather to require advancing through distinct stages.<sup>16</sup> And if that is so, then we can explain why Plato never depicts Socrates as reforming his interlocutor’s way of life by the end of any dialogue *without* having to suppose that he regarded Socrates’ project to be a failure. The explanation is just that Plato believed that such radical change requires more than the length of conversation that reasonably could be depicted in a dialogue. There is additional evidence for this point in the *Apology* and *Gorgias*. In both dialogues, Socrates insists that a longer discussion eventually would persuade his audience. His jurors would be persuaded to acquit him (*Ap.* 37a–b), and Callicles to reform his way of life (*Grg.* 513b–c). Plato seems to hold the reasonable belief that persuasion about matters of importance takes considerable time.

It is thus too quick to infer from Socrates’ depicted track record that Plato considers his project of philosophical conversion to be a failure. An alternative explanation is available, one that, to my mind, we have every reason at this point to think is equally plausible. I shall argue soon that we should favor it. For Plato does in fact regard Socrates’ project as meant to happen in distinct stages over time, and the

16. Cf. G. Scott 2000: “Plato does not seem to allow the possibility of a sudden, complete ‘conversion’ of human character” (167). That Plato thinks value transformation takes significant time is supported also by the seed-sowing metaphor at *Phaedrus* 276b–277a.

dialogues, in large part, depict only the first stage, at the end of which an interlocutor is not meant to be fully converted to valuing wisdom the most.<sup>17</sup>

The second argument, the Argument from *Republic* 7, relies on a particular reading of a well-known passage, in which Plato has Socrates criticize an unnamed “questioner” for causing young people harm by refuting them.

And then a questioner comes along and asks someone of this sort [i.e., someone properly brought up], ‘What is the fine?’ And, when he answers what he has heard from the traditional lawgiver, the argument refutes him [ἐξελέγχη], and by refuting him often and in many places shakes him from his convictions, and makes him believe that the fine is no more fine than shameful, and the same with the just, the good, and the things he honored most. ... And so, I suppose, from being law-abiding he becomes lawless (538d–539a).

The passage continues:

So, if you don’t want your thirty-year-olds to be objects of such pity [for turning lawless], you’ll have to be extremely careful about how you introduce them to arguments. ... And isn’t it one lasting precaution not to let them taste arguments while they’re young? I don’t suppose it has escaped your notice that, when young people get their first taste of arguments, they misuse it by treating it as a kind

17. It will not do, then, to object that, while *Plato* may think that radically reforming a person’s values takes significant time, he depicts *Socrates* trying to accomplish this in just a single conversation, and thus *Socrates* must believe that conversion could happen relatively quickly — a false belief by *Plato*’s lights. What the objection misses is that, for the most part, *Plato* does not depict *Socrates* directly or immediately trying to change his interlocutors’ values. Cf. Szlezák 1999, who notes that, while the dialogues are meant to portray ascent, they “always illustrate only one section of the ascent and make the intentional limitations of the process very clear” (61).

of game of contradiction. They imitate those who've refuted them by refuting others themselves .... Then, when they've refuted many and been refuted by them in turn, they forcefully and quickly fall into disbelieving what they believed before (539a–b).

Many scholars think that the criticism in this passage applies to Socrates. It is not difficult to see why. Socrates routinely exposes young people to examination, both by refuting them and by refuting others in their presence. Worse, he seems implicated on a second score: a group of young people have in fact taken to imitating him by refuting others themselves (see *Apology* 23c). It can thus seem that, as Martha Nussbaum writes, “the characterization of the practice found dangerous points unambiguously to Socrates. ... [The] identification [is] beyond reasonable doubt.”<sup>18</sup> If that is correct, then the passage would go a long way towards showing that Plato *is* deeply pessimistic about Socrates' ability to use arguments to improve people, since it would amount to Plato's condemning Socrates for causing more harm than benefit — specifically, for instilling in young people a moral cynicism, such that they lose conviction in the fundamental claims of conventional morality.

Before continuing, it is worth noting how well this interpretation fits with a familiar way of reading the *Republic*, namely as a parting of ways between Plato and Socrates on the topic of moral education. The *Republic* begins in book 1 with Socrates' spectacular failure to convince Thrasymachus of the value of justice. Argument seems to achieve no good. Then, in books 2 and 3, a way of moral education is proposed that centers on art and gymnastics. Argument is wholly absent. Why? A common answer is that Plato now believes — or perhaps always

18. Nussbaum 1980: 88. She continues: “Plato charges his teacher (ironically, in his teacher's own *persona*) with contributing to moral decline by not restricting the questioning-process to a chosen, well-trained few.” Cf. Vlastos 1988, who writes that Socrates gives “premature exposure” to inquiry about right and wrong (100); and Nehamas 1999: Socrates teaches argument “to very young men”, which risks producing in them an “agnosticism or even cynicism” (60–61).

believed, but is now making clear — that argument is ineffective at the task of improving people, let alone reforming their ways of life.<sup>19</sup> On this reading, then, the passage from book 7, quoted above, is simply a further step in the same direction: not only is argument ineffective — it is also damaging.

But is the passage a criticism of Socrates? (Let us leave aside that, if so, we would have Socrates criticizing *himself* for causing harm, and without any apparent awareness of that: after all, Socrates is the one voicing these complaints.) Two passages elsewhere should make us hesitate. In the *Apology*, Plato has Socrates argue that he has *not* corrupted the young (33d–34b), while also having Socrates admit that he refutes all types (23b; cf. 33a, 37d) and that young people now imitate him (23c) — the exact sort of behavior that *should* be corrupting, if the current interpretation of the *Republic* 7 passage is correct. Further, in the *Sophist*, a dialogue considered to postdate the *Republic*, Plato is enthusiastic about the *benefits* of refutation. It is “the principal and most important kind of cleansing”, and it is a necessary preliminary to learning: “the soul ... won't get any advantage [ὄνησιν] from any learning that's offered to it until someone shames it by refuting it ...” (230c–d). Not only is there no mention of any need to restrict refutation to adults — it would seem that, by the *Sophist's* lights, it is Socrates' *young* interlocutors who especially need to be refuted. Otherwise whatever learning they acquire will not benefit them.

Now, it would be unpersuasive to argue that, because of Plato's views elsewhere, the *Republic* 7 passage must not be saying what it seems to say, i.e., that those who refute young people are at fault. But Plato's views elsewhere do motivate a reconsideration of the passage. And what one notices upon reconsideration is that not refuting young people is presented only as *one* precaution a person might take to avoid the risk of inducing moral cynicism: “... isn't it one [μία] lasting precaution not to let them taste arguments while they're young?”<sup>20</sup> The

19. See, e.g., Klosko 1983, D. Scott 1999: 15, and Woolf 2000: 32–40.

20. On the phrase ‘taste arguments’ (τῶν λόγων ... γεύεσθαι), cf. *Alc.* 1 114a.

passage thus excuses a person who takes some other suitable precaution. The question is whether Socrates does.

The answer is yes. Whenever he refutes young people, as well as whenever a young person is a bystander to a refutation, Socrates consistently acts to ensure that any growing moral cynicism is quickly uprooted. Far from trying to unsettle his interlocutors' convictions in the fundamental claims of conventional morality, he insists that such claims are true: that virtue is good, vice is bad, and virtue and vice are opposites.<sup>21</sup> Further, in places where an interlocutor's definition implies that virtue is not good, Socrates responds by insisting that the implication alone refutes the definition.<sup>22</sup> Such insistence might mean little if young people were distrustful of Socrates, but Plato depicts just the opposite. He depicts young people trusting and admiring Socrates, even consulting him for important life guidance.<sup>23</sup> Socrates' repeated insistence that the fundamental claims of conventional morality are true, then, is precisely the sort of conversational move that would act as a suitable precaution against moral cynicism arising in them. Moreover, it makes sense that the recommended precaution in the passage — *just don't refute young people* — is maximally cautious. The risked harm is serious, and Plato's advice is for practitioners of argument in general. But there are exceptions to general advice, and Socrates' case is one of them. He is "extremely careful" in an alternative way.<sup>24</sup> It is thus mistaken to use *Republic* 7 to motivate a Platonic pessimism about

21. Virtue is good: *Charm.* 159c, 161a, 163e, 175b, 176a; *Laches* 190b–c; *Alc.* 1 109c, 115e, 116c–d, 133b, 134a–135c; *Euthyd.* 279b; *Prot.* 332a–b, 349e; *Grg.* 493d, 498c, 504d–e, 506e–507a, 507a–c, 527d; *Meno* 73b–c, 87d–e, 98e; *Rep.* 1.353e–354a. Vice is bad: *Alc.* 1 115d, 135c; *Euthyd.* 281d–e; *Prot.* 332a–b; *Grg.* 469b, 470e, 472d–473b, 477e, 479c, 498c, 505b, 507a–c, 521b, 522e; *Rep.* 1.353e–354a. Virtue and vice are opposites: *Prot.* 332d–e, 359e, 360d; *Grg.* 507a–c.

22. *Charm.* 160d, 161b, 175b, 176a; *Laches* 192d, 193d; *Prot.* 350b. Cf. *H. Mi.* 376b–c, where Socrates refuses to believe that the one who voluntarily acts unjustly is the good person, despite the argument seeming to show that.

23. See, e.g., *Lach.* 180e–181a, *Alc.* 1 124b, *Prot.* 313c, *Lys.* 205a, and *Phdr.* 269c–d.

24. In fact, he is extremely careful in more than one way. He also refutes those who target people's inherited, fundamental moral convictions, such as Callicles and Thrasymachus.

Socrates' project of philosophical conversion. The passage is not a criticism of Socrates.

The third argument is more complicated than the previous two. It begins from the thought that, at least in the middle dialogues, Plato recognizes that certain desires have the ability to determine our evaluative beliefs. It then seeks to show that Socrates is unable to change those desires, and so is unable to change the beliefs that they control.

I consider just one version of this argument here, a version put forward by Raphael Woolf. I call it the Argument from *Erōs*. Woolf observes that, at best, Socratic argument can only ever convince a person that he should adopt some belief because the reasons support doing so. What it cannot do is instill in him the disposition to change his beliefs in accordance with the reasons. What is required for *that* is that one's desires be rightly directed — more specifically, that a person have *erōs* for "logical consistency" (28; cf. 32). If a person's *erōs* is misdirected — if it is for anything *besides* that — then he will cling to his favored evaluative beliefs even if his reasoning shows them to be false. Thus, no amount of argument will be effective at changing them.

Is it true that Plato thinks of misdirected *erōs* as an obstacle that Socrates cannot overcome? He certainly thinks of it as an obstacle. As Woolf notes, Plato has Socrates diagnose Callicles' *erōs* as the reason why, despite being unable to refute Socrates' arguments, he refuses to change his beliefs (*Grg.* 513b–c). This passage is important, as it is the only spot in Plato's corpus where Socrates diagnoses the reason why an interlocutor has failed to be convinced by his arguments in the here and now. However, immediately after saying that Callicles' *erōs* is to blame, Socrates insists that Callicles *would* be persuaded if they examined the same matters "often and in a better way" (513c). Apparently, then, *erōs* is not an insurmountable obstacle for Socrates. To get the better of it, though, a better sort of argument is needed (e.g., a less confrontational one), and the issue in dispute must be repeatedly considered.

It is possible, of course, to regard Socrates' optimism here as naive — tempting, even, to suppose that Plato intends for us to see it

as such, given that Callicles shows few signs of budging.<sup>25</sup> However, Plato is committed to the thought that repeated arguments have a persuasive power that isolated arguments lack — that repeated arguments succeed where isolated arguments fail, including in disputes about the best way to live.<sup>26</sup> In fact, so much would seem to be implied by the extended analogy in the *Gorgias* between the medical craft and the true political craft (see especially 521a–522b). Often a single medical treatment is not enough to restore health to the body. Likewise, Plato thinks, with arguments and the health of the soul: repeated doses are needed. It is thus implausible that Plato means for us to see Socrates' optimism as misguided.

There is a further reason for supposing that *erōs* is not an insurmountable obstacle for Socrates. It concerns the timing of his first interaction with Alcibiades. In the *Alcibiades*, Socrates says that, though he has been observing Alcibiades for many years, he is deciding to talk with him now because only now is he ready to listen (*Alc.* 1 105e–106a; cf. 124c). Remarkably, what has made Alcibiades ready to listen is his ambition (105a), ambition that Socrates later characterizes as *erōs* for fame (124b4–6). Socrates' claim, then, is that, before Alcibiades developed such *erōs*, conversation with him would have been “pointless” (μάτην, 105e); under the influence of *erōs*, however, he is ready to listen (vῦν γὰρ ... μου ἀκούσαι, 106a).<sup>27</sup>

Now, if misplaced *erōs* — i.e., to go along with Woolf, *erōs* for anything besides logical consistency, such as Alcibiades' *erōs* for fame — poses insurmountable obstacles for the ability of argument to

change a person's values, then Socrates' claim that *erōs* has disposed Alcibiades favorably to the conversation would make little, if any, sense. For Socrates' purpose in the dialogue is explicitly to convert Alcibiades to virtue and the cultivation of himself through philosophical conversation (132b–133c, 134b–135e; cf. 124b–c, 127e). At bottom, then, Socrates' claim is that, at least in Alcibiades' case, a misplaced erotic attachment is a prerequisite for being improved by means of argument — the opposite of the claim that *erōs* condemns such a project to failure.<sup>28</sup>

So we return to the start. If, as I have argued, there is no good reason to suppose that Plato rejects Socrates' project of philosophical conversion, how does he think that it can succeed? How can Socratic argument change what a person values the most? I turn to this question in the next two sections.

## 2. Indirect route

In the *Apology*, Socrates outlines a project of using reason and argument to persuade his interlocutors to value wisdom the most. How is that project meant to go? In this section, I present the first stage of it. I argue that Socrates' immediate aim with his interlocutors is typically to motivate them to pursue wisdom only for its instrumental value, and, further, that such a limited aim is explained by Plato's conviction that a more ambitious attempt is likely to fail.

25. Woolf, e.g., refers to Socrates' optimism as an “act of faith” (31). See also Klosko 1983, who likewise interprets it as a statement of “faith,” one that distances Plato from Socrates: “as Plato is well aware, Socrates is fighting against forces that are too powerful for him ... [Socrates] is inevitably destined to lose” (593, 586). Cf. Plochmann and Robinson 1988: 201.

26. See *Ap.* 37a–b, *Tht.* 177a–b, and *Symp.* 221e–222a.

27. It is disputed whether the *Alcibiades* is written by Plato, though scholars are increasingly in favor of its authenticity. See Denyer 2001: 14–26 and Jirsa 2009 for, to my mind, convincing rebuttals to authorial skepticism, and Smith 2004 for an opposing view.

28. Of course, Socrates' claim would mean little if it were shown to be false — if, e.g., Plato were to depict Alcibiades' *erōs* nonetheless interfering with Socrates' attempts to persuade him. What Plato depicts, however, is Socrates convincing Alcibiades that, unless he gains wisdom, he will not be able to acquire the fame that he intensely desires (see especially 105e, 108e–109a, and 116d–118b). Socrates succeeds, then, not *despite* Alcibiades' *erōs* but *because* of it: he succeeds by exploiting its motivational power to turn Alcibiades towards philosophy. One might object still that the *historical* Alcibiades is a failed case, and thus Plato must have meant for us to regard Socrates' confidence in his ability to persuade Alcibiades as misguided. There is a real question, however, about what Plato means for us to conclude about Socrates from his many depictions of Alcibiades. See Sheffield (forthcoming) for the argument that Plato depicts Alcibiades making moral progress in the corpus, and thus that, by Plato's lights, Socrates was a considerable success with Alcibiades.

Suppose you want to convince a person that she should value  $x$  more than anything that she currently values. One natural way to proceed would be to provide her with reasons why the things that she values now are not so important, and why, in contrast,  $x$  is of ultimate importance. Call such an approach *direct*, in virtue of its attempting to argue a person into a new position (e.g., the position of valuing  $x$  more than anything else) by presenting her with reasons to accept it.

For two reasons, Plato must have thought that a direct approach would face serious problems when the task came to persuading a person to change what he values the most.

The first concerns the drama of the *Gorgias* and *Republic*. When Socrates contests the ways of life of Polus, Callicles, and Thrasymachus — when he provides them with reasons why, roughly, the life of philosophy is better than the life of politics — he meets with stubbornness and defensiveness. None are persuaded. Importantly, these interlocutors are among Socrates' most intransigent. Equally as important, they are the only interlocutors whose conceptions of the good Socrates attempts to change directly.<sup>29</sup> Most often he aims to convince his interlocutors only that they are ignorant about virtue. However, even if successful, there is no reason why this would unsettle their conception of the good. A person can realize that he is ignorant about virtue without that at all causing him to lose confidence in his belief about what the good is.<sup>30</sup>

29. In each case, Socrates specifies that the topic in dispute is the best way to live.

To Polus: "... the heart of the matter is that of recognizing or failing to recognize who is happy and who is not" (*Grg.* 472c–d). To Callicles: the disagreement concerns "the way we're supposed to live" (*Grg.* 500c), and Socrates' explicit goal is to convince him to "choose the orderly life" (*Grg.* 493c). To Thrasymachus: the dispute is about "which whole way of life would make living the most worthwhile for each of us" (*Rep.* 1.344e). With no other interlocutors is the goal so overtly and directly to change their strongly held beliefs about what constitutes the happy life.

30. He simply may not care that he is ignorant about what virtue is (as a whole or any specific virtue): so long as he is convinced that he is pursuing what is good, e.g., wealth, then he can rest easy. Or perhaps he does care, and, after investigating further, it comes to seem to him that virtue is such that being virtuous is incompatible with acquiring the good as he sees it. But that is

The second concerns a problem with trying to decide on the basis of reasons what to value the most. In book 9 of the *Republic*, Plato considers how, in general, we answer questions about value. He thinks that each of us uses a certain "criterion" (κρίτηριον, 582a) or "instrument" (ὄργανον, 582d) when judging things to be valuable (cf. ἐπὶ ἡνείκε ... καὶ ἔψεγεν, 582e). In particular, we use what we value the most, and we judge things to be valuable in light of it. The lover of profit (ὁ φιλοκερδῆς), e.g., makes judgments based on wealth and profit (πλούτῳ καὶ κέρδει ... ἐκρίνετο), i.e., he judges things to be valuable, in the end, insofar as they promote wealth. So, too, with the lover of honor (ὁ φιλότιμος): his evaluative judgments are based on honor and victory and courage (τιμῆ τε καὶ νίκη καὶ ἀνδρεία).

The problem arises when trying to convince a person on the basis of reasons to change what he values the most. To do so, he must evaluate those reasons as showing that something else is more valuable. But this would seemingly never happen, since what he values the most is the final arbiter for him of what is valuable. The result is a sort of evaluative inertia: the lover of profit will dismiss the prospect of anything's being more valuable than money, since, for him, things are valuable in the end just insofar as they promote gaining money. And likewise in other cases: what a person values the most will act as the final arbiter for him of what is valuable, and thus he will not evaluate the reasons as showing that some object is more valuable than what he currently most values.<sup>31</sup>

just as likely to be seen by him as a refutation of his newfound conception of virtue as of his conception of the good — *more* likely, even, as Plato seems to think that, of all one's beliefs, it is one's belief about what the good is that one holds most strongly. All other beliefs conform to it (see especially *Rep.* 8.553c–d and 560a–e; cf. *Phil.* 12a and *Rep.* 9.582d–e).

31. In Plato's corpus, evaluative inertia is expressed most forcefully by Protarchus: "To my mind, absolutely [πάντως], pleasure wins and always will win [as the good]" (*Phil.* 12a). It is found also with Polus, who rejects out of hand arguments that conflict with valuing power the most (*Grg.* 473e, 480e); with Callicles, who does the same in response to arguments that conflict with valuing pleasure the most (*Grg.* 494a–b, 511a, 513c, 521b); and with Thrasymachus, who acts in the same manner when Socrates argues against the value of power and wealth (*Rep.* 1.350d, 354a). In each case, what seems to be happening

What else can be done? What would it mean to experience a conversion here — here, where rational persuasion no longer seems viable?<sup>32</sup> One natural answer is that what would change a person is

is what *Republic* 9 describes: the interlocutor uses what he values the most as his instrument of evaluative judgment, thus blocking off the possibility of judging something on the basis of reasons to be more valuable than it (cf. the phenomenon of “biased assimilation effect” in Lord, Ross, and Lepper 1979 and Corner, Whitmarsh, and Xenias 2012). This view does not require that reason (as a faculty or as a part of the soul) is incapable of deciding on its own what is of ultimate value. It requires only that deductive reasoning is. For a view that closely resembles the one that I ascribe to Plato, see Frankfurt 2004: 23–26 and, to a lesser extent, Frankfurt 1992: 15–16. It is interesting to note that the problem of evaluative inertia can arise both in contexts of value subjectivism (as with Frankfurt) and value objectivism (as with Plato). There are ways to avoid the problem, but each requires supposing that people settle questions about what is valuable in ways other than Plato suggests in *Republic* 9. For example, one can imagine settling this question by first agreeing that the good must meet certain independent criteria, and then arguing that some particular activity or object of pursuit best meets those criteria (cf. *Phil.* 20b–22c, *Nicomachean Ethics* 1.7). For this strategy to succeed, however, one must hold fast to the initial agreement, even when it leads to a result that conflicts with one's prior conviction about the good, and Plato seems to think that, of all one's beliefs, it is one's belief about what the good is that one holds most strongly (and thus clings to when other beliefs conflict with it). Another strategy is to deploy a debunking argument — to argue, e.g., that a person's belief about what is valuable is overly influenced by her upbringing or society. Such an argument, however, does not aim to instill new values but only to disrupt what a person currently values (the only debunking arguments in the corpus that I know of, interestingly, are made by Socrates' interlocutors: see *Grg.* 482e–484c, *Rep.* 1.338c–339a, and *Rep.* 1.343b–334c). An alternative is to use a perceptual or affective experience to persuade someone to change what she values the most. I argue that, indeed, this is Plato's and Socrates' solution to the problem of evaluative inertia, though, importantly, one is meant to undergo the experience by doing philosophy.

32. Cf. McDowell 1998: “[It is a] massively implausible implication that someone who has not been properly brought up ... can be induced into seeing things straight by directing some piece of *reasoning* at him. On the contrary, reasoning aimed at generating new motivations will surely stand a chance of working only if it appeals to something in the audience's existing motivational make-up ... and the trouble ... is that there may be no such point of leverage for reasoning aimed at generating the motivations that are characteristic of someone who has been properly brought up. What it would take to get such a person to consider the relevant matters aright, we might plausibly suppose, is ... something like conversion” (101–102). Similarly, I have argued that trying to reform a person whose values are fundamentally mistaken by “directing some piece of *reasoning* at him” looks to be futile, since such reasoning, to

a perceptual or affective experience.<sup>33</sup> I shall argue that this is exactly what Socrates aims to cause in his interlocutors, but not at all by means that we would regard as mystical or mysterious. He aims to cause it — or, more specifically, to provide the occasion for it to happen — by motivating his interlocutors to keep doing philosophy.

In the next section, I consider how doing philosophy could cause such an experience. First, we need to inquire further into Socrates' typical activities.

If the above argument is at all correct, then Socrates, to convert his interlocutors, must not contest the values that they regard as fundamental, while *somehow* still affecting them such that they move a step closer to valuing wisdom the most. As it happens, that is exactly what we often observe him doing, or in any case trying to do.

To start, consider Socrates' tactic in the *Phaedrus*. Socrates' avowed aim is to convince Phaedrus that “unless he pursues philosophy properly he will never be able to make a proper speech on any subject either” (261a). He thus presents doing philosophy as valuable for the sake of Phaedrus' acquiring rhetorical expertise. Socrates makes a similar move in the *Alcibiades*. He convinces Alcibiades that, unless he gains self-knowledge, no fame or influence will ever come to him, thus motivating Alcibiades, like Phaedrus, to do philosophy for the sake of an external result that he highly prizes. So too in the *Lysis*: Socrates presents wisdom as useful for Hippothales insofar as gaining it will increase his chances of wooing Lysis (210d–e), for Lysis insofar as gaining it will convince his parents to allow him to do whatever he pleases (207d–210c). In all of these cases, Socrates leaves fixed what

be successful, must “appeal[] to something in the audience's existing motivational make-up” — such as what she currently values — and in certain cases, there is nothing to which one can appeal that will do the desired work.

33. One might suspect that such experiences are wholly out of place in Platonic philosophy. But see *Republic* 9.582a, where Plato places experience (*empeiria*) alongside reason (*phronēsis*) and argument (*logos*) as the things by which we should judge, if we are to judge well. See also Edmonds III 2017 for a reading of the ascent in the *Symposium* where the catalyst for change is not the learning of some articulable doctrine but rather a revelatory experience.

his interlocutor fundamentally values — rhetorical expertise, fame and influence, a relationship with Lysis, and the ability to do whatever he pleases — and focuses the entirety of his efforts solely on convincing the interlocutor that, to get what he wants, he needs to begin doing philosophy.

In other cases, Socrates aims to damage something that his interlocutors care about, in such a way that, to repair it, they must do philosophy. I mean for this to cover typical cases of Socratic refutation. Consider, for example, Socrates' refutation of Protagoras. Socrates shows Protagoras to be ignorant of how virtue is teachable and what virtue is (333a–b, 361a–c). Importantly, these are not just any old topics. They are topics that *especially Protagoras* ought to know, given his claim to teach virtue (318d–320c). When he is shown *not* to know them, it impugns his credibility as a reliable educator. Meanwhile potential students and colleagues look on. Protagoras' social standing would thus be lowered — so, too, his self-esteem, to the extent that he takes pride in being a competent educator (and not just in having a reputation as such). In turn, he would be motivated to raise his social standing and self-esteem. Moreover, since his ignorance of virtue is what caused these to be lowered, the way to raise them would be to acquire the knowledge that he was shown to lack. Thus, by refuting Protagoras, Socrates gives him a reason to strive to attain wisdom (in this case, knowledge of virtue) right now, namely that doing so is instrumentally valuable for repairing the things that he cares about.<sup>34</sup>

34. One might suspect that Protagoras already stands in the proper relation to wisdom — he is, after all, a self-professed teacher of virtue who thinks that “wisdom and knowledge are ... the most powerful forces in human activity” (352d) — and thus that Socrates cannot really be trying to motivate *him* to value wisdom more. But Protagoras' behavior in the dialogue betrays him. When he first senses that Socrates has exposed him in an inconsistency, he tries to exit the discussion; he “plays coy, claiming the argument was too hard for him to handle” (333d). Later, he refuses to discuss matters by means of brief questions and answers, claiming that, if he were to do so, it would tarnish his reputation as “a name to be reckoned with among the Greeks” (335a). This is not the behavior of someone who values wisdom the most. It resembles more the behavior of someone who values the appearance or reputation for

A similar account can be given for any refutation that occurs in public and concerns a topic that the interlocutor must know to perform his social role well (thus including refutations of Critias in the *Charmides*; Laches and Nicias in the *Laches*; Ion in the *Ion*; Euthydemus and Dionysodorus in the *Euthydemus*; Hippias in the *Hippias Minor*; Meno in the *Meno*; and Gorgias in the *Gorgias*). Refutations in private — for example, of Euthyphro in the *Euthyphro* — work in like manner, except that, due to the absence of onlookers, the interlocutor's social standing is less vulnerable. If he is to be motivated to do more philosophy, it is his self-esteem that must provide it.

Socrates' immediate aim with many of his interlocutors, then, is to motivate them to pursue wisdom only for its instrumental value. He does not contest what they care about (e.g., rhetorical expertise, social standing). He leaves that fixed, and he persuades them that, to stand in a better relation to it, they must do philosophy. Of course, there is more to be said in each case. I have shown only how a fuller argument might go, i.e., an argument that considers exegetical details at length.<sup>35</sup> I want to emphasize, though, that Socrates' limiting his aim to motivating just the instrumental pursuit of wisdom is exactly what he should be doing, if, as I argued, a more ambitious, direct approach is likely to fail.

On my account, then, Socrates *is* doing what he *should* be doing.<sup>36</sup> It is not, however, as if motivating people to pursue wisdom just for its

wisdom. This point is compatible with the view that Plato thought highly of Protagoras.

35. In any case, the point that Socrates aims to motivate his interlocutors to pursue wisdom as a means is supported by accounts on which Socrates is trying to *shame* his interlocutors, thereby motivating them to undertake further philosophical inquiry as a means to rectifying their shameful condition. See, e.g., Brickhouse and Smith 1994: 25 and 2000: 58–59.

36. At least most of the time. In the *Gorgias* and *Republic 1*, Plato depicts him trying a more ambitious, direct approach — on my account, a mistaken strategy. My account thus implies that Socrates does not always behave in the way most advantageous to converting his interlocutors. Plato nonetheless had good reason to depict Socrates behaving in this way. It dramatizes the problem of evaluative inertia, and thus fills out the explanation for why Socrates tends to adopt a more sophisticated, indirect strategy.

instrumental value is Socrates' ultimate aim — it is not as if *that* is what he means by saying that he “go[es] around doing nothing other than persuading both young and old among you not to care for your body or your wealth in preference to or as strongly as the best possible state of your soul” (*Ap.* 30a–b).<sup>37</sup> No, what he wants is to *transform* the values of his interlocutors — as I put it above, to cause them to value wisdom, the best possible state of their souls, more than anything else. The question, then, is this: How can pursuing wisdom only for its instrumental value lead to valuing it the most? How can doing philosophy for the sake of advancing one's career, or for the sake of repairing one's social standing, or for the sake of wooing an attractive boy, or in any case for the sake of some exterior, non-philosophical goal — how can doing philosophy for the sake of *that* cause one to come to dedicate one's life to wisdom, to value *it* the most?

It is worth noting that the structure of this question is familiar. It is a question of how the means can become the end.<sup>38</sup> It is worth noting, too, that, however this may happen, it frequently *does* happen in the philosophy classroom. For the sake of satisfying a curricular requirement, a student takes a philosophy course, and, by the end of it, she finds that her priorities are transformed, such that she is now strongly motivated to gain something like philosophical wisdom, just for itself. This phenomenon is so common that the observation of it may seem banal. But its reality suggests that Socrates is not misguided in trying

37. Is this claim in tension with what, I have argued, Socrates in fact often does, i.e., allow people to continue caring the most about their bodies or wealth while convincing them that they should care more about their souls as in some way a means to them? No, I think, if we accept that persuasion (*peithōn*, *Ap.* 30a8) can happen in stages, such that I can be considered to be persuading you of *x* even if my immediate aim is only to persuade you of *y*, so long as I am persuading you of *y* so that eventually you will come to believe *x*.

38. More specifically, how the means can become the end in a context where the end would not be chosen for itself. For a useful discussion of mental states that can be brought about only as “by-products of actions undertaken for other ends”, see Elster 1983: 43–60.

to motivate his interlocutors, at first, to *just start doing philosophy*, whatever it takes.<sup>39</sup>

In the next section, I explore one answer as to what would cause such a conversion. First, I want to offer a point of reentry to my overall argument for anyone inclined to doubt my claim above that typically Socrates tries to motivate his interlocutors to pursue wisdom just for its instrumental value. I began by observing that, as scholars widely recognize, there is not one unambiguous instance in Plato's corpus of Socrates' radically reforming his interlocutor's way of life by the end of any dialogue. In every case, it seems as if the interlocutor's fundamental values are unchanged. But Socrates consistently urges his interlocutors to *keep doing philosophy*.<sup>40</sup> We can thus ask a question similar to the one above: What is it about doing more philosophy that Socrates expects to be transformative? What is it about doing more philosophy that Socrates expects to cause his interlocutors to come to value wisdom the most — *especially* when, as I have argued, they will seemingly never do so on the basis of reasons in support of wisdom's ultimate value?

### 3. Pleasure and the good

In the previous section, I argued that Plato depicts Socrates limiting his behavior: despite wanting ultimately to convert his interlocutors to valuing wisdom the most, Socrates typically attempts to persuade them only of wisdom's instrumental value, i.e., he does not aim directly or immediately for his goal. If this is not simply a mistake on Socrates' part, then something must explain it. What explains it, I argued, is the problem of evaluative inertia. In reasoning about what to value, people use what they currently value the most as their evaluative standard, judging things to be valuable in the end based on whether they

39. Cf. the example in MacIntyre 2007: 188 of the child who, initially motivated to play chess only by the promise of candy, eventually, by playing chess, comes to value it for itself.

40. See, e.g., *Euthyphr.* 15c, *Charm.* 169d, *Laches* 194a and 201a, *Prot.* 314b and 361d, *H. Ma.* 295a–b, *Grg.* 461b, and *Euthyd.* 288c.

promote standing in a better relation to it, e.g., acquiring more of it or being closer to it. Thus, people will not reason that something is more valuable than what they currently value the most. If Socrates is to have any success in converting his interlocutors, then, he cannot rely on them to draw the inference that their fundamental values are mistaken.

At first, this claim appears to be in tension with Socrates' habit of encouraging his interlocutors to keep doing philosophy — in tension, too, with his confidence that further arguments will succeed at persuading interlocutors who seem to be intransigent (*Grg.* 513c–d). If procedures of evaluative reasoning block a person from inferring that his current set of values is mistaken, then how is doing more philosophy to be effective at converting him? In this section, I argue that it can be effective not by leading a person to infer that he must change his way of life but rather by causing the experience of intellectual pleasure, an experience capable of reshaping a person's evaluative beliefs.

To start, Plato thinks that doing philosophy causes more pleasure than any other activity. In *Republic* 9, Socrates ranks the pleasures of the philosopher, lover of honor, and lover of money. He concludes: "of the three pleasures, the most pleasant [ἡ ἡδίστη] is that of the part of the soul with which we learn, and the one in whom that part rules [i.e., the philosopher] has the most pleasant life [ὁ βίος ἡδιστος]" (583a1–3).<sup>41</sup> Plato depicts Apollodorus testifying to this fact in the *Symposium*: "my greatest pleasure comes from philosophical conversations" (173c).

Now, experiencing the pleasure of philosophy could motivate a person to do it no longer only for the sake of an external good, e.g., a reputation for intelligence, but also for the sake of an internal good, i.e., the pleasure of philosophy.<sup>42</sup> However, that would fall short of radical

change: a person's motivations could change in this way without it reshaping what he fundamentally values.

The next step, then, is to note that Plato thinks that pleasurable experiences have the psychological effect not only of causing us pleasure. They also influence our beliefs about the good. In the *Gorgias*, Socrates claims that most people would judge a pastry chef to know better than a doctor which foods are "beneficial" (χρηστός, 464d7). Why is that? No direct explanation is in the text, but the salient fact about pastry baking is that it causes great pleasure: it "captivates with what is most pleasant in the moment" (464d). What seems to explain it, then, is that most people operate with a certain psychological tendency, namely the tendency, when something causes them pleasure, to believe that it is good. Just because his pastries cause them pleasure, people wrongly believe that the pastry chef knows which foods are beneficial.<sup>43</sup>

This psychological tendency appears also in the *Phaedo*. Socrates claims that, when we experience "strong pleasure or pain", we are "compelled to believe [ἀναγκάζεται ... ἠγεῖσθαι] at the same time that what causes such feelings must be very clear [ἐναργέστατον] and very true [ἀληθέστατον]" (*Phd.* 83c5–8). Now, the belief that the object is also good is not mentioned, but that is unsurprising: Socrates' point is about what is common to experiences of pleasure and pain. In the case of pleasure, however, it is highly plausible that the experience would instill also the belief that the object causing it is good. Socrates' concern in the context is with the beliefs that influence our habits and ways of life (cf. ὁμότροπός τε καὶ ὁμότροφος, 83d8–9) — that is, with beliefs about the good. This detail clarifies Socrates' point regarding

---

of the practice, and they can be evaluated well only by people who have achieved them before.

41. Russell 2005: 134 argues that Socrates' conclusion is only about "lives as wholes" and not "pleasant characteristic episodes" (cf. Annas 1981: 308–309). But it is about both: the most pleasant pleasure (ἡ ... ἡδίστη [ἡδονή]) and the most pleasant life (ὁ ... βίος ἡδιστος).

42. See MacIntyre 2007: 188–191 for discussion of internal and external goods. External goods are only contingently attached to a practice, and they can be achieved by alternative routes. Internal goods are specifiable only in terms

43. See similar remarks on this passage by Moss 2007: "... most people fail to distinguish between what is pleasant and what is good for them. ... When they experience pleasure, they believe that they are being benefited ... they confuse the state of being pleased with the state of being in good condition" (30–31). See also Moss 2006: "[The pastries] appear to be good, simply because they are pleasant; when someone pleases us, we think he is doing us good" (513).

pleasure: when a person experiences strong pleasure, that experience instills in her the belief that what causes the pleasure is very clear, very true, and very good.<sup>44</sup>

In the *Laws*, a similar psychological tendency is made explicit: “we feel pleasure whenever we believe that we are doing well, and likewise, whenever we feel pleasure, we believe that we are doing well [ὅποταν χαίρωμεν, οἰόμεθα εἶ πράττειν αὖ]” (657c). Here again Plato advances the thought that pleasure influences our evaluative beliefs. It is not said that pleasure instills in us the belief that what *causes* it is good, but that is implicit: if we believe that we are doing well whenever we feel pleasure, then we must believe that what causes the pleasure benefits us.

Now, in the *Gorgias* and *Phaedo*, Socrates laments the psychological tendency that, whenever something causes people pleasure, they believe it to be good. But this tendency is just as capable of turning people in the right direction as leading them astray. Moreover, it explains how the experience of pleasure from doing philosophy can transform one's fundamental values — that is, how *philosophy* can transform one's values, though not by leading a person to draw a conclusion from any premises. Philosophy can cause the experience of pleasure, and thereby instill the belief that it is good. But we can go a step further. It is a corollary of the psychological tendency in the *Gorgias* and *Phaedo* that, when something causes people pleasure, they tend to believe it is good in proportion to how pleasurable it is.<sup>45</sup> Thus, in causing people

to experience the *most* pleasure, the activity of philosophy would instill in them the belief that it is *best*, i.e., more valuable than anything else.<sup>46</sup>

What we wanted, though, was to know how doing philosophy could cause someone to *value wisdom* the most. There are two points here: that the goal is *valuing* (*epimeleisthai*; cf. *Ap.* 29d–30b, 31b, and 36c) and that specifically it is valuing *wisdom*. All I have concluded so far is that doing philosophy could cause a person to *believe* that *philosophy* is most valuable. But believing is not valuing: I believe that space exploration is valuable, but I do not value it, at least not in any way that has a significant influence on my life.<sup>47</sup> And philosophy is not wisdom: conceivably, I could highly value doing philosophy — highly value the activity, say, of navigating conceptual territory — without valuing wisdom all that much. What I value would be the activity, and the activity can come apart from its aim. Likewise, I might highly value playing chess without caring all that much about whether I win.

Nevertheless, in typical cases, one values the aim of an activity just as much as the activity itself (and sometimes even more so: “what I really want is to *win*”). Further, in *Republic* 9, what in particular seems to cause the greatest pleasure is not just *doing* philosophy but *succeeding* at it, i.e., learning: “the most pleasant [pleasure] is that of the part of the soul with which we *learn*”. If so, then it is wisdom that the pleasure from doing philosophy would cause one to believe is most important. As for the transition from *believing* to *valuing*, it is again the default to value what you believe is important — to shape your life by it in some way. Of course, in some cases that will be impossible (I am not really

---

believe it is good in proportion to how pleasurable it is. Cf. the correlation in the *Phaedo* passage of “strong pleasure or pain” with “very clear and very true”.

44. For a similar interpretation, see Ebrey 2017: “the beliefs caused in us by pleasure and pain change our way of life ... we can have certain beliefs forced on us by our experiences and thereby change what we desire and how we act ... pleasures change which things we think are good” (7–8). Cf. Woolf 2004: the “intensity of response” is a “begetter” of “the evaluative stance one adopts towards the relevant sources” (103).

45. After all, Socrates' point in the *Gorgias* is not that all those who cause at least *some* pleasure will be judged to know *equally* as much about what is good and beneficial. It is rather that causing *more* pleasure will incline people to judge that you know *more* about what is good and beneficial (thus the choice of the pastry baker, who is responsible for causing people great pleasure), and what explains that is the tendency, when something causes you pleasure, to

46. For discussion of how and why philosophy causes the most pleasure, see especially *Philebus* 51e–53c and *Republic* 9.585b–586b.

47. Cf. Helm 2001: 71. Similarly, to transition back to the Greek, one can believe that something is important without qualifying as *epimeleisthai*-ing it. Presumably, Meletus believes, as he avows (*Ap.* 24d), that the education of the youth is important; but he is shown not to *epimeleisthai* it (*Ap.* 24c–25c). See *Memoabilia* 2.4.2–4 for *epimeleisthai* functioning similarly in Xenophon.

able to include space exploration in my life), but in most cases it will be a natural outgrowth of the belief.

Now, it is one thing to say that pleasure *can* play this role in philosophical conversion, another to say that Plato *thinks* of pleasure as playing this role. Here we might hesitate. Plato is notoriously suspicious of pleasure, writing that the true philosopher “keeps away from pleasures ... as far as he can” (*Phd.* 83b); that pleasure is “evil’s most powerful lure” (*Tim.* 69d); and that the life *without* pleasure is “the most godlike” (*Phil.* 33b). Further, in the *Gorgias*, he presents pleasure as an unreliable guide to value (see especially 464a–465d; cf. *Phil.* 65c: “pleasure is the greatest impostor of all”).

But things are not so straightforward. To start, it is not as if pleasure is associated only ever with the merely apparent good: some genuinely good things are pleasurable, and the best life for humans, Plato thinks, will involve not only those things but also the pleasure of them (*Phil.* 22a–b, 27d, 63d–64a). More importantly, Plato recognizes that pleasure can play a positive role in convincing someone to adopt a better way of life. In the *Laws*, he writes that “nobody would willingly be persuaded to do [ἐκὼν ἐθέλοι πειθεσθαι πράττειν] something that does not cause more pleasure than pain”.<sup>48</sup> He then recommends that, when trying “to persuade a man to live a just and pious life”, one should argue that it is more pleasant than the alternative (663b).<sup>49</sup> Thus, Plato considers it appropriate and strategic to exploit the motivational power of pleasure when trying to convert a person to a new way of life.

On my view, though, Socrates is not trying to *convince* his interlocutors that the life of philosophy is most pleasurable. He is trying to

convert them by getting them to *experience* the pleasure of philosophy. Is there evidence that Plato considered this experience to be capable of rightly orienting a person?

Yes. In the *Phaedo*, Socrates encourages “seriously concerning” (ἐσπούδασέ) oneself with the “pleasures of learning” (ἡδονὰς ... τὰς ... περὶ τὸ μανθάνειν) — not merely with *learning*, but with the *pleasure* of it (114e). Why is the pleasure of it important? Socrates gives no explanation, so we are left to find one elsewhere in the text. And what is most salient about pleasure elsewhere in the text is that one should avoid certain instances of it because they have the effect of distorting one’s beliefs and way of life (83b–e; cf. 65a–67d). What is most salient, that is, is that pleasures influence what we believe and how we live. If we are to *pursue* certain pleasures, then, it must be because the experience of those pleasures *improves* our beliefs and way of life — exactly the claim that we are concerned to secure.<sup>50</sup>

I do not mean to give the impression that Socrates’ strategy of conversion will succeed always or even most of the time. One problem threatens it in particular — that Socrates’ interlocutors will stop doing philosophy too soon, e.g., before they experience the sort of philosophical pleasure that can be transformative. Socrates is well aware of this danger. He laments in the *Theaetetus* that “many people”, after receiving some initial benefit from his company, leave him “sooner than they should, either of their own accord or through the harmful influence of others” (150e). However, this is precisely why Socrates’ appealing to something that his interlocutors strongly care about to motivate their doing philosophy is strategic. It gives them lasting motivation, thus making it likely that they will persist.

As a last note, it is worth observing that *converting* is different than *fortifying* a conversion. To fortify one’s conversion to valuing wisdom the most, such that it is more likely to persist through hard times (e.g., times when pleasurable experiences are not so readily had), one

48. See Meyer 2015: 270–271 for discussion. The key point is that Plato is not asserting psychological hedonism, but rather the weaker claim that a person will not *willingly* or *wholeheartedly* (ἐκὼν ἐθέλοι) be persuaded to do what causes more pain overall (or even an equal amount of pleasure and pain). Thanks to Katy Meadows for calling my attention to this passage.

49. As Bobonich 2002: 566 n. 95 rightly notes, “this does not entail that pleasure is the only (or the most important) feature that makes [the most pleasant life] good”. It may be that *experience* living the just life shows a person that actually what makes it good is something else entirely. Cf. Paul 2014.

50. It is worth adding, too, that, if you take pleasure in learning, you will be likelier to improve at it, and you will be in a better position to perceive its genuine value, a perception that itself would be pleasurable.

should grasp how and why wisdom is supremely valuable. But that is no more required for converting to valuing wisdom the most than grasping how and why a person is so valuable is required to start valuing them the most, e.g., to fall in love with them. In both the ancient Greek and contemporary contexts, to value something is to believe that it is important and, motivated by that belief, to act on behalf of it and pay attention to it.<sup>51</sup> It is thus possible to begin valuing something the most without fully understanding why it is so important.<sup>52</sup>

#### 4. Conclusion

Conversion is never a matter of certainty. One can never be sure that, if a person is told some claim, or if she has some experience, she will respond to it by reforming her way of life. But that is compatible with supposing that some tactics are better able to accomplish it than others. When it comes to converting his interlocutors to a life dedicated to wisdom, I have argued that Socrates' typical strategy is designed to avoid as much as possible arousing in his interlocutors the defensiveness and stubbornness that is often caused by directly contesting a person's values. His strategy is best understood as consisting of two stages. In the first, Socrates exploits his interlocutor's concern with some external, non-philosophical goal (e.g., rhetorical expertise, power, social status) to motivate him to take up philosophy as a means to it. This is the stage that Plato depicts in many dialogues. In the second, this instrumental pursuit is meant to cause a transformation after which the interlocutor values wisdom the most, a transformation that

51. See especially *Ap.* 24d–25c and *Tht.* 167e in Plato; and *Mem.* 1.2.22, *Mem.* 2.4.2–4, *Sym.* 8.25, *Cyr.* 8.8.8–9, and *Cyr.* 8.8.19 in Xenophon. For valuing in the contemporary context, see Helm 2001: 71–86 and 99–106 and Frankfurt 1998: 82–85.

52. It is reasonable to expect, however, that, given that valuing requires attention, a person will be required *eventually* to grasp how and why the relevant object is valuable, since, if she fails to do so even after considerable time, then that is evidence that she has not been attending to it, and thus has not been valuing it.

happens because of the pleasure of philosophy.<sup>53</sup> Thus, contrary to what some scholars suppose, argument *can* be an effective tool of philosophical conversion, in at least two distinct ways: it can spur people to do more philosophy; and it can cause the experience of intellectual pleasure, thereby reshaping their beliefs about the good.

As a last note, it is worth considering why Plato would have depicted Socrates employing such a strategy. I mentioned one reason already: the alternative, i.e., a direct approach, seems likely to fail, in part due to the problem of evaluative inertia. Thus, Plato is motivated by his own moral psychology to consider philosophical conversion to be an indirect affair. This explanation is internal to the dialogues, but I want to note also an explanation that is external to them. It requires a bit of speculation. One of the few historical facts that we know about Plato is that he founded a school of sorts that later became known as the Academy.<sup>54</sup> So, it is plausible — here is the speculative part — that Plato wanted to attract students, and that the dialogues were meant to function (among other things) as recruitment tools.<sup>55</sup> To recruit students, though, Plato would not have needed to convince his readers that wisdom is of intrinsic value. It would suffice to persuade them of its immense instrumental value — a lesson to which many of them

53. Is the fact that Plato never depicts this second stage evidence that, in the end, he is pessimistic about philosophical conversion? I do not think so. First, it is unclear how the second stage could be depicted (how could one depict that moment of transformation?). Second, it is implausible that, once an interlocutor is motivated in the first stage to continue doing philosophy, merely a second conversation would suffice to be transformative. Likely many more would be needed, and it may be that Plato did not wish to devote the requisite time and energy to depicting them.

54. It is remarkable, though, how little we know about it during Plato's lifetime. Dillon 2003: 1–16 is a useful starting point here.

55. Cf. Kahn 1986: "If we imagine the dialogues read, as they surely were, by the young men themselves, the question becomes: what kind of training are they to pursue? The one thing Plato cannot be saying to them is 'Go study philosophy with Socrates'. But he might well be saying 'Come study philosophy with me'. ... Why should he not want to attract the most gifted minds to come as students or associates?" (10). Consider also Themistius' remark at *Orations* 23.295 that Axiothea of Phlius, upon reading the *Republic*, came to study with Plato, and likewise with the Corinthian farmer and the *Gorgias*.

would be eagerly receptive, given their strong desires for money, power, and honor. Then, once they were in the doors, Plato could expose them gradually to the sort of arguments meant to reshape their values. In this way, Plato's conversion project may well have exhibited the same structure that, I have argued, he depicts shaping Socrates' own: an initial stage of motivating the doing of more philosophy, whatever it takes, followed by a stage of philosophical conversion.<sup>56</sup>

### Works Cited

- Annas, J. 1981. *An Introduction to Plato's Republic*. Oxford: Clarendon Press.
- Barney, R. 2008. "Eros and Necessity in the Ascent from the Cave". *Ancient Philosophy* 28 (2): 357–372.
- Belfiore, E. 2012. *Socrates' Daimonic Art: Love for Wisdom in Four Platonic Dialogues*. Cambridge: Cambridge University Press.
- Benson, H. 1995. "The Dissolution of the Problem of the Elenchos". *Oxford Studies in Ancient Philosophy* 13: 45–112.
- Benson, H. 2000. *Socratic Wisdom: The Model of Knowledge in Plato's Early Dialogues*. Oxford: Oxford University Press.
- Beverluis, J. 2000. *Cross-Examining Socrates: A Defense of the Interlocutors in Plato's Early Dialogues*. Cambridge: Cambridge University Press.
- Blondell, R. 2002. *The Play of Character in Plato's Dialogues*. Oxford: Oxford University Press.
- Bobonich, C. 2002. *Plato's Utopia Recast: His Later Ethics and Politics*. Oxford: Oxford University Press.
- Brickhouse, T. and Smith, N. 1994. *Plato's Socrates*. New York: Oxford University Press.
- Brickhouse, T. and Smith, N. 2000. *The Philosophy of Socrates*. Boulder: Westview Press.
- Burnet, J. 1974. *Plato's Euthyphro, Apology of Socrates, and Crito*. Oxford: Oxford University Press.
- Chance, T. 1992. *Plato's Euthydemus: Analysis of What Is and Is Not Philosophy*. Berkeley: University of California Press.
- Collins, J. 2015. *Exhortations to Philosophy: The Protreptics of Plato, Isocrates, and Aristotle*. New York: Oxford University Press.
- Corner, A., Whitmarsh, L., and Xenias, D. 2012. "Uncertainty, Scepticism, and Attitudes Towards Climate Change: Biased Assimilation and Attitude Polarization". *Climatic Change* 114 (3–4): 463–478.
- de Strycker, E. and Slings, S. 1994. *Plato's Apology of Socrates: A Literary and Philosophical Study with a Running Commentary*. Leiden: E. J. Brill.
- Denyer, N. 2001. *Plato: Alcibiades*. Cambridge: Cambridge University Press.
- Destrée, P. 2017. "How Does Contemplation Make You Happy? An Ethical Reading of Diotima's Speech". In P. Destrée and Z. Giannopoulou, ed. *Plato's Symposium: A Critical Guide*, 216–234. Cambridge: Cambridge University Press.
- Dillon, J. 2003. *The Heirs of Plato: A Study of the Old Academy (347–274 BC)*. Oxford: Oxford University Press.
- Doyle, J. 2012. "Socratic Methods". *Oxford Studies in Ancient Philosophy* 42: 38–75.
- Ebrey, D. 2017. "The Asceticism of the Phaedo: Pleasure, Purification, and the Soul's Proper Activity". *Archiv für Geschichte der Philosophie* 99 (1): 1–30.
- Edmonds III, R. 2017. "Alcibiades the Profane: Images of the Mysteries". In P. Destrée and Z. Giannopoulou, ed. *Plato's Symposium: A Critical Guide*, 194–215. Cambridge: Cambridge University Press.
- Elster, J. 1983. *Sour Grapes: Studies in the Subversion of Rationality*. Cambridge: Cambridge University Press.

56. I am greatly indebted to Rachel Barney, who, over many years, helped me to develop and refine the main ideas of this paper. Thanks also to Brad Inwood, James Allen, Sean Kelsey, Tom Hurka, and Jennifer Whiting, whose insights and objections improved its overall argument. Suzanne Obdrzalek, Sukaina Hirji, Katy Meadows, Doug Reed, Robbie Howton, Willie Costello, Larkin Philpot, and Doug Campbell provided very useful comments at several key points. I would like to thank also the two anonymous referees for their serious and genuinely helpful engagement with the paper.

- Frankfurt, H. 1998. "The Importance of What We Care About". In H. Frankfurt, ed. *The Importance of What We Care About*, 80–94. New York: Cambridge University Press.
- Frankfurt, H. 1992. "On the Usefulness of Final Ends". *Iyyun: The Jerusalem Philosophical Quarterly* 41: 3–19.
- Frankfurt, H. 2004. *The Reasons of Love*. Princeton: Princeton University Press.
- Guthrie, W. 1971. *Socrates*. Cambridge: Cambridge University Press.
- Helm, B. 2001. *Emotional Reason: Deliberation, Motivation, and the Nature of Value*. Cambridge: Cambridge University Press.
- Irwin, T. 1995. *Plato's Ethics*. Oxford: Oxford University Press.
- Jirsa, J. 2009. "Authenticity of the *Alcibiades* 1: Some Reflections". *Listy filologické* 132 (3/4): 225–244.
- Kahn, C. 1986. "Plato's Methodology in the *Laches*". *Revue Internationale de Philosophie* 40 (156/157): 7–21.
- Kahn, C. 1996. *Plato and the Socratic Dialogue: The Philosophical Use of a Literary Form*. Cambridge: Cambridge University Press.
- Klosko, G. 1983. "The Insufficiency of Reason in Plato's *Gorgias*". *The Western Political Quarterly* 36 (4): 579–595.
- Lord, C., Ross, L., and Lepper, M. 1979. "Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence". *Journal of Personality and Social Psychology* 37 (11): 2098–2109.
- MacIntyre, A. 2007. *After Virtue*. Notre Dame: University of Notre Dame Press.
- Mackenzie, M. 1988. "The Virtues of Socratic Ignorance". *Classical Quarterly* 38 (2): 331–350.
- McDowell, J. 1998. *Mind, Value, and Reality*. Cambridge: Harvard University Press.
- Meyer, S. 2015. *Plato: Laws 1 and 2*. Oxford: Oxford University Press.
- Moss, J. 2006. "Pleasure and Illusion in Plato". *Philosophy and Phenomenological Research* 72 (3): 503–535.
- Moss, J. 2007. "The Doctor and the Pastry Chef: Pleasure and Persuasion in Plato's *Gorgias*". *Ancient Philosophy* 27 (2): 229–249.
- Nehamas, A. 1998. *The Art of Living: Socratic Reflections from Plato to Foucault*. Berkeley: University of California Press.
- Nehamas, A. 1999. "What Did Socrates Teach and to Whom Did He Teach It?" In A. Nehamas, ed., *Virtues of Authenticity*, 59–82. Princeton: Princeton University Press.
- Nussbaum, M. 1980. "Aristophanes and Socrates on Learning Practical Wisdom". *Yale Classical Studies* 26: 43–97.
- Paul, L. 2014. *Transformative Experience*. Oxford: Oxford University Press.
- Plochmann, G. and Robinson, F. 1988. *A Friendly Companion to Plato's Gorgias*. Carbondale and Evansville: Southern Illinois University Press.
- Politis, V. 2006. "Aporia and Searching in Early Plato". In L. Judson and V. Karasmanēs, ed., *Remembering Socrates: Philosophical Essays*, 88–109. Oxford: Oxford University Press.
- Robinson, R. 1953. *Plato's Earlier Dialectic*. Oxford: Clarendon Press.
- Rossetti, L. 2011. *Le dialogue socratique*. Paris: Les Belles Lettres.
- Rowe, C. 2007. *Plato and the Art of Philosophical Writing*. Cambridge: Cambridge University Press.
- Russell, D. 2005. *Plato on Pleasure and the Good Life*. Oxford: Oxford University Press.
- Scott, D. 1999. "Platonic Pessimism and Moral Education". *Oxford Studies in Ancient Philosophy* 17: 15–36.
- Scott, D. 2006. *Plato's Meno*. Cambridge: Cambridge University Press.
- Scott, G. 2000. *Plato's Socrates as Educator*. Albany: State University of New York Press.
- Sheffield, F. Forthcoming. "Defending Plato's *Alcibiades*."
- Slings, S. 1999. *Plato: Clitophon*. Cambridge: Cambridge University Press.
- Smith, N. 2004. "Did Plato Write the *Alcibiades* 1?". *Apeiron* 37 (2): 93–108.
- Szlezák, T. 1999. *Reading Plato*. London: Routledge.

- Vlastos, G. 1988. "Elenchus and Mathematics: A Turning-Point in Plato's Philosophical Development". *The American Journal of Philology* 109 (3): 362–396.
- Vlastos, G. 1991. *Socrates, Ironist and Moral Philosopher*. Cambridge: Cambridge University Press.
- Weiss, R. 2006. "Socrates: Seeker or Preacher?" In S. Ahbel-Rappe and R. Kamtekar, ed., *A Companion to Socrates*, 243–253. Oxford: Oxford University Press.
- Wilberding, J. 2004. "Prisoners and Puppeteers in the Cave". *Oxford Studies in Ancient Philosophy* 27: 117–139.
- Wolf, R. 2000. "Calicles and Socrates: Psychic (Dis)harmony in the *Gorgias*". *Oxford Studies in Ancient Philosophy* 18: 1–40.
- Wolf, R. 2004. "The Practice of a Philosopher" *Oxford Studies in Ancient Philosophy* 26: 97–129.
- Yunis, H. 2011. *Plato: Phaedrus*. Cambridge: Cambridge University Press.