Punishment and Psychology in Plato's *Gorgias*

J. Clerk Shaw

Department of Philosophy

University of Tennessee

Knoxville, TN 37996 USA

jshaw15@utk.edu

Abstract

In the *Gorgias*, Socrates argues that just punishment, though painful, benefits the unjust person by removing injustice from her soul. This paper argues that Socrates thinks the true judge (i) will never use corporal punishment, because such procedures do not remove injustice from the soul; (ii) will use refutations and rebukes as punishments that reveal and focus attention on psychological disorder (= injustice); and (iii) will use confiscation, exile, and death to remove external goods that facilitate unjust action.

Keywords

Plato—punishment—psychology—justice—pain

In the *Gorgias*, Socrates argues that just punishment, though painful, benefits the person punished by removing injustice from her soul (476a-481b, 504c-505c, 523a-527a).[1] Socrates' claims raise two questions. First, what forms of punishment does he endorse?

Second, how do these forms of punishment remove injustice from the soul? That is, what is Socrates' penology and his correlative psychology of punishment?

Recent years have seen two main answers to this pair of questions. On one view, Socrates endorses only discursive punishments. This fits with a wider view of Socratic psychology on which action proceeds from an agent's belief about what is overall best for her in her circumstances, and only rational argument guides such beliefs well.[2] Others say that Socrates endorses a wider range of punishments: whipping, beating, imprisonment, confiscation of property, exile, and death. This account of Socratic penology also fits with a wider view of Socratic psychology, on which (as before) action proceeds proximally from an agent's belief about what is overall best for her in her circumstances, but such beliefs can be shaped by non-rational appetites and passions. Thus, a person's beliefs about what is best for her can be altered through painful punishment.[3] I argue for a position between these extremes: Socrates endorses not only discursive punishments, but also confiscation, exile, and death. However, he does not endorse corporal punishment. I offer no general account of Socratic psychology, but I argue that familiar features of Socrates' view explain why he thinks these specific punishments remove injustice from the soul. This effectively undermines both views of Socratic psychology mentioned above. The first cannot capture Socrates' endorsement of confiscation, exile, and death, while the second is motivated in large part by his supposed commitment to the efficacy of corporal punishment.

I begin with two preliminaries concerning Socrates' terminology and whether he offers an ideal or non-ideal theory of punishment. Then, I examine textual evidence that may seem to suggest that Socrates endorses corporal punishment, and argue that he does

not. Indeed, he carefully avoids endorsing corporal punishments as he mentions them, which strongly suggests that he does not endorse them. So does a feature of his analogy between medicine and justice. This is unsurprising; corporal punishment does not remove injustice from the soul, but spurs the punished person to avoid future punishment by being careful in commiting injustice. Next, I show that Socrates considers rebuke and refutation forms of punishment (as some deny). Again, this makes sense: we can see why, on Socrates' psychology, rebuke and refutation can remove injustice from the soul. Third, I argue that Socrates endorses confiscation, exile, and death as forms of punishment. One might wonder why these do not present just the same problem as corporal punishment—namely, that they merely encourage cautious injustice. No doubt they do, but these forms of punishment also remove external goods (wealth, friends, and life itself). External goods provide the means by which injustice is committed and punishment for injustice avoided. Removing external goods thus prevents unjust actions. Finally, I turn more specifically to the death penalty. I suggest that Socrates associates the death penalty with incurable souls, and I briefly consider the post-mortem punishment of both curables and incurables.

Two Preliminaries

Socrates seems to define punishment functionally: whatever removes injustice from the soul counts as punishment. He often talks of the person so treated as κολάζεθαι, which covers not only punishment in the narrow sense but also rebuke. One might say the term is ambiguous, but that is simply to privilege our categories over Greek (or specifically Socratic) categories. Compare: καλόν can be rendered variously in English: 'beautiful',

'admirable', 'noble', 'fine', and so on. To say that καλόν is ambiguous is simply to suppose that our linguistic categories carve nature at the joints but Greek linguistic categories (or revisionary Socratic categories) do not.[4] Similarly with κολάζεθαι. If we do not recognize rebukes and refutations as punishments, it does not follow that they are not forms of κολάσις. Thus, "punishment" below should be read as a gloss on κολάσις, without any presumption that punishment so understood must capture our pre-theoretical views about punishment or, for that matter, Greek pre-theoretical views about κολάσις.[5] If Socrates says strange things about κολάσις on a given reading, it does not follow either that the interpretation is wrong or that Socrates is. Socrates is a radical, and sometimes revisionary theories are correct.

Relatedly, when Socrates says that punishment removes injustice, he means ideal punishment, not necessarily the punishments of positive law.[6] (The *Crito* argues that one should submit to the punishments of positive law in one's city, but not because enduring such punishment necessarily improves one's soul; rather, unjustly avoiding it harms one's soul.)[7] Socrates says that when someone is punished correctly (ὀρθῶς; 476d8, 478a7, 525b1-2)—that is, justly (δικαίως; 476a8, 476e1, 477a6)—she benefits (476b-478b). But not any chance person can punish justly. This is a special case of the ability to *act* justly, which requires a certain power or craft (509e-510a; cf. 464b-c, 477e-478a). Thus, when Socrates recommends that the unjust go to a judge for punishment (478a, 480a-b), he means a true judge—one with the judicial craft, even if not the most skilled possible judge. Socrates does not find true judges in abundance, so nothing suggests that he thinks one should accuse oneself and one's friends and relatives (480a-d) in actually existing lawcourts.[8]

This point might seem to undermine the first: Socrates specifies that he is talking about punishing *correctly*, so he must recognize as punishments some activities that fail to remove injustice. Thus, we cannot say that he defines punishment functionally. However, functional definitions apply even to members of a kind that perform their function poorly. A dull knife is still a knife; a vicious soul is still a soul; false beliefs are still beliefs. At some point, the knife is destroyed by its characteristic vice. Not so the soul, which is therefore immortal (*R.* 608c-611a). The function of belief is to track truth; hence, the *Theaetetus* and *Sophist* ask how beliefs that fail to fulfill their function—false beliefs—are possible. False belief is puzzling, but it is merely one case among many in which a defective F is still an F.

So, again: I use 'punishment' capaciously, so as not to prejudge Socrates' view about what κολάσις involves, and I assume that his primary concern throughout is with an ideal theory of punishment. Socrates' penology does not deliver straightforward practical advice in non-ideal circumstances, i.e. when no expert judge is available or appropriately socially situated. If in one's actual circumstances a likely punishment in a particular case is a whipping, and Socrates thinks the true judge would never impose such a penalty, it does not follow that one should never prosecute such a case.

Corporal Punishment in the *Gorgias*?

I now argue that Socrates never says the true judge will order corporal punishment. The putative evidence that he does is either (i) conditional, (ii) ambiguous, (iii) analogical, or (iv) metaphorical.

(i) As Socrates presses the claim that the unjust person should seek just punishment to rid himself of injustice, he presents a list of cases (480c-d):

> ...if his unjust behavior merits flogging [πληγῶν], he should present himself to be whipped [τύπτειν]; if it merits imprisonment, to be imprisoned; if a fine, to pay it; if exile, to be exiled; and if execution, to be executed.[9]

These claims are all conditional; Socrates does not here affirm that any of these are ever correct, just punishments.[10]

(ii) A little earlier, Socrates may seem to endorse corporal punishment when he describes the person who gets rid of injustice through punishment as (478e2-4):

> ...the man who gets lectured and lashed [ἐπιπληττόμενος], the one who pays what is due.

This translation is disputable. As the verb translated "punish" in these passages, κολάζειν, covers both punishment in a narrow sense and rebuke, so too the verb ἐπιπλήσσειν covers chastisement. Indeed, Socrates contrasts ὁ νουθετούμενός τε καὶ ἐπιπληττόμενος καὶ δίκην διδούς (478e3-4) with the person who μήτε νουθετεῖσθαι μήτε κολάζεσθαι μήτε δίκην διδόναι (478e7-479a1). This replaces the broad ἐπιπλήσσειν with the equally broad κολάζειν. Many other translations and the LSJ use 'chastized', 'rebuked', or 'reproved'.[11] But when Socrates uses conditional language at 480c-d—the passage quoted in (i) above—he reverts to determinate language of physical beatings: πληγῶν and τύπτειν.

(iii) Socrates gives an extended analogy between the soul and the body. In response to Polus' demand to say what he thinks rhetoric is, Socrates calls it a form of flattery, and

Polus proceeds as though he understands his view. Socrates pulls him up short by showing that he does not yet understand how rhetoric differs from another form of flattery, relish-making (ὀψοποιία).[12] Rhetoric and relish-making imitate two different forms of corrective expertise, justice and medicine, directed at the soul and body (462b-465d).

Socrates returns to this analogy throughout the rest of the dialogue. One part of it might suggest that Socrates endorses corporal punishment: he compares going to the judge for just punishment with going to the doctor for surgery or cautery (479a-b):

> I take it that [those who do injustice and escape punishment] have managed to accomplish pretty much the same thing as a person who has contracted very serious illnesses, but, by avoiding treatment manages to avoid paying what's due to the doctors for his bodily faults, fearing, as would a child, cauterization or surgery [τὸ κάεσθαι καὶ τὸ τὲμνεσθαι] because they're painful.

This analogy also recurs right before the conditional claims quoted in (i) above (480c):

> [the orator] should not keep his wrongdoing hidden but bring it out in the open, so that he may pay his due and get well; and compel himself and [his friends and family] not to play the coward, but to grit his teeth and present himself with grace and courage as to a doctor for cauterization and surgery [ὥσπερ τέμνειν καὶ κάειν ἰατρῷ], pursuing what's good and admirable without taking any account of the pain.

This is analogical; nothing commits Socrates to saying that justice's technical procedures include cutting and burning. The analogy requires that justice's technical procedures are

painful, like medicine's, but not that the pains imposed by expert judges are bodily pains. Indeed, the doctor is an expert about the body and the judge an expert about the soul, so we should expect the doctor to perform medical procedures on the body involving bodily pains and the judge to perform judicial procedures on the soul involving psychic pains. And in fact, Socrates describes the true judge as acting on the soul (see below).

Socrates also mentions cutting and burning in an earlier *epagogê*. While arguing that the unjust person is better off being justly punished than escaping just punishment, he says that whenever X φs Y F-ly, Y is φed F-ly by X (476b-d). His examples are: (a) X hits [τύπτει] Y hard or quickly; (b) X burns [κάει] Y severely or painfully; (c) X cuts [τέμνει] Y deeply or painfully. After he argues for the general principle, he applies it to punishment. If X punishes Y correctly (so justly, so admirably, so either beneficially or pleasantly), then Y is punished correctly (so justly, so admirably, so either beneficially or pleasantly, so—since not pleasantly—beneficially) by X.[13] Socrates' examples do not commit him to endorsing corporal punishment. These are not examples of punishment, but examples of actions that follow a certain pattern; Socrates then applies that same pattern to punishment.

Still, Socrates' use of hitting, cutting, and burning in his analogy and in his *epagogê* require explanation—especially since one analogical mention of cutting and burning leads directly into the conditional claim about corporal punishment in (i) above (480c3-8, c8-d3). One might say that the best explanation for Socrates' constant allusions to hitting, cutting, and burning is that he thinks the antecedents of the conditionals are sometimes fulfilled. However, an alternative explanation is at least as plausible.

When Polus earlier denies that the unjust person is better off being punished than escaping punishment, he says (473b-c):

> Suppose that he's caught, put on the rack, castrated [ἐκτέμνηται], and has his eyes burned out [τοὺς ὀφθαλμοὺς ἐκκάηται]. Suppose that he's subjected to a host of other abuses of all sorts, and then made to witness his wife and children undergo the same. In the end he's impaled or tarred.

Polus here expresses his conception of punishment, which stems from prevailing norms.[14] Whether or not it is rhetorically effective, Socrates often uses features of an interlocutor's views to put his points provocatively. He foregrounds Polus' view of punishment by using hitting, cutting, and burning in his *epagogê* and analogy; he thus provokes Polus, even as he carefully avoids saying that expert judges use these procedures. Polus, of course, infers that Socrates thinks one should accuse oneself and one's family and friends in actual law courts, but that simply shows how far he is from grasping Socrates' views. Socrates cannot correct every mistake at once; even trying to do so would be pedagogically misguided. Far better to elicit paradoxes that provoke the desire to understand, as Socrates does with the distinction between what one wants and what seems best to one (467b-c). Indeed, to that end, Socrates actually goes along with the idea that punishment harms (470a-b).

(iv) At the end of the *Gorgias*, Socrates makes further analogical use of hitting or whipping, and he moves quickly into metaphor. The object of his analogy is the soul and certain alterations to it that persist after death. As an analogue, he uses marks on the body that persist after death (524c):

> ...if a man had been a criminal whipped for his crime [μαστιγίας] and showed scars, traces of beatings [πληγῶν] on his body inflicted by whips [μαστίγιων] or other blows [τραυμάτων] while he was alive, his body can be seen to have these marks, too, when he is dead.

This passage mentions whippings as a punishment of positive law to make the point that acquired features of the body persist after death. This provides an analogical basis for a metaphor concerning the post-mortem condition of the soul (524e-525a):

> He [Rhadamanthus] has often gotten hold of the Great King, or some other king or potentate, and noticed that there's nothing sound in his soul but that it's been thoroughly whipped [διαμεμαστιγωμένην] and covered with scars, the results of acts of perjury and of injustice, things that each of his actions has stamped upon his soul.

Here, the whippings are unjust actions, conceived as injuring the soul, and the scars are marks left on the soul as a result of this self-abuse—not as a result of punishment, corporal or otherwise. In fact, part of the point is that those who can do injustice without sanction do more injustice for precisely that reason (on which more below). They 'whip' their own souls more, and so leave more psychic 'scars'.

Finally, Callicles had critiqued philosophy because it offers no protection against being treated unjustly (486a-c; cf. 508d-e and 521b-522e, which introduces the myth). Socrates responds in part by using corporal punishment as a metaphor again (526e-527a):

> When you come before the judge, the son of Aegina, and he takes hold of you and brings you to trial, your mouth will hang open and you'll get dizzy

there just as much as I will here, and maybe somebody'll give you a demeaning knock [τυπτήσει] on the jaw and throw all sorts of dirt at you.

This passage refers to punishments inflicted by an expert judge, Aeacus. However, the corporal language is metaphorical. Death separates the soul from the body (524b) and this judge examines souls after death, stripped of their bodies (523e). Hence Callicles will have no jaw to hit, and no body on which any corporal punishment may be inflicted.[15] So too, Socrates' description of the incurables as 'hung up in prison' [ανηρτημένους…ἐν τῷ δεσμωτηρίῳ, 525c7] cannot be used to show that he endorses corporal punishment. And again, Socrates' use of bodily metaphors is utterly of a piece with his normal practice.

Hence, none of the evidence shows that Socrates endorses corporal punishment in the *Gorgias*.[16] But those who disagree also explain why Socrates might endorse corporal punishment. Before moving on, I consider their reasons and explain why he does not.

Brickhouse and Smith, *Socratic Moral Psychology* say that Socrates thinks non-rational appetites and passions produce human action, but only by influencing the agent's beliefs about what is best for her. Satisfying non-rational appetites and passions strengthens them, which makes it harder to consider reasons not to satisfy them and so makes it more likely that they will (indirectly) guide one's future actions. Unjust action damages the soul precisely by making it harder to consider reasons not to satisfy appetites and passions. In extreme cases, the appetites and passions grow so strong that they destroy the agent's ability to consider reasons not to satisfy them; such people are incurable.[17]

On this view, the pains of punishment themselves repair damage done to the soul by injustice: 'the benefit comes to [the curable] there in Hades through pain and suffering

[δι' ἀλγηδόνων καὶ ὀδυνῶν]. For it is not possible to be rid of injustice in any other way [οὐ…ὅιόν τε αλλῶς]' (525b7-c1).[18] This happens in two ways. First, insofar as injustice is motivated by certain perceived benefits, punishment can sever the perceived connection between injustice and benefit by imposing perceived harms. Thus, punishment makes the results of injustice 'poverty rather than wealth…or shame rather than honor, or pain rather than pleasure' (p. 117). More likely, this will lead the person punished to do injustice secretly (as they note; pp. 118-119) or openly in a setting where they are secure against punishment.[19] Second, punishment not only imposes perceived harms for injustice that alter the unjust person's calculations; it also triggers calculation in those whose passions and appetites are strong enough to suppress it (pp. 121-122). This does not solve the problem of encouraging secret or powerful injustice rather than removing injustice—as they admit (p. 122). So, both proposed explanations for how corporal punishment removes injustice from the soul fail.[20]

So why does pain remove injustice from the soul? And if pain removes injustice, why wouldn't bodily pain imposed by corporal punishment do the job? We should not read too much into the claim that injustice must be removed δι' ἀλγηδόνων καὶ ὀδυνῶν. One can read διά causally, but διά can also simply indicate manner (here, the manner in which something *must* take place). Earlier, for example, Socrates mentions that some crafts can be practiced silently (διὰ σιγῆς; 450c9). Silence obviously is not the means by which painting, sculpture, and many other crafts accomplish their ends. The latter reading of διά at 525b7 is preferable. Socrates tells Polus that the person who commits injustice should seek punishment without regard for its painfulness, just as the ill patient goes to the doctor

without regard for the painfulness of medical procedures (480a-d). Pain is a side-effect of medical treatment that one should disregard, not how the treatment works; so too, pain is a side-effect of judicial treatment that one should disregard, not how the treatment works.[21]

Discursive Punishment

When Socrates discusses punishment with Callicles, he is still concerned with the expert judge, whom he calls the 'skilled and good orator' (504d5-6).[22] The expert judge disciplines the souls of the citizens by keeping them from fulfilling certain desires, making them self-controlled. This recalls not only the earlier conversation about punishment but also Callicles' rejection of self-control and his embrace of unrestrained desire-fulfillment (491d-492c). At this point, Callicles gets annoyed and attempts to leave the conversation (505c1-2). Socrates replies (505c3-4):

> This fellow won't put up with being benefited and with his undergoing the very thing the discussion's about, with being disciplined [κολαζόμενος]!

Similarly, at the end of his argument that doing injustice is worse than suffering injustice, Socrates urges Polus (475d6-7):

> Don't shrink back from answering, Polus. You won't get hurt in any way. Submit yourself nobly to the argument, as you would to a doctor, and answer me. Say yes or no to what I ask you.

In both cases, Socrates describes his refutations as punishments that try to remove injustice from the soul, much as the doctor removes illness from the body (cf. 521e-522c).[23] In one case, with Callicles, he even uses the verb κολάζειν. (This should not be surprising, since

κολάζειν can cover verbal reprimands.) Further, refutation neatly fits Socrates' model of punishment. He thinks refutation benefits the person refuted, whether himself (*G.* 457e-458a, 470c6-8, 506c1-3) or others (*Ap.* 29d-31c, 36d; *M.* 84a-c; *Tht.* 210b-c), though it is painful.[24] His own account of justice in the *Gorgias* makes it clear how refutation removes injustice from the soul. Socrates defines justice and temperance as the soul's order and organization, respectively (τάξις and κόσμος; 503-504). As the carpenter imposes order and organization on wood, nails, screws, and so on, so too the politician imposes order and organization on a soul's attitudes—its beliefs, desires, fears, loves, pleasures, and so on. When Socrates refutes people, he attempts true politics by seeking to remove disorder and disorganization from their souls—that is, from their total set of psychological attitudes. Typically, he uncovers contradictory views; that reveals the subject's lack of knowledge and virtue, which pains her. With this kind of punishment, it is unsurprising for Socrates to say one should reveal one's injustices and those of one's friends, family, and city (480b-d). We actually expect him to say that friends should inquire together into virtue and seek to improve themselves and their own. Indeed, Socrates asks for this sort of 'revenge' on his sons at the end of the *Apology* (41e-42a).[25]

Apology 26a contrasts punishment (κολάσις) with instruction (μαθήσις)—that is to say, teaching and exhortation (διδάσκειν καὶ νουθετεῖν).[26] One might consider this proof that Socrates does not consider instruction a form of punishment.[27] However, he makes a legal claim here: positive law requires one to bring to court only those who need κολάσις. Socrates is involved in a trial, so appeals to the law are apt, and need not reveal his own views of ideal punishment. But what *should* we make of the contrast between punishment

and instruction? Rather than associating punishment with non-discursive procedures and instruction with discursive procedures, we should understand punishment to be directed at removing injustice, while instruction instills justice. On this view, the judge metes out punishment (which can include rebukes and refutations), while the legislator as such does not punish. The politician, who is both judge and legislator (464b-c), punishes to remove injustice and instructs to instill justice. (Their procedures may overlap, but again, the legislator as such does not punish.) This parallels the situation with two crafts concerning the body: medicine removes illness from the body, but gymnastics instills health and strength.

Confiscation, Exile, and Death

However, Socratic penology does not begin and end with discursive procedures. When Socrates distinguishes expert politicians from flattering orators and asks Callicles which politicians are or were experts, the latter suggests several politicians of yore (503a-c). Socrates disagrees and expands on the differences between mere oratory and true politics. Eventually, he returns to Callicles' candidate politicians, and says (517b):

> In redirecting [the city's] appetites and not giving in to them, using persuasion and constraint [πείθοντες καὶ βιαζόμενοι] to get the citizens to become better, [the politicians of old] were really not much different from our contemporaries. That alone is the task of a good citizen.

Those who say that Socratic politics involves only persuasion must explain this passage, but it is unclear how they could.[28] More particularly, Socratic penology includes constraints in

the form of confiscation, exile, and death. One passage may suggest this, while another clearly does. I begin with the weaker evidence, move on to the stronger, and then explain why these punishments make sense given Socrates' psychological views.

When Socrates discusses punishment with Callicles, he characterizes the true orator thus (504d5-e):

> ...this is what the skilled and good orator will look to when he applies to people's souls whatever speeches he makes [τοὺς λόγους προσοίσει ταῖς ψυχαῖς οὓς ἂν λέγῃ] as well as all of his actions, and any gift he makes or any confiscation he carries out [καὶ δῶρον ἐάν τι διδῷ, δώσει, καὶ ἐάν τι ἀφαιρῆται, ἀφαιρήσεται]: he will always give his attention to how justice may come to exist in the souls of his fellow citizens and injustice be gotten rid of, how self-control may come to exist there and lack of discipline be gotten rid of...

At first glance, this passage might seem to indicate a fully discursive penology. Socrates mentions speeches in the simple future indicative, as if to imply that the true orator will in fact apply such treatments. When he turns to gifts and confiscations, we have conditional language again (as at 480c-d). However, a relative clause with subjunctive + ἄν joined to a main clause in the future indicative (here, τοὺς λόγους προσοίσει ταῖς ψυχαῖς οὓς ἂν λέγῃ) has the force of a future more vivid conditional,[29] exactly like the future more vivid conditionals that Socrates uses to talk about gifts and confiscations. In the other direction, these future more vivid conditionals (and their equivalents) might be thought to show that the conditions are sometimes fulfilled. If so, Socrates here implies that the good orator will

sometimes use gifts and confiscations (and speeches). But while more vivid conditions can assign a greater probability to the protasis than less vivid conditions, more vivid conditions can also be used even of impossibilities. So, they need not show that the speaker views the condition as more likely to be fulfilled.[30] Still, other passages show that Socrates endorses the judicial use of speeches (see previous section) and confiscations (as I now argue).

Soon after Socrates first explains his view of rhetoric, Polus argues that orators have great power because of their tyrannical ability to kill, exile, and confiscate (466b-c). During Socrates' subsequent argument that orators and tyrants have least power in the city (466a-468e), he uses these examples—we kill, exile, and confiscate because we think doing these things is better than not doing them (468b). By distinguishing between wanting and thinking best, Socrates gets Polus to agree that we want to kill, exile, and confiscate only if those actions actually benefit us rather than harm us (468c). If they actually harm us, but we do them anyway, then we do not do what we want (468d). However, Polus persists in thinking the bare ability to kill, exile, and confiscate is enviable (468e, 469c). So, Socrates gives an example of this ability (469c-e) that even Polus recognizes as problematic (470a):

> Can you tell me then what your reason is for objecting to this sort of power?
>
> —Yes, I can
>
> What is it? Tell me.
>
> —It's that the person who acts this way is necessarily punished.
>
> And isn't being punished a bad thing?
>
> —Yes, it really is.

Socrates now returns to killing, exiling, and confiscating (470b-c):

> Let's consider this point, too. Do we agree that sometimes it's better to do those things we were just now talking about, putting people to death and banishing them and confiscating their property, and at other times it isn't?
>
> —Yes, we do.
>
> This point is evidently agreed upon by you and me both?
>
> —Yes.
>
> When do you say that it's better to do these things then? Tell me where you draw the line.
>
> —Why don't you answer that question yourself, Socrates.
>
> Well then, Polus, if you find it more pleasing to listen to me, I say that when one does these things justly, it's better, but when one does them unjustly, it's worse.

Here, Socrates clearly thinks that killing, exiling, and confiscating are sometimes just and so better. One might argue that this is purely dialectical—that Socrates merely uses Polus' examples of tyrannical 'power'. After all, he does this mere moments before; he uses Polus' claim that being punished is bad even though he thinks that being justly punished benefits the person punished. However, this strategy will not work with the agreement over killing, exiling, and confiscating. Socrates says twice that they agree these are sometimes good and sometimes bad (470b1-5, 6-8) and he then articulates his own dividing line (470c1-3); the best explanation is that he accepts this claim.[31] Notably, their agreement does not mention corporal punishment. Polus introduces that topic only later (473b-d), but Plato is in charge here. He could easily have put corporal punishment on Polus' initial list of tyrannical

actions and, subsequently, among the actions agreed to sometimes harm and sometimes benefit the agent, but he did not.

Why does Socrates think that such punishments remove injustice from the soul?[32] Briefly, because they remove external goods—wealth, friends, and life itself—that allow one to commit injustice on a wider scale, and (relatedly) to avoid punishment for injustice.[33] So, if an enemy unjustly acquires wealth, the way to harm him is to 'scheme with him not to return it but to keep it and spend it in an unjust and godless way both on himself and his people' (481b). If someone unjustly acquires possessions, 'as he unjustly confiscated them…so having gotten them, he'll use them unjustly too, and if unjustly, shamefully, and if shamefully, badly' (521c). Keeping what one has unjustly acquired is harmful by itself, but further harm lies in spending the money unjustly, as someone who acquired it unjustly will do—not simply because she is spending unjustly acquired money, but because such a person spends in ways that would be unjust even if the money were justly acquired.

External goods enable injustice on a wider scale in large part *because* they prevent punishment for one's injustices. Right after Socrates argues that being correctly punished benefits the unjust person, he says that those who fail to seek out such treatment 'find themselves funds and friends [καὶ χρήματα παρασκευαζόμενοι καὶ φίλους] and ways to speak as persuasively as possible' (479c). Likewise, Socrates later says most incurables are 'tyrants, kings, potentates, and those active in the affairs of cities, for these people commit the most grievous and impious errors because they're in a position to do so [διὰ τὴν ἐξουσίαν]' (525d). When an unjust person has external goods that let her avoid being punished for injustice, that leads her to wider injustice. (Recall Polus, who says that open

injustice is bad only because it leads to being punished [470a], and cf. 510d-e.) Removing an unjust person's funds, friends, or life (through confiscation, exile, or death) removes the enabling conditions for unjust action and so tends to remove injustice from her soul. Socrates may have this in mind when he says the skilled politician keeps an unjust soul away from what it has an appetite for and will only let it do what improves it (505a-b).

Finally, I explain why these punishments are not subject to the same objections as corporal punishment and why corporal punishment cannot be justified in the same way as these punishments. It may seem that removing external goods, like corporal punishment, leads the punished person to seek ways of doing injustice secretly or securely rather than removing injustice from their soul. I grant that this is one result of non-corporal, non-discursive punishment, but removing external goods has two more features not shared by corporal punishments. First, whatever the person punished thinks, losing wealth makes it harder to commit injustice (at least, on the same scale).[34] Second, removing external goods makes it harder to do injustice with the same assurances of safety. Beatings, in constrast, neither make it harder to do injustice in the future nor to avoid sanctions in the future.

Might corporal punishments have these effects? Some polities punish theft by severing the thief's hand, for example, which might make future thieving harder. Socrates does not entertain such possibilities; the only corporal punishments he considers are beating, whipping, and burning, which seem intended simply to cause physical pain. So, we can only speculate about what Socrates would say. Considered purely in terms of bodily ability, this sort of punishment seems unlikely to prevent future injustices; there are ways to steal that do not require both hands. However, Socrates mentions physical defects as

conditions that keep some people from becoming unjust (*R.* 496a-c). And considered as an indelible social mark, cutting off a hand could limit someone's capacity for injustice in the future. None of that shows that physical pain itself removes injustice from the soul.

Incurables, Death, and Post-Mortem Punishment

Socrates mentions that certain souls become incurably unjust.[35] These figures first appear late in the *Gorgias*. Callicles repeatedly stresses the injustices suffered by those who will not act unjustly (485e-486d, 511a, 521a-d). In reply to the second of these passages, Socrates compares rhetoric's ability to preserve life and property with sailing's (511d-512b):

> ...the man who possesses this craft and who has accomplished these feats...is enough of an expert,[36] I suppose, to conclude that it isn't clear which ones of his fellow voyagers he has benefited by not letting them drown in the deep, and which ones he has harmed, knowing that they were no better in either body or soul when he set them ashore than they were when they embarked. So he concludes that if a man afflicted with serious incurable physical diseases did not drown, this man in miserable for not dying and has gotten no benefit from him. But if a man has many incurable diseases in what is more valuable than his body, his soul, life for that man is not worth living, and he won't do him any favor if he rescues him from the sea or from prison or from anywhere else. He knows that for a corrupt person it's better not to be alive, for he necessarily lives badly.

This seems to indicate that Socrates considers the death penalty appropriate for incurables. Likewise elsewhere in Plato: Protagoras designates exile and death as punishments for the incurable (*Pr.* 325a-b); Socrates says that the good judge puts incurables to death (*R.* 410a); and the Athenian suggests exile and death as punishments for incurables (*L.* 735d-e, 854c, 854e-855a, 862e-863a, 876e-877c, 941d-942a). In particular, if human experts cannot cure someone, the death penalty seems appropriate; it makes sense to turn such a person over to divine judges who are better situated. (Presumably, Rhadamanthus sees infallibly whether someone is curable or incurable; cf. 526b.) Thus, one purpose of the myth is to describe the wider context in which skilled human judges operate: the death penalty hands over apparently incurable humans to superior judges.[37] None of this is to deny that removing friends and life through exile and death can also, as described above, restrain the unjust person from further unjust acts.

However, incurables are also punished post-mortem, and a dead incurable cannot be executed. (Annihilation would be better for them, but that is impossible; cf. *Phd.* 107c). Socrates describes their post-mortem condition thus (525c):

> From among those who have committed the ultimate wrongs and who because of such crimes have become incurable [ἀνίατοι] come the ones who are made examples of. These persons themselves no longer derive any profit [ὀνίνονται] from their punishment, because they're incurable. Others, however, do profit from it when they see them undergoing for all time the most grievous, intensely painful and frightening sufferings for their

> errors, simply strung up [ἀνηρτημένους] there in the prison in Hades as examples, visible warning to unjust men who are ever arriving.

As I have said, dead incurables have no bodies to be whipped or strung up. They also have no wealth to be confiscated. One might suppose antecedently that incurables are exiled in the afterlife, but arriving curables witness their condition, and they do not spontaneously suffer from the condition of their souls, but are punished by another (ὑπ᾽ ἄλλου; 525b). This leaves one possibility: the incurables are punished discursively—refuted and rebuked. For the same reasons, these must be the same sorts of punishment that curables receive in the afterlife.[38] That directly connects the fears that curables have at seeing the condition of incurables and the effectiveness of the same sort of punishments on them. Those who can respond to such punishments, seeing the condition of those who cannot, are motivated to respond. So, while just punishment is not entirely discursive in this life, it is in the next.

The pain involved in post-mortem punishment, then, is much the same as the pain experienced in Socratic refutation or rebuke. Socrates makes others aware of the flaws in their souls. While embodied, other things can distract them from this deep problem and let them ignore the flaws that Socrates uncovers. But once disembodied, just as the judges will not be distracted by good looks, high reputation, and political office, so too the soul can no longer avoid the painful awareness of its own flaws by distracting itself with its good looks, high reputation, and political office. Thus, there is no way to remove injustice (the soul's disorder) except through (in the presence of) pain.[39] Indeed, the pains of refutation and rebuke even seem to play a causal role in removing injustice—not by themselves, but by

directing the subject's attention to the very flaws in her soul that need treatment. However, this point obviously will not extend to the painfulness of other forms of punishment.

Conclusion

Many readers of Plato are drawn to Socrates' revolutionary claim that one should never return wrong for wrong or harm for harm (*Cr.* 49b-d).[40] In keeping with that claim, as we have seen, Socrates thinks punishment must be just and (when possible) beneficial for the person punished. Some interpretations of Socratic penology erase the radical appeal of his view by rationalizing even corporal punishments; others naively say that only discursive punishments are just. The interpretation above preserves Socrates' radically revisionary view of political life without supposing that discourse alone can always remove injustice. So understood, Socrates' view is worthy of consideration alongside contemporary moral education theories of punishment (though these differ in important ways).[41] In short, I hope to have offered a reading here that is not only textually superior to the alternatives but also one that holds more philosophical interest.[42]

Department of Philosophy

University of Tennessee

---

[1] An anonymous referee notes that the ostensible topic of the *Gorgias* is rhetoric, and wonders whether Socrates might shape his comments about punishment to suit that context, rather than offering a theory of punishment in its own right. However,

punishment is directly relevant to rhetoric; Socrates calls the rhetor an imitator of the true judge, who imposes punishments to remove injustice from the soul.

[2] T. Penner, 'Socrates and the early dialogues,' in R. Kraut (ed.), *The Cambridge Companion to Plato* (Cambridge: Cambridge University Press, 1992), pp. 121-169, pp. 161-162n51; T. Penner, 'Socrates' in C. J. Rowe and M. Schofield (eds.), *The Cambridge History of Greek and Roman Political Thought* (Cambridge: Cambridge University Press, 2000), pp. 164-189; C. J. Rowe, *Plato and the Art of Philosophical Writing* (Cambridge: Cambridge University Press, 2007), Ch. 4.

[3] T. C. Brickhouse and N. D. Smith, 'Incurable Souls in Socratic Psychology', *Ancient Philosophy*, 22 (2002), pp. 21-36; T. C. Brickhouse and N. D. Smith, *Socratic Moral Psychology* (Cambridge: Cambridge University Press, 2010), esp. Ch. 4.

[4] Cf. A. W. H. Adkins, *Moral Values and Political Behavior in Ancient Greece* (New York: W. W. Norton, 1972), pp. 3-6, though I have not used his solution of transliterating such terms. In recognizing that the ordinary sense of κολάσις extends more widely than the ordinary sense of 'punishment', I do not mean to deny that Socrates' view, or something like it, might be a good account of punishment. If that were so, then 'punishment' would be a bad translation—but it is not a bad translation.

[5] So, for example, an anonymous referee wonders whether the procedures described as punishment here are really "pre-emptive action to prevent future injustice", rather than "punishment for past injustice".

[6] See M. M. MacKenzie, *Plato on Punishment* (Berkeley: University of California Press, 1981), pp. 181-182; Rowe, *Philosophical Writing*, pp. 148-151. Contrast Brickhouse and Smith, *Socratic Moral Psychology*, p. 131, who say that Socrates is talking about "those [corrections] the state provided".

[7] Contrast Brickhouse and Smith, *Socratic Moral Psychology*, p. 110, who use *Crito* 51b to argue that Socrates endorses beatings as a form of punishment. They also cite *Hippias Major* 292b, where Socrates says that if he took on Hippias' account of the fine (291d-e), his alter ego would justly beat him—but this is clear hyperbole.

[8] Compare Socrates' shock that Euthyphro is prosecuting his father (*Eu.* 4a-b).

[9] All translations are taken from D. J. Zeyl, *Plato: Gorgias* (Indianapolis: Hackett Publishing Company), sometimes with slight alterations. If γε at 480c were limitative, Socrates would be casting special doubt on the first, corporeal example. In fact, though, μὲν γε there simply introduces a series of examples (cf. H. W. Smyth, *Greek Grammar* [Cambridge, Mass: Harvard University Press, 1984], §2829).

[10] See also J. Moss, 'The Doctor and the Pastry Chef: Pleasure and Persuasion in Plato's *Gorgias*', *Ancient Philosophy*, 27 (2007), pp. 229-249, p. 232n8. Brickhouse and Smith, *Socratic Moral Psychology*, p. 109n15 reject this point without argument, and they lean heavily on this passage, citing it five times in one sentence as evidence for which punishments Socrates endorses (110). (See n.7 above on their other evidence.) Rowe, *Philosophical Writing*, ignores this passage. An anonymous referee suggests that if, as I argue, Socrates ultimately endorses confiscation, exile, and death, that creates a

presumption in favor of the other modes as well. (I disagree, but report the suggestion here in case others agree.)

[11] Examples include B. Jowett, *The Dialogues of Plato* (Oxford: Clarendon Press, 1953); W. Woodhead in E. Hamilton and H. Cairns (eds.), *The Collected Dialogues of Plato* (Princeton: Princeton University Press, 1971); and T. Irwin, *Plato: Gorgias* (Oxford: Clarendon Press, 1979). Cf. Rowe, *Philosophical Writing*, p. 149n14 and surrounding discussion. The LSJ lists no Platonic use of ἐπιπλήσσω meaning "strike". See passages listed there *s.v.* and *Tht.* 200c7, which says that the argument chastises.

[12] Cf. Socrates' first line, in which he compares Gorgias' rhetorical display to a feast (ἑορτῆς; 447a).

[13] For criticism of this argument, see MacKenzie, *Plato on Punishment*, Ch. 11 and App. I.

[14] Cf. C. Kahn, 'Drama and Dialectic in Plato's *Gorgias*', *Oxford Studies in Ancient Philosophy*, 1 (1983), pp. 75-121, who says that Polus represents the position of contemporary common sense.

[15] The soul may have extension, and so shape, post-mortem (cf. S. Broadie, 'Soul and Body in Plato and Descartes', *Proceedings of the Aristotelian Society*, 101 (2004), pp. 295-308). *Phaedo* 81c-e might suggest that corrupt souls are literally corporeal after death, but I read this in terms of (the objects of) their desires. The *Phaedo* does not use the corporeality of corrupt souls in describing their punishments (113d-114b).

[16] Even the *Laws* does not countenance corporal punishment *for citizens*; see T. J. Saunders, *Plato's Penal Code: Tradition, Controversy, and Reform in Greek Penology* (Oxford:

Oxford University Press, 1991), pp. 336-337. Saunders notes 794b as a possible exception, but the Stranger uses a form of κολάζειν, and the context (explaining who has authority to punish in certain circumstances) makes it appropriate to read the verb broadly. In two cases, citizens may be beaten (762c, 784d), but the state never administers a beating. Rather, it declares that any citizen may beat certain others without sanction.

[17] Brickhouse and Smith, *Socratic Moral Psychology* use their type of Socratic intellectualism to explain both how unjust action harms and finally ruins the soul and how punishment removes injustice from the soul. I focus on punishment and how it removes injustice, not on how unjust action harms and ultimately ruins the soul. I do reject their other claims, though; briefly, I see no reason why unjust actions cannot directly strengthen the false evaluative beliefs from which they stem, or why unjust actions cannot make those false evaluative beliefs indelible, so that the agent is incurable.

[18] Brickhouse and Smith, *Socratic Moral Psychology*, p. 123 say Socrates tells Callicles that "*only* corporal punishment rids some souls of injustice". This must be the passage they have in mind, but Socrates speaks of pain and suffering, not corporal punishment specifically. Again, post-mortem corporal punishment is impossible.

[19] It may even reinforce the view that poverty (for example) is terribly bad, making the soul more unjust.

[20] Later, they say that "What is required…is that the agents make the *judgment* that the satisfaction of an appetite at which their action aims is not worth the suffering to which they are liable as a result of such action. Thus, they must form the judgment that the apparent good to which they are attracted is merely apparent" (p. 124). At first glance, these two sentences are in tension with each other. The first says that the agent weighs good X against bad Y and decides that X is not worth Y. The second says that the agent weighs good X against bad Y and decides that X is not good at all (cf. p. 122).

[21] Saunders, *Plato's Penal Code*, p. 166 says that Socrates' exhortation for the unjust to seek punishment without regard for pain is "nonsense, for what point can there by in a whipping etc., if not to cause pain?". He fails to reconsider whether Socrates ever endorses corporal punishment. As we shall see, other punishments also involve concomitant pain, but do not remove injustice by means of pain—with the possible exception of discursive punishments, as I suggest at the end of the paper.

[22] Ordinary rhetoric imitates expertise in justice and pretends to be good in the way justice is, so the expert can reasonably be described as the truly skilled and good orator.

[23] Does Socrates think he is the true politician, as many have thought this sort of comment requires (cf. esp. 528d6-8)? No, he merely attempts true politics; see J. C. Shaw, 'Socrates and the True Political Craft', *Classical Philology*, 106 (2011), pp. 187-207.

[24] At *Gorgias* 457e-458a Socrates says it would please him to be refuted. However, he differs in this way from those he refutes, who either get angry with Socrates or angry with (or ashamed at) themselves.

[25] What about rebuke? Whereas refutation uncovers some inconsistency (disorder, injustice), rebuke keeps the subject's awareness focused on some already-uncovered inconsistency (disorder, injustice)—e.g., some failure to live up to her own standards.

[26] Cp. *Gorgias* 488a-b, where Socrates tells Callicles not to bother exhorting him (νουθετεῖν) if he agrees to what Callicles says in the conversation but doesn't change his ways in light of that agreement. Cf. also *Sophist* 227c-230e, where the Visitor divides soul-cleansing: punishment removes psychological disease; teaching removes psychological ugliness.

[27] Brickhouse and Smith, *Socratic Moral Psychology*, pp. 103-104, 110; they also reveal this view by placing discussion of *Gorgias* 505c3-4 in their chapter on education (p. 139), which they contrast with punishment. MacKenzie, *Plato on Punishment*, p. 185 says Socrates is joking at 505c3-4. NB if Socrates only ever *punishes*, then he never *teaches*.

[28] They might point to how Callicles responds to Socrates' claim to be disciplining him--ὡς βίαιος εἶ!—but this is how Callicles sees his approach, because of its painfulness, not how Socrates sees it. An anonymous referee suggests that Socrates' mention of "persuasion and constraint" could be limited to the putative politicians under consideration. I have added the last sentence of the quoted passage to show that Socrates clearly describes the true politician's task as involving constraint, and argues

that neither present nor past Athenian politicians use those tools correctly, as the true politician would.

[29] See W. W. Goodwin, *Syntax of the Moods and Tenses of the Greek Verb* (London: Macmillan, 1889), §529 or Smyth, *Greek Grammar*, §2565.

[30] See Smyth, *Greek Grammar*, §2322. The conditionals at 480c-d also express future more vivid conditions, though with participles in the apodoses. Socrates starts accumulating circumstantial participles with διώκοντα (480c7) and μὴ ὑπολογιζόμενον, and continues with παρέχοντα (d1), which takes the infinitives τύπτειν (for beating) and δεῖν (for imprisonment), and then apparently takes the participles ἀποτίνοντα (d2), φεύγοντα (d2), and ἀποθνήσκοντα (d2-3), though the LSJ says this is rare (*s.v.* παρέχω, A. II. 2).

[31] He might ask the question twice to be sure that Polus has given up on admiring without qualification the ability to kill, exile, and confiscate. Were that his aim, though, he surely would not emphasize twice that they agree on this point; nor would he go on to articulate his own dividing line.

[32] An anonymous referee suggests that these might be used as preconditions for discursive punishment, much as imprisonment is used to enable discursive punishment in *Laws* X. I agree that they might be, but I see no good reason to think that Socrates envisions this as their only proper judicial use.

[33] Brickhouse and Smith, *Socratic Moral Psychology*, §4.2.4 use this to explain why certain punishments *do not harm* the person who is punished. I say rather that this is the

means by which non-discursive punishments remove injustice from the souls of the unjust and so *benefit* them. They use the language of benefit too, but only in passing (p. 115), and they surely do not think this can account for the putative benefits of a whipping.

[34] To this extent, Brickhouse and Smith, *Socratic Moral Psychology*, p. 124 are wrong to say that effective punishment must operate through the punished person's conception of the good. At least some punishment works by altering the agent's circumstances and not (directly) her attitudes.

[35] Socrates describes incurables as having nothing healthy, everything warped, and nothing straight (524e-525a; cp. *Tht*. 173a-b). Brickhouse and Smith, *Socratic Moral Psychology*, §4.2.7 say they can no longer "manage, rule, deliberate, and all other such things" (*R*. 353d3-6). They have some true evaluative beliefs, but cannot apply them to action. Punishment does not work as it does on curables; it neither (i) presents them with reasons not to commit injustice nor (ii) triggers calculation, making them aware of reasons not to commit injustice. They also say that incurables engage in means-end reasoning, but can no longer reason about ends (p. 129n23). But means-end reasoning sounds like a form of management, rule, or deliberation—indeed, the form that allows one to appreciate punitive reasons not to commit injustice (e.g. to avoid a beating). Further, it is unclear how beating curables is supposed to prompt them to reason better about ends.

[36] Socrates says earlier that rhetoric is not a craft, in contrast with what he says about sailing here; however, he soon redescribes ordinary rhetoric as a subordinate craft (517c-518b).

[37] J. Annas, 'Plato's Myths of Judgment', *Phronesis*, 27 (1982), pp. 119-143, p. 119 says the myth must be relevant to "the dialogue's main moral argument", and I agree. But she also says the myth focuses on punishment because Plato cannot defend justice without appealing to post-mortem punishment, and here I disagree.

[38] Rhadamanthus deems souls curable or incurable in advance of any punishment (526b4-c1), presumably to decide who receives what punishments. But that is consistent with the claim that both curables and incurables are punished only discursively; there are different forms of discursive punishment.

[39] More particularly, Socrates thinks the *real* magnitude of pain involved in confronting the bad condition of one's soul is greater than the real magnitudes of bodily and reputational pains. The latter, like bodily and reputational pleasures, *seem* greater because of contrast effects; cf. J. C. Shaw, *Plato's Anti-Hedonism and the Protagoras* (Cambridge: Cambridge University Press, 2015), Ch. 6. Hence, incurables suffer the superlatively "most grievous, intensely painful, and frightening sufferings" (τὰ μέγιστα καὶ ὀδυνηρότατα καὶ φοβερώτατα πάθη; 525c).

[40] See also *Republic* I, where Socrates argues that the just person would never harm anyone.

[41] See especially J. Hampton, 'The Moral Education Theory of Punishment', *Philosophy and Public Affairs*, 13 (1984), pp. 208-238. Hampton gives a better account than do Brickhouse and Smith of how pain by itself might morally educate, and she offers reasons to reject corporal punishment and the death penalty (p. 223; cf. pp. 231-232 on incurables). I cannot give a systematic and detailed comparison here.

[42] Thanks to Eric Brown, Fay Edwards, Rachana Kamtekar, Rachel Singpurwalla, and an audience at the Society for Ancient Greek Philosophy (especially Kelly Arenson, Emily Austin, and Catherine McKeen) for comments on earlier versions of this material. Thanks also to an anonymous referee for *Polis*, whose generous comments led to several changes in the final version. He or she also suggested more ambitious changes—e.g., saying more about punishment in the *Laws*, about the art of purification in the *Sophist*, and about whether there are in fact any incurables, and if so how we would know. I leave these questions for another occasion, or for others to explore.