

Journal of the
HISTORY *of*
PHILOSOPHY

AN OFFPRINT FROM

OCTOBER, 1972

Volume X, Number 4

The Happy Philosopher— A Counterexample to Plato's Proof*

SIMON H. ARONSON

THE PAST FEW YEARS have seen a renewed interest in the main argument of Plato's *Republic* concerning the relationship of justice and happiness. Professor Sachs reopened the discussion¹ with his suggestion that in the *Republic* Plato employs (at least) two different conceptions of justice. Sachs went on to charge that Socrates, by failing to connect these two senses of justice, commits the fallacy of irrelevance, for whereas Thrasymachus, Glaucon, et al. have challenged Socrates to prove that the "vulgarly just man" is happy, all that Socrates does in fact show is that the "Platonically just man" is happy. A number of replies² were soon forthcoming, attempting to show either that the alleged failure is more apparent than real, or that Sachs had misconstrued the nature of the original challenge, or that Plato never particularly cared to vindicate the "vulgar morality" of the many. Recently Professor Vlastos has entered the debate³ with his attempts at clarifying Plato's two "definitions" of justice and providing "an argumentative bridge" between them.

In what follows I approach the problem from a somewhat different direction. While in this essay I reach the same conclusion as did Sachs, i.e., that Plato failed to meet the challenge of proving that the just man is happier than the unjust man, my argument neither deals explicitly with Sachs' allegations nor does its validity depend on what position one takes in the Sachs controversy.⁴ Quite

* I am much indebted to members of the University of Chicago Philosophy Department and to a reader for this *Journal*, whose valuable criticisms of an earlier draft of this paper led me to correct certain mistakes and to fill certain gaps in my argument.

¹ "A Fallacy in Plato's *Republic*," *Philosophical Review*, LXXII (April 1963), 141-158.

² See Demos, "A Fallacy in Plato's *Republic*?" *Philosophical Review*, LXXIII (1964), 395-398; Weingartner, "Vulgar Justice and Platonic Justice," *Philosophy and Phenomenological Research*, XXV (1964), 248-252; Schiller, "Just Men and Just Acts in Plato's *Republic*," *Journal of the History of Philosophy*, VI (1968), 1-14; Alan Gewirth's paper "Comments on 'A Fallacy in Plato's *Republic*'" (unpublished).

³ "The Argument in the *Republic* that 'Justice Pays,'" *Journal of Philosophy*, LXV (Nov. 7, 1968), 665-674, hereafter referred to as JP; "Justice and Psychic Harmony in the *Republic*," *Journal of Philosophy*, LXVI (Aug. 21, 1969), 505-521, hereafter referred to as PsH.

⁴ My own view, for which I do not argue here, is that Sachs is correct in his first charge that Socrates fails to successfully demonstrate that all Platonically just men are also vulgarly

apart from whatever problems are posed by any fallacy, lacuna, or oversight existing in Plato's argument, the thesis of this paper is that the *Republic* is open to a more fundamental criticism. I argue that certain elements in Plato's theory, elements essential to his description of the ideal city, tend both to undermine Plato's own argument and in fact to strengthen Thrasymachus' position. Briefly, I want to argue what may at first sound paradoxical, namely that the philosopher-ruler is himself a counterexample to the thesis that maximal happiness is compatible with justice. I submit that in his portrayal of the philosopher-ruler Plato lets slip the startling admission that by committing an injustice the philosopher will be happier than if he did not commit one. I suggest that this dilemma of the philosopher in the ideally just city well illustrates the thrust of Thrasymachus' claim that one must choose between one's own happiness and that of one's fellow citizens.

I begin by summarizing what Plato means by 'justice' in the *Republic*, and in what sense it can be said that there are two different "kinds" or "conceptions" of justice. I then examine the challenge which Thrasymachus, Glaucon, and Adimantus pose, and indicate what Socrates must show to meet this challenge. Finally I argue that in a crucial passage Socrates himself provides a counterexample which undercuts his own case, and which to an extent vindicates Thrasymachus' original claim.

I. JUSTICE

For Plato there is only one idea (or Form) of justice. As a Form, or universal, it will be participated in by many things in the world; thus there will be many different instances of justice. For each of the various instantiations, however, a formal statement of what the justice of that thing consists in would amount to the same *λόγος*, or definition, of justice.⁵

'Justice', as Plato uses it, is a term whose primary application is to groups or communities, to "wholes" which have "parts." Plato generally seems to limit his use of the term so as to apply it mainly to human groups, or to individual humans considered as communities of psychic parts; in addition to speaking of just men and just cities, he mentions the justice and injustice of such associations as a clan, an army, and even a band of robbers or pirates (351c-e).⁶ Socrates makes it clear that, whatever different "whole/part" he is talking about, it is nevertheless the *same* conception (Form) of justice that is said to apply (368d, 435a-e, 442d).

just, see n. 16 *infra*, but that Sachs is wrong in expecting Socrates "to prove that his conception of the just man applies to—is exemplified by—every man who is just according to the vulgar conception" (Sachs, p. 153). Socrates does not take this latter position; some men are "vulgarly just" merely by habit, accident, or right opinion, and they are not guaranteed maximal happiness. Cf. Gewirth, *op. cit.*

⁵ Cf. *Meno* 72-73.

⁶ While Plato does not do so, one could see how he might apply his concept of justice to, say, a family or an empire. By extending the term slightly, we can understand how Plato might even apply 'justice' analogically to non-human "wholes" as well, e.g., to an artificial product such as a watch, or even to the universe taken as a whole. Cf. *Gorg.* 508a.

A whole is said to be just when all of its parts are functioning in a certain way. When Plato specifies what this way is, we see that 'justice' is for him an avowedly normative term. For justice is said to obtain when each part is "minding its own business," when each part performs that function in the whole for which its nature best suits it (433ab). It is in this condition that a whole achieves its best interests, by all the parts cooperating together in the "common enterprise" of the whole (351c). A whole accomplishes whatever it accomplishes because of the justice that obtains among its parts (351c, 351e, 352c). When justice obtains among the parts, the whole achieves its "good";⁷ conversely, when some whole is in fact good, justice (as well as the other virtues) will be present in it. Socrates makes this point explicit when he argues that if the city is perfectly good (τελέως ἀγαθὴν,) then it is plain that it is wise, courageous, moderate, and just (427e). Socrates avows that this connection between a thing's being just and its attaining its own good was assumed from the start: "we founded one [a city] as best we could, knowing full well that justice would be in a good one" (434de).⁸

I shall not undertake in this paper to analyze Plato's further identification of "good" and "happy."⁹ It is clear that in the *Republic* the request to show that justice is a good in itself (366e) is taken to be equivalent to the request to show that the just life is happier than the unjust life (344a, 352d). In crucial passages Socrates uses the terms interchangeably: at one moment he speaks of having founded the city to be as good as possible (427e, 434e); at the next he unhesitatingly recalls having founded the city to be as happy as possible (420b-e, 466a). The judgment in Book IX Socrates calls a consideration of "the good life and the bad one" (578c), and then immediately goes on to decide this question in terms of happiness and wretchedness. For the purposes of this paper I shall

⁷ 'Justice' is also variously identified in the *Republic* with what is 'beneficial', 'profitable', 'gainful', 'advantageous'. Cf. *Rep.* 336d and 337bc; 339b; 348c.

⁸ As Robinson points out, *Plato's Earlier Dialectic*, 2nd edition, pp. 211-212, this assumption begs the important question of whether justice is a good. This, however, is not as damaging as might at first seem, for Plato goes on to check his definition against common usage (433b-e; 442e-443a).

⁹ Cf. Sachs *op. cit.* p. 142 n. 5; Vlastos JP, p. 665, "'good' is an ellipsis for 'good for the just man himself', i.e., a contribution to his well-being or happiness (*eudaimonia*)"; Gulley's discussion of Socrates' psychological *eudaimonism*, *The Philosophy of Socrates* (London: Macmillan, 1968), esp. pp. 77, 87; Versenyi, *Socratic Humanism* (New Haven: Yale, 1963), pp. 79-82; *Symposium* 204-205 where happiness is said to consist in the possession of the good.

I am uncertain whether Plato wants (1) to simply identify "happiness" with "goodness," or (2) to say that "happiness" is something which is other than "goodness" but which invariably accompanies it. If the latter, then it is incumbent upon Plato to explain this connection.

For Plato, "goodness" or "being a good X" is often rather far removed from our sense of "morally good." At times in the dialogues it is almost equivalent to a functional or instrumental sense of 'good', i.e., being a good specimen of that kind of thing, performing well the function that this sort of thing is supposed to do—hence the Platonic notion that goodness lies in realizing what one is naturally suited to do, and that an excellence is logically connected to "being good," i.e., performing well. Plato regularly uses εὖ πράττειν to effect the transition from "doing well" (i.e., functioning well) to "faring well" (i.e., living well or happily) (*Rep.* 353e). Plato is frequently faulted for equivocating at this point, but if "goodness" means "functionally good," and if happiness is identified with being good, Plato is saved from much of this criticism. Cf. Thayer, "Plato: the Theory and Language of Function," *Philosophical Quarterly*, XIV, 57 (Oct. 1964), 303-318.

assume this interchangeability between "good" and "happy." Thus the logical connection explained above between a whole's being just and being good will also hold between a whole's being just and being happy; Socrates says as much at 420b: "In founding the city we are not looking to the exceptional *happiness* of any one group among us, but, as far as possible, that of *the city as a whole*. *We supposed we would find justice most in such a city*" (my italics). Thus, on Plato's use of 'justice', it is necessarily true that a whole which is just is happier than a similar whole whose parts are not working in such a just manner.

It is precisely this fact, that for Plato a just Φ (where a Φ is a "whole" as discussed above) cannot but be happier than an unjust Φ , that explains why the main argument of the *Republic* takes the course it does. For Plato, one can doubt whether a just Φ is happier only if one does not see clearly what being just consists in; if one saw "what justice is," there would be no question (354bc). This is simply another instance of Socrates' frequent insistence on the priority of the "what is X" question; once one knows the nature of X, one will know or "see most clearly" (445b7) whether X is good or bad, teachable or not, shameful or noble, etc.¹⁰ Thus the main argument of the *Republic*, which occupies Books II-IV and resumes in VIII and IX, consists almost entirely in laying out the nature of justice/injustice—or more precisely, the justice/injustice of cities and men. That is why, once these "natures" are laid out, there is no need for further independent argument as to which is happier; Glaucon and Socrates agree that further inquiry would be "absurd" (445ab) since "these qualities have manifested their characters in our description." The only reason that Socrates delays this judgment at the end of Book IV is that, having attained a good vantage point, he feels he might as well make good use of it (445b)—there are various forms of vice and degeneration that are "worth looking at." But once this "laying out" of the degenerate natures is completed (576b), the final judgment is quickly concluded; it consists in simply spelling out what we see when we view the city or man "as a whole" and look into its internal workings.¹¹ The answer, Glaucon states, is "clear to everyone."

If this is so, then what is the problem Thrasymachus poses? Are his claims that being just is foolish and that the unjust man lives a better (\equiv happier) life (344a) simply based on ignorance or semantic confusion? And in what sense, if any, does his challenge involve relating or connecting *two* "senses" of justice?

II. JUST MEN AND JUST CITIES

Any "whole" is just, and thus good and happy, if it instantiates the one conception or Form of justice discussed above. In the *Republic* Plato considers two "wholes," a man and a city, each of which is potentially happy or miserable. In each the just condition, i.e., the proper functioning and ordering of its

¹⁰ Cf. *Meno* 71ab; *Gorgias* 448e, 451d, 463d.

¹¹ 576d-580c. The two other proofs which follow are clearly intended to be appendices to the judgment of the happier life, the result of which is announced at 580b. Plato then tries to show that such a just (\equiv happy) life is also the most pleasant. He explicitly distinguishes (581e8) this latter decision concerning pleasure from the question of the good/bad life. Cf. Murphy, *The Interpretation of Plato's Republic*, pp. 59-60.

“parts,” is the condition in which that whole is most happy. In the individual man the “parts” are the various elements in his soul; in the city the “parts” are men, or more precisely, classes of men. Thus a man may be viewed from either of two perspectives, i.e., with respect to two distinct “justices”: he is a “whole” with respect to the justice of a man, while he is a “part” with respect to the justice of his city. And it is only in the former case that the justice involved entails that the man is maximally happy.

It is important to realize that Socrates keeps these two “justices,” that of the man and of the city, quite separate as he develops his argument. At the start of his inquiry (368d-369a) he sets forth his intended procedure: he will use a larger “whole” to provide a more perspicuous view of “what justice is” and then apply this knowledge to his consideration of a smaller “whole”; thus he first examines the justice of cities, and then that of individual men. From 369 to 427 Socrates, with some help and prodding from Glaucon and Adimantus, “founds” or describes a city that is “perfectly good” (427e). They then search for the various virtues in it, arriving finally at justice. The justice of the city consists in the very rule which governed the founding of the city, namely, “each one must practice one of the functions in the city, that one for which his nature made him most naturally fit” (433a6). This is the definition, or *λόγος*, of *the justice of the city*. For the city as a whole to be as happy as its natural limitations allow will necessitate that men (parts) act in certain ways, e.g., some will rule; some will fight and act as police; others will sow crops and mend shoes. What is essential to note is that the justice of the city demands of each man not that he do whatever he is best suited by nature to do *simpliciter*, but that he do that which he is best suited by nature to do *of the functions in the city*.¹² In the course of establishing this just city Socrates states explicitly that his concern is to maximize not the happiness of its individual members, but that of the city as a whole (420b ff.).

Having completed his investigation of what constitutes the justice of a city (434d), Socrates switches gears and considers the nature of man. At 436-441b he finds that a man, like a city, can be considered to have three “parts,” i.e., three parts of his soul. At 441c-e he sums up the conditions of the letter analogy: when he applies the newly discovered formal definition of justice to this new whole (man) with a new set of parts (divisions of the soul), he will be able to state in what the justice of a man consists.

It is at this point that I think Professor Vlastos errs in interpreting Plato’s argument. He evidently forgets Plato’s account of the letter analogy, wherein Plato explicitly states that we are using the just city as a way of seeing what justice *writ large* is—we are discovering the “formal definition” of justice. We are then going to apply this definition to man, and then check our results against common usage. The argument at 441c-e is *not* an attempt to show that the justice

¹² This is necessary for at least two reasons. First, not all functions are performed in a just city, e.g., imitative poetry. Second, as Plato is fond of pointing out, one who is best suited to perform a particular task is usually also well-suited to perform its opposite, e.g., a good guard also makes a good thief, cf. 333e-334a. But a just city will require one to do only that task which promotes the well-being of the city.

of the city and the justice of a man "must always be satisfied together."¹³ Rather, at 441c Plato is spelling out what must be the case if the letter analogy is to apply. It is at what Vlastos (in PsH, p. 517) calls step (E) that he mistakes what Plato is doing. The point is worth considering in detail. Socrates says at 441d3:

And we shall say, O Glaucon, that a man is just in the same way in which a city is just.

It seems plain that what Plato is saying is that a city and a man are both just in the same way, i.e., *by the presence of the same Form*: the justice of the man and the justice of the city will each consist in the parts of that whole performing the function for which they are best suited for the good of that whole. Plato confirms that this is what he means by proceeding to apply what he says, concluding two steps later:

Therefore, let us bear in mind that also in the case of each one of us, whosoever is such that each of the three kinds [of elements] in him does its own, he is a just man *and a man who does his own*. (441d12-e2; Vlastos' "step G," p. 517)

Vlastos, partly because he is already eagerly looking to find some argument, and partly because he assumes that the terminal phrase in G 'a man who does his own' must mean "does his own job in the city,"¹⁴ decides that the crucial argu-

¹³ Vlastos' claim, PsH p. 515. Vlastos wants to find an argument that links social justice with individual justice, because he is (justifiably) concerned that otherwise Socrates might indeed be "guilty as charged" of committing Sachs' fallacy of irrelevance. I am not sure what is gained by saving Socrates from that fallacy at the price of having him commit (twice) the fallacy of equivocation, but I do not think the attribution of either fallacy here is warranted or necessary. Socrates' attempt to meet Glaucon's (and Sachs') requirement of showing that the Platonically (internally) just man acts in accordance with vulgar, i.e., social, standards occurs not here but one page later, at 442d10-443b2. Note that there Socrates specifically mentions many of the very acts which characterized the standards he was supposed to meet in the original challenges (343d-344b, 360a-c, 362b). It is true that the argument there amounts to little more than an appeal to intuition, and offers merely Glaucon's affirmation of Socrates' asserted conclusions. But this is entirely in accordance with regular Socratic procedure: the long and hard task is in trying to decipher what the nature of a particular concept, e.g., the justice of a man, *is*; once it is known, its "effects" (supposedly) follow as a matter of course. But, regardless of whether one accepts the conclusions of the argument at 442d-443b (I don't), the important point is that Socrates *does try* to make the crucial argument linking internal justice with social standards *there*: he says explicitly, "Isn't the cause of all this [behavior] that so far as ruling and being ruled are concerned, each of the parts in him minds his business?" (443b1-2). The value of knowing that Socrates attempts the argument there is that it saves us from feeling compelled to find it where it really is not attempted.

¹⁴ It is true that in the earlier discussions of the justice of the city the phrase did mean this. Here, in switching to a discussion of the justice of a man it seems natural to interpret "does his own" as referring to the man's own "true" nature, i.e., his soul. Socrates, explaining the definition at 443c-e, carefully distinguishes the way he is using 'own' here from any social sense:

In truth justice was, as it seems, something of this sort; however, *not with respect to a man's minding his external business, but with respect to what is within, with respect to what truly concerns him and his own*. He does not allow the principles within him to do other work than their own, nor the classes in his soul to interfere with another, but really sets his own house in good order and rules himself. (443c-e; my italics).

ment linking social justice with individual justice is to be found in the above step E. He finds the argument (a fallacious one) by determining that Socrates equivocates on the two uses of 'just' in E. He states (p. 518), "In its first occurrence [in the line 441d3 quoted above] we would expect 'just' to carry its natural sense when applied to an individual. So it would have to mean 'just₁' here unless some warning to the contrary had been given, as none is."¹⁵ But Vlastos is prejudging the case—what we should be expecting here is precisely *not* its "natural sense," but its newly found formal definition. A rather clear warning has in fact been given: at 434e Socrates announces his intended procedure:

Let us apply what came to light there [i.e., in the city] to a single man, and if the two are in agreement, everything is fine. But if something different should turn up in a single man, we'll go back again to the city and test it.

That is, Socrates explicitly allows for the possibility that the Form of justice which he is applying to man will *not* accord with the "natural sense" of what we do call "just men"; if that happens, we shall have to return to the city and look harder. If Vlastos were correct in thinking that Plato was initially introducing the "natural sense" of 'just' here, then what would be the point of "testing our justice in the light of the vulgar standards" (442e-443a)? On Vlastos' interpretation this would be an empty, tautologous procedure. On the interpretation I am offering, Socrates straightforwardly examines one half of the analogy to find out what justice is, applies the result to the other half, and then checks his application to make sure it accords with ordinary usage.¹⁶ Vlastos is driven, on his interpretation, to attribute to Plato an argument riddled with equivocation; then, not surprisingly, Vlastos concludes that "the deduction is worthless; the demonstration has collapsed" (p. 518). But the plain fact is that Plato never attempted it.¹⁷

¹⁵ For Vlastos, "justice is the social virtue par excellence" (p. 520), and thus 'just' "in its primary sense must be a relational predicate" (p. 519). He thus explains 'just₁': "in its primary signification 'just' is a relational predicate; to speak of a person as having this property is to think of the way in which he habitually relates himself to persons or groups of persons in his conduct. This is all too plainly true on the popular notion of justice" (p. 517).

This may be true of common notions—it may even be true of any "correct" notion of justice—but it is inverted Plato. For Plato, the vulgar use of language is precisely what is *not primary*—the ordinary man's use takes the "image," the "phantom," the "shadow," the "external consequence" as though it were "the real meaning" (*Rep.* 382bc, 422e, 493bc, 500a, 515b, 597a). For Plato, justice is *primarily* a property of a whole, and only derivatively can be said to belong to a part. (See my discussion of J_p below.) On Plato's way of thinking, Vlastos (and ordinary language) is guilty of making the part prior to the whole, the external prior to the internal. (Plato, as it were, inverts Wittgenstein 180 degrees; one might almost say that Plato believes an outward process stands in need of inner criteria [cf. *Philosophical Investigations*, par. 580].) I am in no sense defending Plato's notion of how language operates, but failure to recognize these features both produces distortions in Plato's arguments where they do not occur and inhibits recognizing where Plato does in fact err.

¹⁶ My remarks here should not be construed as implying any agreement with Socrates' conclusion that this new "true" usage does in fact meet the "vulgar tests" to which he puts it. Indeed, I am dubious as to whether Socrates could easily defend his rather cavalier conclusions at 442e-443a, that the man who is thus internally just, i.e., instantiates the Form of justice among his psychic parts, would necessarily refrain from performing the acts indicated.

¹⁷ The fact is that when Socrates arrives at step E in the argument we have only been exposed to *one* meaning of 'just', i.e., a ϕ is just if its parts are so arranged, etc., and it seems

Let us return to Socrates' consideration of the justice of a man. As stated, Socrates applies the previously discovered Form of justice to man, and concludes:

. . . in truth justice was, as it seems, something of this sort; however, not with respect to a man's minding his external business, but with respect to what is within, with respect to what truly concerns him and his own. He does not allow the principles within him to do other work than their own, nor the classes in his soul to interfere with another, but really sets his own house in good order and rules himself. (443c-e)

Let us refer to such a man, i.e., he who considered *as a whole* instantiates justice among his psychic parts, as J_w . Concerning the actions of one who is J_w Plato states that

he acts, if he does act in some way—either concerning the acquisition of money, or the care of the body, or something political, or concerning private contracts. *In all these actions he believes and names a just and fine action one that preserves and helps produce this condition*, and wisdom the knowledge that supervises this action; while he believes and names an unjust action one that undoes this condition. (443e; my italics)

Let us call such acts J_wA . It will then be true that the J_w who continually performs J_wA (and refrains from performing U_wA , i.e., unjust actions as defined above) will be as happy as the limitations of his nature allow.¹⁸

But because a man can be viewed with respect to either of two distinct justices (or, more correctly, two distinct instantiations of justice) we find that we can now talk of a man's being just in yet a second sense. For with regard to the justice of the city we can call a man "just" in the derivative sense that he performs those acts which are required of him if the city as a whole is to be just; in this new secondary usage we say that a man is just *as a part* if and only if he practices that function which the definition of the city's justice (433a6 above) necessitates. Let us refer to a man who is just in this sense, i.e., as a part, as J_p , and let us call his typical acts J_pA . Depending on the nature of the particular J_p , such acts would include ruling and making political decisions, fighting in the army and some minor

almost a non sequitur to pull in another sense of 'just' out of the blue. It seems much more plausible to attribute to Socrates the good sense of doing what he has said he would do all along. Vlastos has to clean up too many loose ends on his interpretation (cf. n. 44); he also has to attribute to Socrates a second commission of the equivocation (in G) (n. 45), in order to make sense of the fact that the argument thereafter is plainly about an internally just man. I cannot share Vlastos' surprise (n. 12 and n. 24) that other commentators have "missed" finding the argument.

¹⁸ Since men have different natures (370a, 415a) this will not mean that all men, if they are as happy as possible, will end up engaging in the same sorts of activities. It is not true, as Weingartner claims it is, that "Plato urges us all to become philosophers" (*op. cit.*, p. 252). In urging us each to become internally just (592b), Plato is telling us we will be as happy as possible when we realize that for which our natures best suit us. If I am most naturally equipped, say, to work with my hands, while my natural capacities for perceiving what *is* are poor, the last thing Plato would urge is that I try to become a philosopher. Indeed, Plato would argue, I imagine, that I will be happier being a top-notch craftsman than by being a bungling philosopher.

police work, or engaging in various economic pursuits. From this it follows that if all the men in the city are J_p and continue to perform J_pA , then the city will be just and as good and as happy as possible.

With this framework in mind, we see that there are two different criteria by which to judge whether a particular person or act is just: a man may be either J_w or J_p (or both); and an act may be either J_wA or J_pA (or both). But as the above explication lays bare, a man's being as happy as possible is logically connected only to the former of each pair. The justice of the city, and thus a man's being J_p and performing J_pA , seems *prima facie* to be compatible with a man's attaining less than his maximum possible happiness; and a man's being J_w and performing J_wA (and thus being maximally happy) seems compatible with his performing U_pA and thus being U_p , i.e., being an unjust part vis-à-vis his city. And it is here that Thrasymachus must be heard.

III. THE CHALLENGE

When Thrasymachus raises his initial objections, one might fairly question whether he in fact distinguishes between two kinds of justice. Certain passages make it plausible to attribute to him a distinction between "justices" which differ at least in degree, if not in kind (338cd; 344a; 344c; 348d); of these he makes it clear that he is talking about the "grander," larger scale" one. But regardless of whether we say that he only mentions one kind of justice, or that he mentions two and then chooses to speak of the "larger scale" one, it is clear that the justice of which he does speak is the justice in a city. From the outset he dismisses Socrates' concentration on a particular man (338d3), and focuses the discussion on cities, law, and rulers (338de). He talks of justice in democracies, tyrannies, and other regimes as a device used to promote the advantage of the ruling class (339a); injustice is thought of in terms of not petty crimes, but political coups (344b). All of his examples deal with political office, and he considers whether such offices should be subverted from promoting the good of the ruled to promoting the advantage of those who hold power.

What is equally clear is that at this point Thrasymachus in no sense thinks of justice as being *within* one individual. It seems quite doubtful that it has ever occurred to him that justice might be thought of as internal to one man; Socrates does not suggest this possibility until 351e. Thrasymachus' claim that the unjust life is better (343-344) is stated in terms of committing injustice in the city; at no time does Thrasymachus ever suggest that a man, all of whose parts are functioning well and working harmoniously, will be less happy than a man not so ordered.

What Thrasymachus is suggesting seems rather straightforward and unambiguous. As he sees it, justice in the city consists in performing actions which promote *not* the good of the agent, but someone else's good (343c).¹⁹ Since

¹⁹ Kerferd has argued persuasively that Thrasymachus maintains one consistent position throughout Book I, namely, that 'justice' means "(the promotion of) another's good." Thus, from the point of view of the ruled, justice would be (the promotion of) the advantage of the ruler. Since Thrasymachus operates on the assumption that everyone is out for himself,

Thrasymachus is only interested in his own good, he sees no reason to perform such actions; in fact, he sees them as detrimental to his own good. He challenges Socrates to show him that one who performs such actions will be happier than one who refrains from performing them. He wants to be shown how being just vis-à-vis one's city "pays." Or, to state the problem in terms of Socrates' analysis of the justice of a city and of a man, Thrasymachus wants to be shown that $J_p A$ do not conflict with $J_w A$. He wants to know how being J_p will promote—or at least be compatible with—being J_w .²⁰ An indication that Thrasymachus' point may have some basis is shown by the example, provided by Socrates himself, of the just robber band (351-352). Perhaps unwittingly, Plato offers here a case of a "whole" which has "a kind of justice" within it, and yet still is an unjust "part" of a larger society.

If we step back a moment, and view the problem in more general (and somewhat less precise) terms, we could say that what Thrasymachus is asserting is that interests, even enlightened "true" or "real" interests, of the various "wholes" in the world conflict. He is suggesting that one must choose between being either a happy "whole" or a not-so-happy "part" of some larger whole; that no state of affairs can exist which satisfies simultaneously the "goods" of all "wholes." He is challenging the Platonic belief that all "true" interests are compatible, or to put it in more Platonic language, that they all "form a harmony."²¹

In this paper I do not deal directly with the important question of whether in the *Republic* Plato attempts to meet this challenge. It may be that Plato simply chooses to concentrate his attention on the problem of the "good life" for man²² and feels that Thrasymachus' concern with J_p is only a peripheral problem, since it deals with what is at most an "image" (εἰδωλόν 443c3) of a man's "true" justice. Or, more plausibly, perhaps Plato believes that although he discovers and discusses the two justices separately, his description will reveal to what extent the justices of the city and of man not only can coexist but in fact are complementary;²³ it seems clear that Plato wants to maintain that, for the vast majority of men, if they realize their own "true" capacities (J_w), they will in fact

he thinks it foolish to do what is just. It is on this basis that Thrasymachus argues that unjust action—by which he means action which aims at the good of the agent—will make a man happier. See "The Doctrine of Thrasymachus in Plato's *Republic*," *Durham University Journal*, n.s. IV (1947-1948), 19-27; also Kerferd's "Thrasymachus and Justice: a Reply," *Phronesis*, IX (1964), 12-16.

²⁰ Even this restatement of the problem in Socratic terms is not entirely fair to Thrasymachus, for it smuggles in a rather important shift in "whose good" the justice of the city promotes. Whereas Thrasymachus sees a conflict between one "part" versus another "part," e.g., ruler vs. ruled, on Socrates' analysis the problem becomes "part" vs. "whole." It might be a significantly easier task to prove that a part benefits when its "whole" does, than to show that one "part" benefits when some other "part" does. But for the purpose of this paper, I disregard this shift; let us assume that Thrasymachus accepts Socrates' description of what the city's justice involves, and wants to know why some "part" should promote the good of the city as a whole, when such action seems to be at the expense (420b ff.) of the maximum happiness of that "part."

²¹ *Gorgias* 508a.

²² Cf. 344e, 352d, 578c, 618c.

²³ Cf. 435e, 442a, 590de.

also be performing their "proper" (J_p) social functions. For such men, the choice between one's own good and the city's thus presents a false dichotomy.

IV. THE HAPPY PHILOSOPHER—A COUNTEREXAMPLE

In what follows I argue that in at least one crucial instance, namely in the case of the philosopher who has been nurtured and trained in the ideal city, Plato baldly admits that this man's J_wA are incompatible with his J_pA . And if this is so, then such a philosopher would be happier (J_w) by committing an injustice (U_pA) than by acting justly (J_pA).

The oft-quoted passage where this question arises is *Rep.* 519c-521b.²⁴ Here Socrates discusses the necessity of compelling those of philosophic nature who have been trained to "see the good" to "return to the cave," i.e., to give up some time from their life of contemplation to rule in the city. The passage is the locus of two important tenets of Plato's political philosophy: first, the fact that the philosopher must sacrifice some of his own happiness during his tenure as ruler, and second, the "paradox" that only one who does not want to rule will make an effective ruler. But what to my knowledge has not often been noticed²⁵ is that here Plato makes a very un-Platonic distinction—he draws a sharp and clear contrast between what is necessary for this philosopher's happiness, and what is just for him to do in his city. And I submit that in allowing this "split" to enter his argument, Plato has opened a Pandora's box of obstacles to any attempt at refuting Thrasymachus.

Plato is quite clear on both the following points:

1. *That the philosopher's maximum happiness does not lie in ruling.* It is stated explicitly several times that there is a "better life" than ruling for philosophers (519d9; 520e5; 521b9). Much earlier (at 419a), in the midst of founding the city, Adimantus had objected that Socrates, in describing the lives of the guardians, was not making them happy. There (after reminding Adimantus that as founders of a city the maximum happiness of any one class was not their concern) Socrates could reply that Adimantus was using a rather low-level standard of what constituted happiness. Adimantus had objected to the fact that the guardians would not have property of their own, or much use of money; Socrates dismissed such materialistic considerations as merely "what is reputed to be happy" (420a7) and as a "foolish adolescent opinion about happiness" (466b8). Here, however, the alternative life is no childish chimera of happiness, but that activity wherein the greatest happiness lies, i.e., contemplation of the Forms. The "better life" is based on a "true" standard of happiness, rooted in the divine order of things. The

²⁴ Limitations of space prevent quoting the passage in full; the reader would do well, when considering the following argument, to keep a copy of this passage before him.

²⁵ M. B. Foster noted the exceptionally "un-Platonic character" of this passage some thirty-five years ago in "Some Implications of a Passage in Plato's *Republic*," *Philosophy*, XI (1936), 301-307. He saw the uniqueness of this passage not, however, as undercutting the main argument of the *Republic* but as a precursor to the truths of Christianity.

philosopher is a lover of that which truly *is*, and happiness for him lies in being close to what is real, communing with “the region inhabited by the happiest part of what is” (526e). The highest ecstasy a man can attain—compared with which concern for laws and institutions is only an “elementary initiation”—is being in the presence of the Forms themselves.²⁶

The cave image makes it clear that one must choose between turning the soul towards the unchanging Forms which *are* or towards the multiform ever-changing human affairs. If left to themselves, the philosophers “would not be willing to go down again among those prisoners or share their labors and honors” (519d). One must take seriously this reluctance to return, for it cannot be argued that these philosophers do not yet know what is in their own interest. *Ex hypothesi* (517b-d) they are the ones who have already seen the good, yet we are told that we should “not be surprised that the men who get to that point aren’t willing to mind the business of human beings, but rather their souls are always eager to spend their time above.”²⁷ A true philosopher “despises political office” (521b); ruling for him is “drudgery” (540b). It is quite true, but also quite irrelevant, that Plato acknowledges that each philosopher will only need to serve a short term in office (540b) and can spend the major part of his later years philosophizing. The fact that Plato feels it necessary to offer this apology is proof that Plato does recognize that the happiness of the philosopher is not promoted by his ruling his city.

2. *That the justice of the city necessitates that philosophers rule.* When analyzing the argument of the *Republic* it is of central importance that the reader keep in mind the frequent shifts in the main analogy between the city and the man. At any given point it is essential to ask whether it is the just city or the just man which Socrates is at that point considering. (One helpful but often unnoticed hint which Plato provides is that when it is the city which is under consideration, he

²⁶ *Symposium* 211. It is sometimes suggested that, for Plato, the activity of philosophy includes “returning to the cave,” or in some vague sense is only fulfilled when applied to the “world of becoming.” I find no justification for this in any of Plato. A man’s nature is defined by what he loves; a philosopher is explicitly said to be a lover of what *is* (474 ff.) and explicitly denied to be a lover of ruling (521b4). The philosophic nature is, understandably, suited for philosophy, and, as Plato succinctly puts it, “it is that ascent to what *is* which we truly affirm to be philosophy” (521c7). It is not part of realizing the philosophic nature to rule; all Plato ever claims is that nothing in that nature makes it impossible for such a man to rule (484-487). Indeed, the philosophic nature is the best nature for ruling; if you need a ruler, get a philosopher. But from this it does *not* follow that if you want to philosophize, you need to rule.

²⁷ The basic problem I am posing will perhaps be clarified somewhat if stated from another point of view. At the level of ordinary men, in ordinary cities, a man’s desires will not always be for what is in his true interest—as Plato puts it, what seems to him to be best may not coincide with what a man really wants (cf. *Gorgias* 466 ff.). Only of a man who knows the good will it be true that his desires (as they appear or seem to him) will necessarily coincide with his real or rational wants, i.e., with what is in his true interest. The problem of whether the (internally) just man is compatible with the just city, i.e., whether $J_w \equiv J_p$, can be posed as follows: in an ideal city, and for a man who has seen the good, will his wants be compatible with what is in the interests of the city? I am suggesting that even at this ideal level, Plato admits that there is a conflict: what the philosopher (who has seen the good [517b-d]) wants or wills (517 cd) stands in conflict with what is in the true interests of the ideal city. Or, to put the dilemma into Platonic terms: this man’s rational wants are opposed to performing $J_p A$.

frequently has Socrates refer to himself and his interlocutors as “we founders.”) The section comprising Books V through VII is explicitly said to concern the city, its realizability and the makeup of certain initially puzzling features, e.g., the equality of women, the commonality of property and children, etc.²⁸ The entire question of a philosopher’s ruling comes up only in this connection, in answer to the question of what would be functionally necessary for such a city to come into being.

When Glaucon first hears that the philosophers must sacrifice some of their own happiness, he replies, “Are we to do them an injustice, and make them live a worse life when a better is possible for them?” (519d). Glaucon, remembering the previous arguments, is perfectly correct in calling actions which lead to a less than best life “unjust”—if one is talking about the justice (and the happiness) of a man, i.e., J_w . But Socrates reminds him that we are not talking about individual men here:

My friend, you have again forgotten that it’s not the concern of law that any one class in the city fare exceptionally well, but it tries to bring this about for the whole city. . . . It produces such men in the city not in order to let them turn whichever way each wants, but in order that it may use them in binding the city together. (519e-520a)

Thus, vis-à-vis the city, Socrates explains, we (“founders,” 519c) will not be doing an injustice in requiring them to rule. In addition to compelling them, we shall try to persuade them by “saying just things” (520a). This “saying just things” can only refer to J_p (the justice of the city), for if it meant J_w (the internal justice of the individual man), this would mean saying such things as, “You owe it to yourself to do what your nature is best suited for, to mind your own internal business”; these latter statements would encourage the philosopher towards contemplation, not towards ruling. Instead, Socrates appeals to the city, to the debt the philosophers owe for their rearing and to the fact that their rule would be good for the city. (And such persuasion is bound to work—for these philosophers have been nurtured in the city, baby-fed on the “noble lie” [414] and on the upbringing that gives them a sense of kinship with the city [462-466]. They have been raised to be J_p .²⁹)

Whereas Glaucon has previously called the act of compelling philosophers to rule an “injustice” with reference to the better life available for that individual man, on hearing Socrates’ present justification—one offered entirely in terms of men considered as parts in their city—Glaucon now concludes, “We are laying just

²⁸ The equality of women and the communality are both justified in this context because they are good for the city (456e-457a; 462b).

²⁹ Whether or not the noble lie is successful in convincing the philosopher to rule does not militate against the truth of my point. The fact that a man is successfully persuaded that his true interest coincides with the interest of the city does not make that of which he is persuaded true. If the man has a philosophic nature, it is in his true interests to develop that nature and then to philosophize. In saying he ought to rule, it might be “noble” to place the interest of the city paramount, but it is still a “lie” with respect to what is in the true interest of the philosopher.

commands on those who are just" (520e). This statement makes sense again only if the "justice" referred to, of both the commands and those to whom they are directed, is the justice of the city: it is functionally necessary that philosophers rule *if the city is to be best* (473d, 520d); and such men will obey these commands because they are J_p .

If both these points are true—and it seems clear Plato wants to assert they are—then there seems to be an incompatibility between a philosopher's being completely happy and his ruling the ideal city. Moreover, not only does the passage make this incompatibility clear, but it shows further that Plato is aware of the conflict. If Plato does opt³⁰ for making the city happy, and thus devises ways of persuading the philosopher to perform J_pA , his recognition of the possible need to "compel" (προσαναγκάζειν 520a) indicates his awareness that the tension is a real one.

V. SUMMARY AND CONCLUSION

I have argued that Thrasymachus has challenged Socrates to show that for any individual both (1) being as happy as possible is compatible with being just vis-à-vis his city, and (2) that being unjust vis-à-vis his city is incompatible with that degree of happiness. I have argued that in at least one instance Plato admits that neither of these challenges can be completely met. In the case of the philosopher in the ideally just city, maximizing his happiness entails looking at the Forms as much as possible, while justice in his city entails his giving up some of this happiness to work at a task he does not enjoy. Thus, unfortunately for Plato's proof, it appears that a philosopher who has been educated by his city to see the good would be happier if he were unjust (U_p), i.e., if he "copped out" on his ruling task and instead philosophized. And, of course, we are entitled to assume by the conditions of the original challenge (361a) that he successfully gets away with his injustice. But regardless of whether he commits this particular U_pA or not, the fact seems to be that, at least for the best of men, there seems to be an irremediable conflict between developing oneself to perfection and the justice which demands sacrificing a part of oneself on behalf of one's fellow men.

Before concluding, I want to consider briefly two objections which have been raised against earlier drafts of this paper, and which may perhaps be in the minds of some readers. Both objections deny that the philosopher can in fact be maximally happy without his ruling the city. The objections, summarily stated, run as follows:

1. Socrates argues (420 ff., especially 420d and 421a) that if the best natures do not rule, the city will be destroyed. And if the "whole" is destroyed, *a fortiori*

³⁰ I am *not* claiming that Plato does opt for this choice, nor does the existence of the "noble lie" provide evidence that Plato does so opt. The "noble lie" occurs in Book III, in the context of the founding of the ideal city. Socrates' argument is simply that if one wants to found the happiest possible city, such a lie will be functionally necessary.

the "part" is destroyed along with it. Thus, according to this objection, the philosopher will destroy himself if he does not rule.

2. Socrates argues (347a-d) that in fact the philosopher will have to rule *in his own interest*, because otherwise he will suffer the penalty of having a worse man rule him. Only by *his* rule can he avoid the consequence of living under the rule of an inferior.

Neither of these objections seems to be tenable. I have already cited above Plato's many explicit statements that a better life than ruling does exist. But apart from this, both objections misconceive the nature of Plato's argument for the following reasons:

1. The first objection mistakes what it is that Socrates says is destroyed. It is not the man himself who perishes, but "the guardian," i.e., the man *qua part*. What Socrates says is that if we were to maximize the guardian's happiness, we would "turn him into something other than a guardian," i.e., he would become a full-time (happy) philosopher. What would be destroyed is *this man's being a guardian*. It is true that (perhaps because of the ineptness of second-raters at calculating the nuptial number) the city might soon perish, and with it the species of (future) philosophers, at least insofar as they are programmatically produced in such cities. But this is not an objection as to why this individual, already-existing philosopher will not be happiest while contemplating.

2. Even if this second objection were valid, it would have to be seriously modified. For at 347d1 Socrates says this argument applies if "they have no one better than *or like themselves* to whom to turn the rule over." But the philosophers do have others like themselves to rule, since the city developed a class of philosopher-rulers. It would be to the advantage of each to commit the injustice and leave ruling to the rest of those like themselves. And, by the conditions of Thrasymachus' challenge, we must consider only the life of one who successfully "gets away with" his injustice. But in fact these considerations are academic, because this second objection forgets what Socrates is trying to prove. This objection seeks to justify being J_p by reference to the consequences or "wages" that stem from performing J_pA , and it is precisely this kind of consideration that Socrates explicitly rules out. Plato evidently wants to underscore that the "negative sanction," the penalty of having to be ruled by an inferior, definitely is to be counted as a wage ($\mu\iota\sigma\theta\acute{o}\varsigma$), for he has Socrates call it such over the objection Glaucon raises *to this very word* (347a9). And then Glaucon uses this very word (358b7) to tell Socrates what he does *not* want justice praised for. To justify ruling as "good" because of what comes from it, while it itself is "drudgery," would be to place it in the category of medicine, i.e., it would be to admit that "being just" is not "good in itself."

If what I have argued throughout this paper is correct, or even approximately correct, I think it sets the problems concerning the argument about justice and happiness in a new light. Heretofore the problems of Plato's *Republic* have been seen largely in terms of a conflict between a Platonic and a vulgar notion of justice; on my interpretation, there is a real conflict between two senses of

justice, both of them Platonic. Moreover, unless Thrasymachus' challenge can be met, it would seem that what Plato is saying is that the development of a social conscience in the best of men, while necessary if mankind as a whole is to be benefited, may be at the expense of some happiness on the part of those who are wisest. Perhaps Plato, realizing this, nevertheless chose to dedicate some of his years attempting to reform the political life of Sicily; the fact that this sacrifice of philosophy resulted in only "attempts to kill the man who tried to lead them up" (517a), might account for what some have sensed as an aura of "Platonic pessimism."

University of Chicago