THE COHERENCE OF THRASYMACHUS

RALPH WEDGWOOD

There is a passage in Book I of the Republic, or so I shall argue, where Plato aims to horrify us. In the character, beliefs, and desires of Thrasymachus, he aims to personify some of the most terrible dangers that the human soul confronts us with. In this way, the role that Thrasymachus plays for Plato is akin to the role that for Hobbes is played by the bellum omnium contra omnes, the war of all against all, which would allegedly be the inevitable result of a ‘state of nature’, where human beings have no government to terrorize them into obedience. It is also akin to the role that for Kant is played by the ‘radical evil’ that is allegedly an indelible feature of human nature itself. As I shall suggest, the desires that characterize Thrasymachus are what are described in Book IX as ‘lawless’ desires, desires of the insane kind that dominate the tyrannical soul; and his beliefs reflect these lawless desires, in an alarmingly systematic and coherent way.

Many recent commentators on the Republic take a rather different view. According to some of these scholars, Thrasymachus is just a muddle-headed oaf, with no clear coherent position of his own;¹ others describe him as simply an ill-mannered blowhard—missing the element of

horrors that I shall try to identify here.\(^2\) According to the reading that I shall argue for here, Thrasy
machus’ beliefs and desires are horrifying, but they are only a particularly stark and radical form of the values that underlie the challenge that Glaucon and Adeimantus articulate in Book II. In this way, the whole central argument of the Republic is a response to Thrasy
machus. If that response fails, it would represent the triumph of an essentially Thrasy
machean position, with all the horrors that that entails. The stakes for Plato could not be higher.

1. Thrasy
machus’ account of justice

Thrasy
machus begins his contribution to the dialogue by defining what is just as what is to the advantage of the more powerful (τὸ τοῦ κρείττονος συμφέρον, 338c3).\(^3\) Over the course of the next six pages, he reiterates the very same definition three times:

i. ὡστε συμβαίνει τῷ ὀρθῶς λογιζομένῳ πανταχοῦ εἶναι τὸ αὕτῳ δίκαιον, τὸ τοῦ κρείττονος συμφέρον.

Anyone who reasons correctly will conclude that everywhere it is the same thing that is just—what is to the advantage of the more powerful (339a2).

ii. ὡστε, ὅπερ ὡς ἄρχης ἔλεγον, δίκαιον λέγω τὸ τοῦ κρείττονος ποιεῖν συμφέρον.

So, as I said at the beginning, I say that it is just to do what is to the advantage of the more powerful (341a3–4).

\(^2\) For example, a recent description of Thrasy

\(^3\) The Greek text cited here is that of S. R. Slings (ed.), Platonis Rempubicam ..., (Oxford, 2003); the translations are my own modifications of G.M.A. Grube (trans.), Plato, Republic, second edition, revised by C.D.C. Reeve (Indianapolis, 1992).
As I said at the beginning, what is just is what is to the advantage of the more powerful, while what is unjust is profitable and advantageous to oneself (344c7).

Since Thrasygachus states this definition of justice no fewer than four times in total, it seems most plausible to interpret this as his official account of justice. In this section, I shall offer an interpretation of what Thrasygachus means by defining justice in this way; in the following section, I shall propose an interpretation of his reasons for doing so.

The discussion of justice in Book I seems to be focused on *just actions*. Starting in Book II, Plato will turn away from an act-focused conception of justice. Instead, Plato holds, the true nature of justice is revealed, not by any common feature of the external actions that count as just, but by the inner functioning of the just person’s soul. Here in Book I, however, the focus appears to be on the justice of actions. According to the first definition—which Socrates treats as having been given by Cephalus—justice consists in telling the truth and returning what is owed (331c1). According to Polemarchus’ definition, it consists in helping one’s friends and harming one’s enemies (332d2–5). In a similar way, I propose, according to Thrasygachus’ definition, justice consists in doing what is to the advantage of the more powerful (338c3).

Some commentators have been puzzled by the following feature of Thrasygachus’ position. According to the official account, to act justly is to do what is to the advantage of the more powerful. But suppose that you are *yourself* the tyrant, and hence the most powerful of all

---

4 In particular, the text seems to me to make this interpretation more promising than the proposal of Kerferd (‘Doctrine’, 25), that Thrasygachus’ fundamental definition of justice is as ‘another’s good’—a formulation which appears only once, long after the formula ‘to the advantage of the more powerful’ has been repeated several times.
persons whose interests you can affect in any way. Then, surely, Thrasymachus should say that for you to act justly would be for you to do what is to your own advantage (since in this case, you yourself count as ‘the more powerful’). Yet he does not say this. On the contrary, he says that the tyrant acts unjustly (344a4). Does this point show that Thrasymachus is simply confused and incoherent?

In fact, however, there is a simple solution to this puzzle. The perspective from which just acts are defined is clearly that of the person who is doing the just act. Thus, according to Cephalus’ definition, for me to act justly is for me to return what I owe, to those to whom I owe the debts in question; according to Polemarchus, for me to act justly is for me to help my friends and to harm my enemies; and according to Thrasymachus, for me to act justly is for me to act in a way that is to the advantage of the more powerful. Presumably, what Thrasymachus means by this is that for me to act justly is for me to act in a way that is to the advantage of someone who is more powerful than I am—and more generally, for any person x to act justly is for x to act in a way that is to the advantage of someone who is more powerful than x.5

---

5 As the editor of this journal has pointed out to me, this interpretation of Thrasymachus’ definition, as the thesis that for x to act justly is for x in some way to serve the interests of those who are stronger than x, has already been independently articulated by two other scholars—specifically, by C.D.C. Reeve, ‘Glaucon’s Challenge and Thrasymanechnanism’ ['Thrasymanechnanism'], *Oxford Studies in Ancient Philosophy* 34 (2008), 69–103, at 95, and by R.T. Long, 'Thrasymanechnus and the Relational Conception of Authority' ['Relational'], in P. Hanna (ed.), *An Anthology of Philosophical Studies*, Vol. 3 (Athens, 2009), 27–36, at 30. There are some subtle differences between Reeve and Long; see note 7 below.
If this is correct, it immediately explains why Thrasymachus later claims that justice is ‘the good of another’ (ἀλλότριον ἀγαθόν, 343c3). According to Thrasymachus’ definition, whenever any person \( x \) acts justly, that person \( x \) is doing something that is to the advantage of someone who is more powerful than \( x \). Since \( x \) cannot be more powerful than himself, it follows that whenever \( x \) is acting justly, \( x \) is doing something that is to the advantage of someone other than \( x \).

At all events, Thrasymachus’ claim that justice is the ‘good of another’ implies only that whenever any agent acts justly, that act is to the advantage of someone other than the agent. It does not imply the converse—the thesis that every act that is to the advantage of someone other than the agent is just. For example, suppose that you were to do something that serves the interests of someone other than yourself who is not more powerful than you. Then this act need not be just—unless it happens also to serve the interests of someone who is more powerful than you. In this way, I propose, the claim that justice is the ‘good of another’ is not designed to be a further definition of justice; it is designed merely to articulate a necessary feature that all just acts

---

\(^{6}\) For the claim that the initial definition of justice as ‘what is to the advantage of the more powerful’ conflicts with the claim that justice is ‘the good of another’, see Everson, ‘Incoherence’, 114–5. Before Everson, a similar claim was made by J. Annas, *An Introduction to Plato’s Republic* (Oxford, 1981), 45. T.D. J. Chappell, ‘The Virtues of Thrasymachus’ ['Virtues'], *Phronesis* 38 (1993), 1–17 at 7ff., relies on this passage to argue that Thrasymachus does not regard justice as a vice, or injustice as a virtue. Chappell’s argument depends on the assumption that if the most powerful person acts to his own advantage, he is doing what is ‘to the advantage of the more powerful’, and so acting justly. However, according to the interpretation that I am here defending as the simplest solution to the puzzle that makes Thrasymachus seem incoherent, all these arguments rest on a misinterpretation of Thrasymachus’ formula: according to this interpretation, for someone to do what is ‘to the advantage of the more powerful’ is for him to do what is to the advantage of someone who is more powerful than he is.
possess. This claim is not only perfectly compatible with the official definition of justice as *what is to the advantage of the more powerful*; it is an immediate corollary of that definition.

Thrasymachus’ account of justice has a further corollary as well. Suppose that you are yourself the tyrant, and so the most powerful of all people whose interests you can affect in any way. Then it is clearly no longer possible for you to do *anything* that serves the interests of anyone who is more powerful than yourself. In that case, given this definition of justice, you no longer have any opportunity for acting justly at all. So it is clear why Thrasymachus does not say that the tyrant acts justly: on his conception of justice, the tyrant has no opportunity for acting justly at all.\(^7\)

Thrasymachus only explicitly offers a definition of *justice*. But he seems to have an implicit conception of *injustice* as well. As I shall argue later, his background assumption seems to be that in any significant human interaction, there are two parties—one party that emerges as the *winner*, and another that emerges as the *loser*. He also seems in effect to be ignoring cases where one party wins by some lucky accident rather than intentionally, through their greater power or skill. Thus, he assumes that the winning party is always the more powerful, and the losing party is the weaker. For \(x\) to act unjustly, then, according to this conception, is for \(x\) intentionally to

---

\(^7\) Reeve (‘Thrasymacheanism’, 95) appears to interpret Thrasymachus to mean that for \(x\) to act justly is for \(x\) to obey the laws created by *everyone* who is stronger than \(x\); he concludes that, if one is oneself the strongest, and one obeys one’s own laws, then one’s action is just, because it ‘is—albeit vacuously—advantageous for everyone stronger than oneself, since that class is empty’. Long (‘Relational’, 31) seems to interpret Thrasymachus’ definition as implying that for \(x\) to act justly, \(x\) must serve the interests of *someone* who rules over \(x\); so, as he says, ‘the ruler could act justly only by submitting to some still further ruler over himself’. Since Thrasymachus describes the tyrant as unjust, Long’s interpretation seems preferable to Reeve’s on this point.
emerge as the winner—serving their own interests, and not the interests of anyone who is more powerful than x.$^8$

This conception of injustice tallies nicely with what Thrasymachus says. First, it explains why he says that petty criminals are unjust (344b1–3), given that he also thinks that petty crime is beneficial to the perpetrator (348d6). Secondly, it explains why he says that the person who commits the most ‘complete injustice’ (344a4)—the tyrant—is supremely happy and blessed (344b5), given that he thinks that, more than anyone else, the tyrant promotes his own interests, and not the interests of anyone who is more powerful than him. Finally, it explains why he thinks quite generally that ‘injustice is profitable and advantageous to oneself’ (344c8).

It seems reasonable to assume that Thrasymachus’ definitions of justice and of injustice are intended to give necessary and sufficient conditions for having the property in question.$^9$ On this interpretation, if someone sacrifices his own interests to help a weaker party, then by Thrasymachus’ definitions this act is neither just nor unjust. Is this a problem for this interpretation? Thrasymachus would probably regard such cases as rare and of marginal significance: normally, when one party gains at a second party’s expense, this is because the first party has somehow overpowered or manipulated the second. Moreover, as I shall argue, he also thinks that acting against one’s own interests is inconsistent with their manifesting practical intelligence (εὐβουλία, 348d1). In that sense, such acts are in his view simply stupid. It seems to cohere well with the rest of Thrasymachus’ view if he is committed to dismissing such marginal acts of stupidity as neither just nor unjust.

$^8$ On this interpretation, if someone sacrifices his own interests to help a weaker party, then by Thrasymachus’ definitions this act is neither just nor unjust. Is this a problem for this interpretation? Thrasymachus would probably regard such cases as rare and of marginal significance: normally, when one party gains at a second party’s expense, this is because the first party has somehow overpowered or manipulated the second. Moreover, as I shall argue, he also thinks that acting against one’s own interests is inconsistent with their manifesting practical intelligence (εὐβουλία, 348d1). In that sense, such acts are in his view simply stupid. It seems to cohere well with the rest of Thrasymachus’ view if he is committed to dismissing such marginal acts of stupidity as neither just nor unjust.

$^9$ T.D.J. Chappell, ‘Thrasymachus and Definition’, Oxford Studies in Ancient Philosophy 18 (2000), 101–7, at 103, questions whether Thrasymachus’ definition is intended to be of this form, pointing out that this strict conception of what definitions must do was peculiar to Socrates and his followers. But the real question is not about the historical orator Thrasymachus, but about the literary character who appears under the name ‘Thrasymachus’ in Plato’s Republic. As I shall argue, Plato is presenting this character as a radical alternative to the view of justice that he regards in the spirit of Socrates’ philosophical mission. For this reason, it seems reasonable to read Thrasymachus as offering an answer to the strict definitional question that Socrates was concerned with.
assumption, however, these definitions are highly revisionary—that is, they have consequences that diverge sharply from what would strike most people as intuitively plausible. Suppose that someone defies the most powerful person in the community, in order to protect the rights of an innocent person whom the powerful person wishes to oppress. As an example, we could consider the way in which Socrates—at least according to the *Apology* (32c4–d8)—disobeyed the Thirty Tyrants when they ordered him to arrest Leon of Salamis. According to Thrasymachus’ definition, since Socrates’ disobedience is not to the advantage of anyone who is more powerful than the agent, it is not really a just act at all; anyone who thinks that it is just is mistaken.¹⁰

Moreover, Thrasymachus’ definition also implies that whenever you do something that is to the advantage of the more powerful, you are acting justly. But this too seems extremely revisionary. Suppose that the tyrant orders you to become his accomplice in violating what most people would regard as the rights of others. In this case, if you comply with the tyrant’s order, and thereby act in a way that is to the advantage of the most powerful, Thrasymachus’ definition would imply that your act was *just*—which would strike many people as another highly implausible implication.

Finally, Thrasymachus’ conception of injustice also seems revisionary. I have suggested that he conceives of an unjust act as any act that intentionally serves the interests of the agent himself and not the interests of anyone who is more powerful than the agent. But this definition implies that whenever anyone evades the tyrant’s oppression, this act of evasion is also unjust—which seems to be yet another counterintuitive implication of this definition. (However, since as we

¹⁰Everson (‘Incoherence’, 118) discusses this sort of apparent counter-example, which he takes to show how confused and muddle-headed Thrasymachus is. As I shall argue in the next section, we should not take the mere fact that Thrasymachus’ account is so revisionary as evidence that he is confused—indeed, as we shall see, Plato’s own account of justice is also seriously revisionary in its own way.
shall see, Thrasymachus regards injustice as a *virtue*, the fact that evading the tyrant’s oppression is unjust does *not* in his view imply that one *ought not* to resist.)

Given that Thrasymachus’ definition of justice is so revisionary, we need to understand what reasons he could possibly have for endorsing it. I shall investigate these reasons in the next section of this paper.

Since Thrasymachus offers such a revisionary conception of justice, it would not be surprising to find that he also has a revisionary conception of ‘advantage’, and perhaps even of ‘power’, as well. In fact, it seems plausible that his conception of what is ‘advantageous’ to human beings is also strikingly revisionary. He appears to argue that since my acting justly benefits someone more powerful than me, it *cannot* simultaneously benefit me; on the contrary, my acting justly must actually *harm* me. As he says:

> ἡ δικαιοσύνη καὶ τὸ δίκαιον ἄλλοτριον ἁγαθὸν τῷ ὄντι, τοῦ κρείττονός τε καὶ ἄρχοντος συμφέρον, οἰκεία δὲ τοῦ πειθομένου τε καὶ ὑπηρετοῦντος βλάβη.

Justice, what is just, is in reality the good of another—what is to the advantage of the more powerful and the ruler—but self-harm to the one who obeys and serves (343c3–5).

Similarly, the only acts that he ever recognizes as ones that actually benefit the agent are acts of ‘outdoing’ other people (*πλεονεκτεῖν*, 344a1), such as cheating people out of what one is contractually obliged to do for them (343d3), or, like the tyrant, kidnapping, enslaving, and appropriating the possessions of one’s fellow citizens (344b4).

Why does Thrasymachus argue in this way? It seems to be because he believes that, quite generally, the interests of distinct human beings are necessarily antagonistic or opposed to each
other. In effect, all human beings are locked into a constant ‘zero sum game’—that is, a situation in which the only way in which anyone can benefit to any significant extent is at someone else’s expense. There is no prospect of a ‘win–win’ outcome, in which everyone is made better off than they would have been had they not interacted at all. In every social interaction in which anyone’s interests are significantly affected at all, one party is the winner, and the other party is the loser: that is, one party—the victor—ends up better off than they would have been if the interaction had never occurred, while the other—the defeated party—ends up worse off than they would have been had the interaction never occurred.

This conception of human interests lies behind many of Thrasymachus’ claims. As I have suggested, this conception of interests leads him to claim that acting justly is a kind of self-harm. This in turn leads to Thrasymachus’ remarkable further claim, that injustice is a virtue while justice is a vice. At 348c5–10, he explicitly accepts that out of the pair consisting of justice and injustice, one quality is a virtue while the other is a vice, and yet denies that justice is virtue or that injustice is a vice, stating instead that his claim is ‘the opposite’—that is, the claim that injustice is a virtue and justice a vice. To confirm that he has understood it correctly, Socrates repeats the claim (348e2):

&alpha;λλά τόδε ἔθαύμασα, εἰ ἐν ἀρετῆς καὶ σοφίας τιθεῖς μέρει τὴν ἑδικίαν, τὴν δὲ δικαιοσύνην ἐν τοῖς ἐναντίοις.

But I found this astonishing, that you place injustice in the category of virtue and wisdom, and justice among their opposites.

---

Thrasyilmachus explicitly confirms Socrates’ interpretation of his claim: ‘That is precisely where I place them’ (ἀλλὰ πάνυ οὖτω τίθημι, 348e3).\footnote{So why does Thrasyilmachus balk at assenting to the thesis that justice is a vice at 348c9–10? Perhaps it is merely because Thrasyilmachus is not completely happy with the moralistic connotations of the word ‘vice’, and so prefers to use terms like ‘naivety’ instead. Chappell (‘Virtues’, 11) argues that Thrasyilmachus does not regard justice as a vice—but instead regards justice as a ‘device’ that the powerful use to manipulate the weak. I certainly would not dispute the claim that Thrasyilmachus regards justice as useful to the powerful in this way; but that is compatible with his also regarding justice as a vice. In fact, as I explained in note 6 above, Chappell’s main argument for the conclusion that Thrasyilmachus does not regard justice as a vice rests on a reading of the formula ‘the advantage of the more powerful’ which is incompatible with the interpretation that I have defended here.}

Why exactly does Thrasyilmachus regard justice as a vice, and injustice as a virtue? It seems to be because he believes that unjust action reveals practical intelligence (εὐβουλία, 348d1), and unjust agents are prudent (φρόνιμοι, 348d2), while only fools act justly (348c12). But why does he believe that? At this point, as often happens with Plato, we are forced to speculate. But the following speculation seems promising to me. Suppose that Thrasyilmachus accepts two further principles. First, there is the principle that it is never the case that when one’s action manifests practical intelligence, one is acting incorrectly—that is, acting as one has decisive reason not to act. Secondly, there is the ‘rational egoist’ principle that, all things considered, one has always decisive reason not to perform any act that harms one’s own interests. Together, these two principles entail that no act that harms the agent’s own interests can manifest practical intelligence. So, since Thrasyilmachus holds that acting justly harms the agent’s own interests, he is committed to the view that acting justly never manifests practical intelligence. This may be the best explanation of why Thrasyilmachus thinks that justice is not a virtue.
If this is the correct explanation of why Thrasymachus thinks that justice is not a virtue, it may be, more specifically, that he takes the view that justice always reveals a kind of ignorance. He would be committed to this view if he accepts the Socratic principle that no one knowingly acts incorrectly—that is, no one knowingly acts as they have decisive reason not to act.\textsuperscript{13} If no one willingly acts incorrectly except through ignorance, and it is always incorrect to act justly—that is, we always have decisive reason not to act justly—it follows that anyone who willingly acts justly must be doing so out of ignorance.

On this interpretation, then, Thrasy machus holds that virtuous individuals always pursue their own interests, and the interests of distinct individuals are invariably antagonistic to each other. In this way, his view entails that virtuous human beings are in effect always locked into a bellum omnium contra omnes, a war of all against all. The only way in which anyone can ever act correctly—that is, as a person of practical intelligence would act—is by advancing his own interests; and the only way in which anyone can significantly advance his own interests is at the expense of others—that is, by emerging as the winner of the interaction, and forcing the other party into the position of the loser.

Why does Thrasy machus hold this shockingly bleak view of human life? It is tempting to speculate that he believes that out of all the goods that are available in human life, the dominant and most supremely beneficial good is power over others. He speaks of the ‘more powerful’ person, the ruler, as the most blessed of all (344b5), which suggests that he believes that what is of greatest benefit to anyone is simply to be, and to remain for as long as possible, in this position of dominating others.

\textsuperscript{13} For places where Socrates appears to endorse this principle, see Protagoras (354e–356c), Gorgias (466a–472e) and Meno (77b–78c).
It seems that his conception of ‘power’ is tied to the idea of a hierarchy of domination and subordination: for you to have power over others is for those others to be your subordinates in this social hierarchy. This seems to be why, when he praises successful injustice, it is as ‘stronger and freer and more masterly’ (ἰσχυρότερον καὶ ἐλευθεριώτερον καὶ διεσποτικώτερον) than justice (344c4): the successfully unjust person like the free slave-owning master, while the just person is like an unfree powerless slave.¹⁴ Since a dominant position in a social hierarchy is an essentially positional good (it is determined by one’s position in relation to others), the logic of such positional goods is necessarily that of a zero-sum game: you can only improve your position with respect to such positional goods by worsening the position of others.

If this speculation is along the right lines, then there may even be a serious thought behind Socrates’ seemingly playful question about whether by ‘the more powerful’, Thrasydamus means someone like Polydams the pancratiast, who was famous for his physical strength (338c8). There are some forms of strength or power with respect to which it is perfectly possible for every single member of society to become more powerful. This could be true of the kind of power that comes from physical strength: if a community instituted a rigorous diet and training program for all its citizens, it would be possible for everyone in the community to become physically stronger. It could also be true of the kind of power that comes from knowledge: if a society instituted a well-designed education system, every single member of the society might become more knowledgeable, and in that way more powerful. This could not be true, however,

¹⁴ Chappell (‘Virtues’, 13) suggests translating ἐλευθερία as ‘unrestraint’; but it seems more plausible that the central connotations of this word in early 4th century Greek were with the social status of being free rather than a slave. As is argued by K. Raaflaub, The Discovery of Freedom in Ancient Greece, first English edition, trans. R. Franciscono (Chicago, 2004), 261, a conception of ἐλευθερία that was typical of the ‘late Periclean period’ held that ‘only the master who ruled over others was truly free.’
for power over others: if one member of a society gains power over others, then some others must lose. It seems to be power of this latter kind that Thrasymachus regards as the dominant component of an individual’s good.

Thrasymachus also develops a slightly more precise and detailed conception of this position of being in power over others. In his view, it is not a position that someone could occupy while being asleep or in a coma for years. In this way, having power over others is not like having a social rank or position. It necessarily involves actively ruling over others.

Thrasymachus insists that ruling over others requires intelligence and good judgment, and in response to the objection (339c–e) that rulers are fallible and liable to make errors, he quickly resorts to the idea that ruling is a skill. He compares the ruler to a craftsman (δημιουργός, 340e4), and he implies that being a craftsman requires the deployment of knowledge (ἐπιστήμη, 340e3). If ruling is indeed a skill, it is perfectly reasonable for Thrasymachus to accept that for you to be genuinely more powerful than me, you must be exercising the skill of ruling over me.

With every genuine skill, as Thrasymachus and Socrates both agree, there is a standard of correct performance. Genuinely exercising the skill involves meeting the standard; those who fail to meet the standard are at best attempting but failing to exercise the skill (340d1–341a2).

Presumably, the skill that Thrasymachus has in mind is broadly of the same kind as the skill that Machiavelli advised autocratic rulers to acquire—the skill of securing and reinforcing one’s position of dominating others, and keeping them subject to one’s will, for as long as possible. At all events, Thrasymachus’ idea of the correct exercise of the skill of ruling seems quite compatible with his central definition of what it is for you to act justly, as your doing what is to

---

the advantage of the more powerful—that is, to the advantage of someone who is exercising the skill of ruling over you.

According to Thrasymachus, then, truly virtuous agents will act unjustly, promoting their own interests, by exercising the skill of ruling over others and subjecting others to their power—while just agents will foolishly act in ways that advantage those who rule over them. As we have already remarked, this is an extremely revisionary conception of justice and virtue—which yet more pressingly raises the question: How does Thrasymachus defend this conception?

2. Thrasymachus’ defence of his account

The fact that a philosopher defends a revisionary conclusion does not immediately show that the philosopher is confused. When Hume attacks the ‘monkish’ virtues, such as ‘celibacy, fasting, penance, mortification, self-denial, humility’, and the like, and concludes ‘we justly therefore transfer them to the opposite column, and place them in the catalogue of vices’, the overall form of Hume’s conclusion—asserting that what a long and distinguished tradition of ethical thought has deemed to be a virtue is really a vice—is essentially the same as that of Thrasymachus’.16 A similar point could be made about the ‘revaluation of all values’ (Umwertung aller Werte) that was advocated by Nietzsche.17 It would certainly be too quick to conclude at once that Hume and Nietzsche are simply confused.

---

16 See D. Hume, Enquiry concerning the Principles of Morals, revised edition (London, 1777), Section 9, ¶3.
17 See F. Nietzsche, Der Antichrist (Leipzig, 1894), §13. For the interpretation of Thrasymachus as a proto-Nietzschean, see P. Shorey (ed.), Plato: The Republic, with an English translation, revised edition, Loeb Classical Library (Cambridge, Massachusetts, 1946), x. Chappell (‘Virtues’, 12) rightly points out that Thrasymachus is not strictly revaluing all values—since his valuing of freedom and practical intelligence is eminently in line with
Moreover, the view that Plato himself will be defending later in the *Republic* also turns out to be significantly revisionary—or at least at variance with the ethical views of the great majority of his own contemporaries. At the beginning of Book II, Glaucon and Adeimantus lay out what they describe as the opinions that ‘most people’ have about justice and injustice (358a3). It is clear that Plato sets out to show these opinions are wrong. But neither Socrates nor any other character in the dialogue ever disputes the attribution of these opinions to ‘most people’. So it seems clear that Plato himself would accept that his own view of justice are inconsistent with that of ‘most people’.

Revisionary philosophical accounts of this kind face a distinctive sort of objection—the objection that they are in fact covertly *changing the subject*, by tacitly switching from using the central terms in their ordinary everyday sense to using them in some special technical sense instead. This is the basic idea behind the objection, which was notably pressed by David Sachs, that Plato himself is guilty of a ‘fallacy of irrelevance’.

To answer this sort of objection, something needs to be said to dispel the concern that the revisionary theory is no longer using these crucial terms in their ordinary everyday sense.

As I shall argue, Thrasymachus attempts to address this sort of concern in 338d6–10, where he gives his most explicit defence of his account of justice. In this passage, he first explains that in each city—whether it is a tyranny, a democracy, or an aristocracy—the ruling class is what counts as the ‘more powerful’ in the relevant sense. Then he continues with the following claims (338e1–339a4):

---

traditional views. (The same can surely be said of Nietzsche too, however. To that extent, the phrase ‘Umwertung aller Werte’ is hyperbole.)

---

τίθεται δέ γε τοὺς νόμους ἐκάστη ἢ ἀρχή πρὸς τὸ αὑτῇ συμφέρον, δημοκρατία μὲν δημοκρατικοὺς, τυραννίς δὲ τυραννικοὺς, καὶ αἱ ἄλλαι οὕτως: θέμεναι δὲ ἀπέφηναν τοῦτο δίκαιον τοῖς ἀρχομένοις εἶναι, τὸ σφίσι συμφέρον, καὶ τὸν τούτου ἐκβαίνοντα κολάζουσιν ὡς παρανομοῦντα τε καὶ ἀδικοῦντα. τοῦτ’ οὖν ἐστιν, ὃ βέλτιστε, ὃ λέγω ἐν ἀπάσαις ταῖς πόλεσιν ταῦτάν εἶναι δίκαιον, τὸ τῆς καθεστηκυίας ἀρχῆς συμφέρον: αὕτη δὲ που κρατεῖ, ὡστε συμβαίνει τῷ ὀρθῶς λογιζομένῳ πανταχόω εἶναι τὸ αὐτὸ δίκαιον, τὸ τοῦ κρείττονος συμφέρον.

And each ruling class makes laws to its own advantage. Democracy makes democratic laws, tyranny makes tyrannical laws, and so on with the others. And in making these laws, they declare that this—what is to their advantage—is just for their subjects, and they punish anyone who deviates from this as lawless and unjust. This, then, Socrates, is what I say is just: it is the same in all cities, the advantage of the established rule. Since the established rule clearly has the power, anyone who reasons correctly will conclude that everywhere it is the same thing that is just—what is to the advantage of the more powerful.

Several recent scholars have argued that the defence that Thrasymachus gives of his account of justice in this passage implicitly seeks to make it clear how he is not changing the subject, but using the term ‘justice’ so that it refers to the very same phenomenon that is commonly referred to by that term in everyday language. For example, C.D.C. Reeve argues that in this passage, Thrasymachus first presents a ‘nominal definition’ of what ‘justice’ means in a particular city—
as consisting in ‘obedience to [the city’s] laws’—and then moves from this nominal definition to a ‘real definition’ of justice as ‘what is advantageous to the stronger’.\textsuperscript{19}

It certainly seems plausible that Thrasymachus’ account of justice as ‘what is to the advantage of the more powerful’ is intended as a ‘real definition’ of the \textit{essence} of justice (like ‘water is H\textsubscript{2}O’ or the like), rather than as a conceptual analysis of the ordinary meaning of the word ‘justice’, or as a purely stipulative redefinition of the term.\textsuperscript{20} However, I shall argue that it is not plausible to read Thrasymachus as presenting ‘obedience to the city’s laws’ as a nominal definition of ‘justice’. The key point in this passage, as I shall argue, is that Thrasymachus is treating certain sociological facts—about how in every city the rulers make laws that are to their advantage, and ‘declare’ that it is just for their subject to obey these laws—as \textit{evidence} for his account of justice. To rebut the charge of ‘changing the subject’ in giving a real definition of a phenomenon, it is not necessary to give a full-blown nominal definition of the word that is being used to pick out the phenomenon. It is enough if the real definition is defended on the basis of what is \textit{generally agreed} to be \textit{relevant evidence} about that phenomenon.

In this passage, it seems clear that Thrasymachus presents what he takes to be evidence in support of his account of justice. This evidence consists of two alleged sociological facts: first, the fact that in every kind of city the ruling class ‘makes laws to its advantage’; and secondly, the fact that in every city the ruling class ‘declare’ that their laws are just. But how exactly does Thrasymachus take these two sociological facts as evidence in favour of his definition of justice, as what is to the advantage of the more powerful?

\textsuperscript{19} See Reeve, ‘Thrasymacheanism’, 85.

\textsuperscript{20} On this point, I am in agreement, not just with Reeve (‘Thrasymacheanism’, 92), but also with R.C. Cross and A.D. Woozley, \textit{Plato’s Republic: A Philosophical Commentary} (London, 1964), 38–41.
In the ensuing discussion, Socrates asks three times whether Thrasymachus has asserted that it is just for subjects to obey their rulers:

i. οὐ καὶ πείθεσθαι μέντοι τοῖς ἀρχουσιν δίκαιον φης εἶναι;
Don’t you also say that it is just to obey the rulers? (339b7)

ii. ἃ δὲ ἂν θώνται ποιητέον τοῖς ἀρχομένοις, καὶ τοῦτό ἐστι τὸ δίκαιον;
And whatever the rulers lay down must be done by their subjects, and this is what is just? (339c9)

iii. οὐχ ὀμολόγηται … ἃ δὲ ἂν προστάττωσιν οἱ ἀρχοντες δίκαιον εἶναι τοῖς ἀρχομένοις ποιεῖν;
Haven’t we agreed … that it is just for the subjects to do whatever the rulers order? (339d7)

On each of these three occasions, Thrasymachus answers affirmatively. Socrates repeats the point again at 339e4:

iv. τοῖς δὲ δίκαιον εἶναι φῆς ταῦτα ποιεῖν ἃ ἐκεῖνοι προσέταξαν.
You say that it is just for the [subjects] to do what [the rulers] order.

It is at least somewhat puzzling that Socrates interprets Thrasymachus in this way, since Thrasymachus did not in fact say at 338d6–10 that it is just for subjects to obey their rulers. On the contrary, he said only that the rulers make a certain declaration about the laws—namely, the declaration that it is just for their subjects to obey these laws and unjust for them to disobey. This is a quite different claim—and indeed a much more plausible claim. Arguably, nothing counts as a ‘law’ at all unless the authority that promulgates the law also represents the law as something
that those who are subject to the law have a reason of justice to obey.\textsuperscript{21} Presumably, the claim that Thrasymachus actually makes, about what the rulers ‘declare’, will only support any conclusions about the nature of justice if it is also true that when the rulers ‘declare’ that a certain act is just, that is evidence that the act in question really is just.

It is not clear from the text what reasons Thrasymachus has for treating the rulers’ declarations about justice as evidence of what really is just. There are in fact many reasons that philosophers might have for thinking that these declarations are evidence in this way. For example, it may be that quite generally, as Karl Marx and Friedrich Engels said in \textit{The German Ideology}, ‘The ideas of the ruling class are in every epoch the ruling ideas: i.e., the class which is the ruling material force of society is at the same time its ruling intellectual force.’\textsuperscript{22} It may also be that these ‘ruling ideas’ are guaranteed, by the very nature of the concepts involved, to be evidence of what really falls under those concepts. (This could be because the reference of each of these concepts is partly determined by the requirement that it should be a ‘charitable’ interpretation, which makes these ‘ruling ideas’ as close as possible to the truth.) At all events, Thrasymachus himself never offers any account of why the rulers’ declarations about what is just should be evidence of what really is just; but he assumes that Socrates and the other interlocutors will agree that these declarations are evidence in this way. If this is generally agreed, that is enough to rebut the charge that Thrasymachus is changing the subject.

\footnote{21} This point is accepted by many different philosophers of law. These philosophers of law clearly include the ‘natural law’ theorists like J. Finnis, \textit{Natural Law and Natural Rights} (Oxford, 1980); but they also include at least those legal positivists, like H.L.A. Hart, \textit{The Concept of Law}, revised edition (Oxford, 1994), who hold that it is an essential feature of law that there are certain key legal officials who take an ‘internal view’ on the law.

Suppose that the rulers’ declaration that a certain act is just is indeed evidence that the act really is just. Then we can use these declarations to identify a set of putatively just acts. We can then search for the most fundamental unifying feature that (i) all or at least most of these acts have in common, and (ii) typically causally explains the ruler’s declarations that these acts are just. This feature will then be the deep phenomenon underlying our superficial evidence concerning justice; and it will be plausible to identify this feature with the real nature of justice.

According to Thrasymachus, the acts that the rulers declare to be just invariably include acts of complying with the laws that they have made. So we should search for the most fundamental unifying feature that (i) is common to all or at least most acts of compliance with the law, and (ii) typically causally explains the rulers’ declarations that these acts are just. According to Thrasymachus, at least typically, the acts that the rulers declare to be just are to the rulers’ advantage. Moreover, when this happens, it is no accident: the fact that these acts are to the rulers’ advantage causally explains the rulers’ declarations. So, the fundamental feature that we should identify with justice is the feature of being to the advantage of the ruler. This, I propose, is how Thrasymachus’ sociological claims count as evidence for his account of justice.

Socrates initially responds to this argument by interpreting it as treating facts about what rulers ‘declare’ to be just as conclusive evidence of what it actually is just for the rulers’ subjects to do. However, it turns out that according to Thrasymachus, the positive declarations of those who occupy the social position of rulers are not conclusive evidence. As we have seen, at 340d–341a Thrasymachus maintains that ruling is a skill, and those who occupy the social position of rulers are not always true rulers; they are not always genuinely exercising this skill. So some of the acts that are positively declared to be just by those who occupy the social position of rulers will fail to be to the rulers’ advantage, and so will not be just. Thus, in Thrasymachus’ view, the
positive declarations of those who have the social position of rulers are only *defeasible* evidence of what is just.

However, whenever *true rulers* order their citizens to act in a certain way, that order will be to the ruler’s own advantage. It also seems hard to see how a true ruler could give an order to his subjects without expecting the order to be obeyed. So, if a true ruler gives an order to his subjects, it must be to the advantage of the ruler for the subject to obey. Thus, it seems that Thrasymachus does accept that it is always just for subjects to obey what a *true* ruler orders them to do. On the other hand, there is no reason to think that he accepts the converse of this. That is, it seems quite possible that certain acts could count as just—that is, according to Thrasymachus, as being to the advantage of the more powerful—even though these acts do not count as ‘obeying’ an ‘order’ or ‘law’ that the ruler has made. Suppose for example, that without being ordered to, I give an effusive eulogy of the tyrant, lavishing praise on our great leader. In this case, I have done something that is to the advantage of the more powerful—and so also, according to Thrasymachus, something just—even though I have not obeyed any order or law that the ruler has given me.²³

Moreover, Thrasymachus never says that it is just for the *ruler* to obey the law. On my interpretation, indeed, the ruler has no opportunity to act justly at all. Whether or not the ruler obeys the law is irrelevant: since the ruler advances his own interest, at the expense of the interests of everyone else, he acts *unjustly*—quite regardless of whether he obeys the law.

---

²³ The editor of this volume has suggested the extremely pertinent example of what the Germans call ‘*vorauseilender Gehorsam*’—the tendency of German citizens during the Nazi era to carry out what they *expected* that the authorities would demand of them, even without receiving any explicit order. By Thrasymachus’ lights, if it genuinely served the *Führer*’s interests, such *vorauseilender Gehorsam* would be just.
Thus, there are three reasons for thinking that Thrasymachus would not accept any ‘legalist’ definition of justice as ‘compliance with the law’. Indeed, no such legalist account could serve as either a real definition or as a nominal definition, since both kinds of definition must consist of a formula that is at least *necessarily coextensive* with the *definiendum*. First, if ‘the law’ includes the positive laws made by those who occupy the social position of rulers, but are not *true* rulers, then complying with the law may fail to be to the advantage of the more powerful, and so would not be just. Secondly, although every act that complies with the orders of the true ruler is indeed just, the converse need not always hold: not every act that is just need be in compliance with any explicit order—for example, the fawning citizen who gives an effusive eulogy of the tyrant has acted justly even if he was not complying with any explicit order. Thirdly, for the ruler himself, obeying the law is not just at all: if the ruler is a true successful ruler, then whatever he does is unjust—which it conforms to the law or not.

In addition, there is a further reason why the legalist account could not serve as a *real* definition of justice. The law—which itself consists of the rulers’ declarations about what is just and what is unjust—is only a *symptom* of a deeper phenomenon, which is the ruler’s maintenance of his power and advantage over his subjects: it is this deeper phenomenon, and not the more superficial symptoms, that gives the true nature or essence of justice.24

---

24 Indeed, a ‘legalist’ definition of justice as *what the rulers declare to be just* would surely be subject to the very same objections that Plato developed against Euthyphro’s definition of piety as *what the gods love* (*Euthyphro* 10a–11a). It is implausible to interpret Plato as making Thrasymachus commit the very same fallacy as Euthyphro, given that (i) Plato has already refuted this fallacy in *Euthyphro*, and (ii) he never comments on this fallacy here in the *Republic*. 
Several commentators have read the text as revealing that Thrasymachus endorses such a ‘legalist’ (or ‘conventionalist’) view of justice.\textsuperscript{25} But in fact, there is little textual support for this reading. As we have seen, in three of the four places where Socrates restates Thrasymachus’ claims, Socrates says only that Thrasymachus has claimed that \textit{it is just for subjects to obey their rulers}. But this obviously does not entail that obedience to rulers is the definition of justice. It only entails that obeying one’s rulers is \textit{an example} of justice. In general, a statement of the form ‘It is just to φ’ only entails that φ-ing is an example of justice, not that φ-ing is the essence or definition of justice.

There is admittedly one passage that could possibly be read as expressing a ‘legalist’ definition of justice:

\begin{quote}
 ámb δ’ ἂν θὸνται ποιητέον τοίς ἄρχομενοις, καὶ τοῦτο ἐστί τὸ δίκαιον.
\end{quote}

And whatever they lay down must be done by their subjects, and this is what is just (339c8).

But in fact, this passage does not have to be read as a definition of justice. The phrase here translated ‘what is just’—‘τὸ δίκαιον’—means literally ‘the just thing’. There are contexts in which ‘the just thing’ can refer to the universal property of justice as such, but these are by no means the only contexts—or even the commonest contexts—in which this phrase can be used.

The phrase ‘the just thing’ is a definite description, singling out the unique just item from some contextually salient domain of quantification. (For example, in many contexts, there will be a particular agent and a particular time that are contextually salient, and this domain will include all and only the various courses of action that are available to that agent at that time; in these

\textsuperscript{25} Thrasymachus was interpreted as a legalist by Hourani, ‘Thrasymachus’ Definition of Justice in Plato’s \textit{Republic}, \textit{Phronesis} 7 (1962), 110–120. To my way of thinking, a completely convincing reply to Hourani on this point is given by G.B. Kerferd, ‘Thrasymachus and Justice: A Reply’, \textit{Phronesis} 9 (1964), 12–16.
contexts, ‘the just thing’ will refer to the unique just course of action available to that agent at that time.) Often, however, this phrase is used in a way that is implicitly generalized, so that it effectively quantifies over the items that count as ‘the just thing’ in each of a range of different cases. Sometimes—and perhaps especially in philosophical contexts—the generalization is unrestrictedly universal, so that the phrase ‘the just thing’ in effect quantifies over everything that counts as ‘the just thing’, in any case whatsoever. But in many other contexts, the generalization is not quite so universal, and covers a narrower range of cases instead.

At 339c8, it seems to me, the most natural interpretation is to take the relevant range of cases to include, not every possible case in which there is a unique just course of action, but only the cases in which the true rulers give some of their subjects an order. In this range of cases, obeying the order is ‘the just thing [for the subjects to do]’. There is no implication here that there could not be other cases in which one might act justly without obeying one of the rulers’ orders.

Why have so many commentators been tempted to read Thrasymachus as endorsing a legalist definition of justice at 338e1–339a4? My guess is that these commentators have correctly seen that in this passage, the rulers’ declarations about what is just are treated as evidence about the real nature of justice; and then these commentators assume that the rulers’ declarations about what is just could not be evidence about the real nature of justice unless justice is somehow defined in terms of the rulers’ declarations. On reflection, however, it is far from obvious that evidential relations must rest on any such definition (whether a nominal definition or a real definition or a conceptual analysis or anything of that sort). That is, it may be that the rulers’ declarations about what is just count as evidence of what really is just, even if justice cannot in any sense be defined in terms of complying with the rulers’ orders. In short, once we have a
larger conception of evidential relations, there is no longer any reason of any kind to see Thrasymachus as espousing a legalist conception of justice.  

3. The significance of Thrasymachus

I suggested at the outset that the figure of Thrasymachus plays a role for Plato that is akin to the roles played for Hobbes by the *bellum omnium contra omnes*, the war of all against all, and for Kant by the ‘radical evil’ that is allegedly an ineradicable feature of human nature. In each of these cases, the philosopher sets out to depict a particularly horrifying possibility for human life, and then suggests that this is only the most extreme and radical form of what is in effect the *only* viable alternative to the ideal that is being recommended in the philosophical work in question. For Hobbes, the war of all against all is a danger that constantly threatens human life, and the only effective way of minimizing this terrible risk is by following his prescriptions for organizing civil society. For Kant, human nature contains an indelible element of radical evil, and the only way to avoid succumbing to it completely is by following the *a priori* moral principles that he articulates. Similarly, as I shall argue, for Plato, the beliefs and desires of Thrasymachus represent a perennial terrible danger in human life; but they are merely a

---

26 Interestingly, Everson’s objections to Thrasymachus (‘Incoherence’, 102–3) are similar to the objections that G.E. Moore, *Principia Ethica* (Cambridge, 1903), 66–7, made to the ‘proof’ of utilitarianism that was given by J.S. Mill, *Utilitarianism* (London, 1879), Chapter 4, ¶3. Moore assumes that Mill must be *defining* the desirable as what is desired—whereas in fact Mill is merely assuming that the fact that something is desired is *evidence* for the conclusion that the thing in question is desirable. In a similar way, Everson assumes that Thrasymachus must be *defining* the just as what the rulers legislate—whereas in fact Thrasymachus is only assuming that the fact that something is commanded by the rulers is *evidence* for the conclusion that it is just.

particularly stark and radical form of what is in Plato’s view the only viable alternative to the very different beliefs and motivations that are defended over the course of the Republic as a whole.

The significance of Thrasymachus for the Republic as a whole is most clearly revealed when Glaucon raises his challenge at the beginning of Book II by stating that he is ‘renewing the argument of Thrasymachus’ (358b7). Admittedly, there are some respects in which Glaucon’s challenge seems to differ from Thrasymachus’ argument. First, while I have argued that Thrasymachus’ account of justice focused on individual just acts, it is arguable (albeit controversial) that Glaucon’s challenge fundamentally concerns, not just acts, but rather the state of having a just soul or character. Secondly, unlike Thrasymachus, Glaucon emphasizes the instrumental benefits to most people of being just—the ‘rewards and other consequences’ of justice, as he calls them (357c7). Nonetheless, Glaucon’s view of the intrinsic benefits of justice and injustice is precisely the same as Thrasymachus’. This is clear from the example of the two agents who can get the normal rewards of justice, and escape the normal consequences of injustice, without actually being just—the agent who like Gyges’ ancestor has a magical ring of invisibility (360a–d), and the unjust person who has equipped himself with ‘the greatest reputation for justice’ in the ‘comparison of lives’ passage (362b–c). Both these agents are described as acting in exactly the manner of Thrasymachus’ tyrant.28 In raising this challenge, Glaucon aims to articulate what he claims to be implicit in the beliefs of ‘most people’ (358a3). In this way, he suggests that the beliefs of ‘most people’ are actually closer to Thrasymachus’ revisionary views than they at first appear.

---

28 This point is rightly stressed by Reeve, ‘Thrasymacheanism’, 85.
Famously, the central argument of the *Republic*, from the middle of Book II until the end of Book IX, is designed to answer this challenge. Suppose that Plato has articulated the response to this challenge that is most likely to succeed. Then, if Plato’s response fails, the only viable alternative would be the rival broadly Thrasymachean view that Glaucon and Adeimantus advocate in making their challenge. In this sense, the only viable alternative to the ideals and values that Plato defends in the *Republic* is this broadly Thrasymachean view.

The point that the central argument of the *Republic* is an answer to Thrasymachus is emphasized at the beginning of Book VIII, when Socrates and Glaucon pick up the argument that they broke off at the beginning of Book V, and turn to giving a description of the four bad types of soul, which directly parallel the four bad types of constitution. Socrates begins this discussion by saying (545a2–b2):

```
ἄρ’ οὖν τὸ μετὰ τοῦτο διιτέον τοὺς χείρους … ἵνα … ἡμῖν τελέα ἡ σκέψις ἡ, πῶς ποτε ἢ ἀκρατος δικαιοσύνη πρὸς ἀδικίαν τὴν ἀκρατον ἔχει εὐδαιμονίας τε πέρι τοῦ ἔχοντος καὶ ἀθλιότητος, ἵνα ἢ Θρασυμάχῳ πειθόμενοι διώκωμεν ἀδικίαν ἢ τῷ νῦν προφαινομένῳ λόγῳ δικαιοσύνην;
```

Shouldn’t we next go through the inferior souls … in order to … complete our investigation into how pure justice and pure injustice stand, with regard to the happiness or wretchedness of those who possess them, and either be persuaded by Thrasymachus to practise injustice or by the argument that is now coming to light to practise justice?

Coming at this crucial point in the structure of the *Republic*, this passage serves to remind us that the whole central argument of the *Republic* is a response to Thrasymachus.

If the central argument of the *Republic* is a response to Thrasymachus, why does his argument have to be ‘renewed’ by Glaucon and Adeimantus? A complete answer to this question
would take us too far afield. I suggest, however, that the problem is not fundamentally with Thrasy
aphus’ argument, but with the counter-arguments that Socrates gives against Thrasy
naphus in Book I. For some reason, Plato views these counter-arguments as merely provision
al, and in need of further elaboration and clarification. This is why it seemed neces
sary to Plato to continue the debate between the positions of Thrasy
naphus and of Socrates. As for why Plato replaces Thrasy
naphus with Glaucon and Adeimantus, we can only speculate. One possible answer is that Plato is no longer sanguine that it will be possible to get an interlocutor who is as vicious as Thrasy
naphus to concede the premises that are necessary for achieving adequate insight into the ethical issues that are under discussion. This may be why he chooses to replace Thrasy
naphus with interlocutors who are playing devil’s advocate, rather than defending their own view.

As we have seen in the previous section, Thrasy
naphus’ conclusions are revisionary, but only in the same way as Plato’s own conclusions are. Plato and Thrasy
naphus both attempt to find a deep account of the underlying nature of justice, even if this account ends up diverging sharply from what ‘most people’ accept. We have already explored Thrasy
naphus’ method for finding this deep account of the nature of justice. Plato’s method is more intricate—involving an elaborate comparison between the just individual and the ideal community, along with complex arguments for identifying the feature of the ideal community that makes it just, and similarly complex arguments about the structure of the individual person’s soul. Still, at a very general and abstract level, Plato’s own method is broadly similar to that of Thrasy
naphus.

29 I suspect that this is particularly true of the two arguments that seek to show that justice is a virtue—the pleonexia argument (349b1–350d4) and the civic-and-psy
cic unity argument (351c–352a). These arguments have been correctly argued to be absolutely essential to Socrates’ refutation of Thrasy
naphus in Book I by D. Scott, ‘Aristotle and Thrasy
Even though Plato’s and Thrasymachus’ methods are similar, and Plato thinks that the only viable alternative to his own view is the essentially Thrasymachean view that forms the basis of Glaucon’s challenge, Plato’s also makes it clear that Thrasymachus’ views are as different from his own as they could possibly be. Thrasymachus’ views are not just incompatible with Plato’s, but diametrically opposed to his. According to Plato, Thrasymachus’ values are the inversion of the truth—like a photographic negative in which what is really dark is depicted as bright, and what is really bright is depicted as dark. In Plato’s view, what is in fact the very worst possible life—the life of the tyrant—is praised by Thrasymachus as the most blessed, while what is in fact the best life—the life of the perfectly just person—is despised by Thrasymachus as the most wretched and contemptible.

As I have argued, Thrasymachus’ views represent the values that form the basis of Glaucon’s challenge. As I shall argue in what follows, however, they represent these values in a particularly stark and extreme form. As I have argued in Section 1 above, Thrasymachus’ view of justice flows from his view of human interests—and in particular from his vision of human interests as fundamentally antagonistic and opposed to each other. In this way, Thrasymachus believes that every rational human being is fundamentally an adversary of everyone else. Your overriding interest is in winning and in forcing others to be losers. Specifically, your overriding interest is in dominating others, while their overriding interest is in direct conflict with this—namely, in dominating you, along with everyone else.

This view of human interests, as fundamentally antagonistic to each other, stands in an obvious contrast with the vision that is developed later in the Republic, of the interests of the citizens of the ideal community as perfectly coinciding and harmonizing with each other. One of the most important features of the ideal state, as Socrates describes it (420b–421c), is that it
reveals that the social arrangements that are best for any member of the community are also the ones that are best for all its members. In this way, Plato is pointing to the ultimate ‘win–win’ outcome, the outcome that is the best for everyone, and so the most triumphant possible escape from the grim ‘zero-sum’ game into which Thrasymachus views human beings as permanently locked.

Indeed, the idea that such ‘win–win’ outcomes are possible is stressed very early on in the rest of the Republic. Even Glaucon and Adeimantus argue that each person is better off if everyone treats everyone else justly than they would be if everyone treated each other unjustly (358e–359a). Admittedly, Glaucon and Adeimantus maintain, contrary to what Plato believes, that the society in which everyone treats each other justly is only the second-best outcome for each person, while the very best outcome for each person is the situation in which he unjustly takes advantage of others, while no one else unjustly takes advantage of him. Still, Glaucon and Adeimantus agree with Plato in thinking that this outcome is better for everyone than the situation in which everyone treats everyone else unjustly.

The same idea about the availability of ‘win–win’ outcomes also lies behind the basic economic point which Plato makes even in discussing the ‘cities of pigs’ (369c–d), that a social division of productive labour, combined with a system of exchange, makes every single person better off than they would be without such a division of labour. Even this basic economic point reveals that human interests are far from being as antagonistic as Thrasymachus supposes.

I suggested in Section 1 above that the reason why Thrasymachus thinks that human interests are so antagonistic is that he believes that the dominant component of success in life is power over others. As I shall now argue, this belief reflects his desire to dominate and subordinate others. As we learn in Book IX, the tyrannical soul is full of beliefs that act as ‘bodyguard’ to its
desires (574d6). In a similar way, I suggest, Thrasymachus’ view of the greatest good in human life—domination over others—is the ‘bodyguard’ of his desire for domination; and this desire would have been regarded by Plato as an insane ‘lawless’ desire—a desire of the kind that characterizes the wretched tyrannical soul—although Plato also holds that these wild lawless desires exist, at least in an embryonic form, in all of us. As Socrates says, ‘there is a terrible, wild, and lawless form of desire in each of us, even in those of us who seem to be entirely moderate; this is made clear in our dreams’ (ὡς ἂρα δεινόν τι καὶ ἄγριον καὶ ἄνομον ἐπιθυμιῶν εἰδος ἐκάστῳ ἑνεστί, καὶ πάνυ δοκοῦσιν ἡμῶν ἐνίος μετρίος εἶναι: τούτο δὲ ἂρα ἐν τοῖς ὑπνοις γίνεται ἐνδηλον, 572b3–4).

In their discussion at the beginning of Book IX, Socrates and Adeimantus take pains to emphasize how multifarious and chaotic the desires of the tyrannical soul are. Still, it seems that the desire to enslave and dominate others is one of the desires that will be particularly fierce in a soul of this kind. As Socrates says, if the tyrant has difficulty in subduing his city, he will ‘chastise his fatherland, if he can, by bringing in new friends and making his fatherland … their slaves and keeping them that way, for this is surely the end at which such a man’s desires are directed’ (575d3–e1). It seems to be because the tyrannical soul is so devoted to enslaving and dominating others that it is incapable of relating to others as a friend—instead, as Socrates says, ‘those with a tyrannical nature live their whole life without ever being friends with anyone, being

---

30 Why are there these terrible desires in our souls? Why is there a propensity to evil in human souls at all? The reason seems to be that our lack of knowledge of the good makes us blind to certain forms of goodness, and susceptible to certain illusions in which things seem better than they are, or things that are not good at all seem good to us. For an exploration of this issue, see H. Cherniss, ‘The Sources of Evil According to Plato’, Proceedings of the American Philosophical Society 98 (1954), 23–30, at 26f. The lawless desires are just the worst of these illusions about the good.
always a master to one person or a slave to another, and never getting a taste of freedom and true friendship’ (576a4–6).

Thrasymachus explicitly asserts that the true rulers—who are in his view the most skilful and sagacious individuals that there are—desire power over others. At 345e2, Socrates asks:

σὺ δὲ τοὺς ἀρχοντας ἐν ταῖς πόλεσιν, τοὺς ώς ἄληθῶς ἀρχοντας, ἐκόντας οἴει ἄρχειν;

Do you think that those who rule cities, the true rulers, rule willingly?

To this, Thrasymachus replies:

μὰ Δί᾽ οὐκ, ἔφη, ἄλλ᾽ εὖ οἶδα.

By god, I don’t think it, I know it (342e3).

In a sharp contrast with this, Plato maintains that the rulers of the ideal community will be reluctant to rule. This point emerges particularly clearly in the passage where Socrates describes what they will say to the philosopher-kings to persuade them to go back down into the cave, to rule those who have not been enlightened as they have (520c6–d3):

καὶ οὕτω ὑπαρ ἡμῖν καὶ ὑμῖν ἡ πόλις οἰκήσεται ἄλλ᾽ οὐκ ὄναρ, ώς νῦν αἱ πολλαὶ ὑπὸ σκιαμαχοῦντων τε πρὸς ἄλληλους καὶ στασιαζόντων περὶ τοῦ ἄρχειν οἰκοῦνται, ώς μεγάλου τινὸς ἄγαθοῦ ὄντος. τὸ δὲ ποὺ ἄληθὲς ὃδ᾽ ἔχει: ἐν πόλει ἢ ἡκιστα πρόθυμοι ἄρχειν οἱ μέλλοντες ἄρξειν, ταῦτην ἄριστα καὶ ἀστασιαστότατα ἀνάγκη οἰκεῖσθαι, τὴν δ᾽ ἐναντίους ἄρχοντας σχοῦσαν ἐναντίως.

So the city will be awake, governed for us and for you, rather than dreaming like the majority of cities nowadays, which are governed by people who battle each other over shadows and fight civil wars in order to rule—as if that were a great good. No, the truth is surely this: a
city in which those who are going to rule are least eager to rule is necessarily best and freest from civil strife, whereas a city with the opposite kind of rulers is governed in the opposite way.

The reference to ‘shadows’ obviously refers to the parable of the cave earlier in Book VII. The contrast between being ‘awake’ and ‘dreaming’ refers back to 476c5–6, where ‘dreaming’ is defined as ‘thinking that what is similar to something is not similar but the same as what it is like’—which also seems a perfect description of the deluded souls at the bottom of the cave.\(^{31}\) In this way, the desire to have power over others, purely for its own sake, seems characteristic of the deluded minds at the bottom of the cave.

If any soul best typifies the delusions of the bottom of the cave, it is surely the wretched tyrannical soul. In an echo of these descriptions of the cave, the lawless desires that characterize the tyrannical soul are described as most evident in ‘dreams’ \((572b3–4)\). This seems to provide some support for my suggestion that the desires that characterize Thrasymachus—the desire to dominate others, purely for its own sake—would indeed be viewed by Plato as one of the wild lawless desires that are allowed free rein in the tyrannical soul. The shockingly rude, aggressive, bullying tone of Thrasymachus’ contributions to the dialogue also seems to mirror the untrammelled violence that, according to Socrates \((574e5)\), the tyrant is happy to resort to.

In this way, then, Thrasymachus does not just praise the life of the tyrant. He seems to desire to lead the tyrant’s life himself; and this desire is a lawless desire, of the sort that typifies the wretched tyrannical soul. In this respect, Thrasymachus himself foreshadows the figure of the

tyrant, who reappears at the beginning of Book IX.32 Indeed, while Thrasymachus foreshadows the worst of the four defective types of soul that are discussed in Books VIII and IX—the tyrannical soul—it is also possible to see the other interlocutors of the Republic as comparable to the three other defective types of soul: Cephalus the money-maker can be compared to the oligarch; the oligarch’s son Polemarchus, who resembles a zealously partisan politician, can be compared to the democrat; and Glaucon himself is described as similar to the ‘timocratic’ or ‘Laconian’ soul by Adeimantus (548d7).

After Book I, Thrasymachus is mentioned again four times in the Republic. Three of these references come at structurally crucial points in the overall argument. We have already discussed the crucial allusion to Thrasymachus at the beginning of Book VIII, where the main line of argument is resumed after the digression of Books V–VII; this allusion seems designed to remind us that the central argument of the Republic is a response to his praise of injustice.

Two of the other mentions of Thrasymachus also play a similar structural role.33 First, at the beginning of Book V (450a5–b5), we are reminded that he is still present, and (somewhat

32 R. Barney, ‘Callicles and Thrasymachus’, in E.N. Zalta (ed.), The Stanford Encyclopedia of Philosophy (Winter 2011 Edition), §2, writes: ‘Thrasymachus is depicted as torn between the characteristic drives of the two lower parts of the soul identified in Book IV of the Republic: the appetitive part [epithumêtikon], which lusts after money, and the spirited part [thumos], which loves competition and victory’. I cannot argue for the point at length here, but I believe that this is a misinterpretation of the spirited part of the soul: the spirited part does indeed ‘love competition and victory’—but it loves them for the sake of glory and renown. Thrasymachus seeks to win in debate, and in life, in order to dominate others for its own sake—a desire which in my view belongs, not to the spirited part, but to the wild lawless desires of the appetitive part.

33 He is also mentioned in Book VI, in the middle of the discussion of the ideal philosopher, where it is suggested—perhaps in jest—that Socrates and Thrasymachus have now become ‘friends’, and that Socrates won’t
surprisingly) still taking part in the conversation. This too reminds us that even though Books V-VII are officially a digression from the main line of argument, the focus on responding to Thrasymachus has not been forgotten. Then, at the end of Book IX (588a–592b), as the central argument of the Republic is being completed, Socrates gives an image of what the soul of the unjust person is like. The three parts of the soul are compared to a human being, a lion, and a hundred-headed monster, in order to underscore the point that it is in everyone’s interest for the rational part of the soul to rule the two irrational parts of the soul. In the course of this discussion, Socrates makes the following statement about those whose rational part is incapable of ruling their spirited and appetitive parts (590c7–d6):

οὐκοῦν ἵνα καὶ ὁ τοιοῦτος ὑπὸ ὀμοίου ἄρχηται οἴονπερ ὁ βέλτιστος, δούλον αὐτὸν φαμεν δεῖν εἶναι ἐκεῖνοι τοῦ βελτίστου, ἔχοντος ἐν αὐτῷ τὸ θείον ἄρχον, οὐκ ἐπὶ βλάβη τῇ τοῦ δούλου οἰόμενοι δεῖν ἄρχεσθαι αὐτῶν, ὡσπερ Ὄρασιμαχος ἑκτὸ τοὺς ἄρχομένους, ἀλλ᾽ ἡς ἀμείνων ἐν παντὶ ὑπὸ θείου καὶ φρονίμου ἄρχεσθαι, μάλιστα μὲν οἰκεῖον ἔχοντος ἐν αὐτῷ, εἰ δὲ μὴ, ἐξωθεν ἐφεστῶτος, ἵνα εἰς δύναμιν πάντες ὀμοίων ὁμεν καὶ φίλοι, τῷ αὐτῷ κυβερνώμενοι;

So it is in order to ensure that someone like that is also ruled by something similar to what rules the best person that we say that he should be the slave of that best person who has the divine ruler within himself. It is not to harm the slave that we think he should be ruled, as Thrasymachus thought was true of all subjects, but because it is better for everyone to be ruled by a divine and wise ruler—preferably one that is his own, which he has inside himself;

‘relax his efforts’ to ‘convince him and the others, or at any rate, do something that may benefit them in a later incarnation, when they, reborn, they happen upon these arguments again’ (498c6–d4).
but otherwise one imposed on him from outside, so that we may all be as alike and as friendly as possible, because we are all steered by the same thing.

This passage recapitulates the contrast, which was already clear in Book I (341c–342e) between Thrasymachus’ and Socrates’ views of the relationship between ruling and human interests. According to Thrasymachus, ruling over others is in the interest of the ruler, and harmful to those who are ruled; whereas for Socrates, at least in the ideal case, ruling—even ruling as a master rules over a slave—does not serve the interests of the ruler, but rather the interest of those who are ruled.

In conclusion: our examination of the text has, I believe, supported my conclusion that Thrasymachus is presented as a horrifying figure, personifying some of the most terrible dangers that the human soul confronts us with. Part of what makes Thrasymachus so dangerous is precisely the fact that his view of justice, and of human life and human interests, is not simply confused. If his view were simply incoherent, it would not be as pernicious and seductive as it actually is, and Plato would have less reason to argue against it. For Plato, Thrasymachus’ view is utterly wrong-headed from beginning to end; but it is not incoherent. Commentators who depict Thrasymachus as a muddle-headed bruiser are trivializing the awful dangers that Plato is desperately trying to warn us against.34

University of Southern California

34 This paper was originally written in 2011, when it was presented at a meeting of Oxford University’s weekly Ancient Philosophy Workshop, and then not revised until almost six years later. In rewriting the paper, I received invaluable assistance from two anonymous readers, and also from the editor of this journal. The motivation for revising the paper came in part from the political events of 2016–17, which seemed to me to confirm that Plato’s warnings against Thrasymacheanism are as vitally important today as when they were first written.
**BIBLIOGRAPHY**


Kant, I., *Die Religion innerhalb der Grenzen der bloßen Vernunft* (Königsberg, 1793).


Nietzsche, F., *Der Antichrist* (Leipzig, 1894).


