

## The Psychology of Justice in Plato: a Response to Cooper

Previously unpublished comments on John Cooper's paper "The Psychology of Justice in Plato" (*American Philosophical Quarterly* 14 (1977) 151-157). These comments were presented, as they appear here, along with Cooper's paper, in a symposium at the Pacific Division meeting of the American Philosophical Association in Portland, Oregon, March 1977. Since Cooper's paper has been several times anthologized and is by now (May, 2009) considered a classic in the field, it seemed appropriate to make these comments publicly available. They are unchanged from the *APA* version except that references to Cooper now conform to the pagination in the version subsequently published in *APQ*.

Plato addresses himself, in the *Republic*, to so many philosophical problems that we are now inclined to think of as separate from and unconnected with one another that it becomes tempting to treat his various discussions of them piecemeal. But although it might be congenial to a contemporary historian of philosophy to compartmentalize the *Republic* (devoting, say, two lectures to its metaphysics, two more to its political theory, and another to its epistemology), to do so would be to misrepresent Plato. It is clear that, for him, the various parts we now label metaphysics, moral philosophy, etc., were interwoven in crucial ways, all contributing to his overall aim of identifying human excellence and specifying how it is to be achieved. What can easily strike us as a collection of theories on separate topics were, for Plato, parts of a single unified theory.

I take Professor Cooper's project to be that of illustrating one important aspect of the unified theory. That is, on Cooper's account, the metaphysics of the *Republic* is intimately connected to its moral psychology; properly understood, the *Republic's* metaphysics dictates what the psychology and behavior of the just man (on Plato's conception of him, at any rate) will be. On the metaphysical side, there is the Form of the Good, its functional properties and its substantive nature; on the side of moral psychology, there is the theory of justice as a kind of psychic harmony in which reason rules.

Now on the surface, at least, these two facets of Plato's thought seem easy to connect. For, on Plato's account, the rule of reason requires knowledge, and the object of this knowledge must be the Good-itself. (I will use the labels "the Good-itself" and "the Form of the Good" interchangeably, as Plato seems to have done.) But the connection that Cooper sees runs far deeper than this obvious one, for he offers us not merely the platitude that Plato's just man possesses, and acts out of, knowledge of the Good-itself. In Cooper's view, once we appreciate the role of the Good-itself in Plato's unified theory we can see why the just man will behave in otherwise unpredictable ways: why he will "re-enter the cave" and thereby turn away from the full-time contemplation of the Form of the Good that his dialectical training has prepared him for. We will also discover something about his motivation, for, as Cooper concludes, "the just man in Plato is no egoist, and no altruist either, but a sort of high-minded fanatic" (p. 157).

Here we have the two most striking and important claims in Cooper's paper: (1) that the metaphysical conception of the Good-itself has consequences for Plato in the area of

moral motivation, and (2) that the traditional understanding of Plato's moral psychology as a kind of psychological egoism is mistaken. In my comments I will concentrate on the first claim; but I want to begin by saying something (briefly) about the second, for the two are intimately connected.

Cooper's rejection of an egoistic interpretation of Plato's moral psychology is based on the idea that the Platonically just man acts out of a concern not for *his own* good but for *the good* (p. 155). But this is not to say that the Platonically just man will ever choose to do something that is not in his own self-interest. Indeed, Cooper admits that the philosopher who, opting for the mixed life, decides to re-enter the cave, "will actually be more *εὐδαιμόνων* than any who opts for the purely intellectual life" (p. 157). And if the philosopher can be expected to know this (as seems reasonable) we seem here to have, after all, a case in which the philosopher chooses to do something that he judges to be in his own self-interest. The charge of egoism can now be avoided only by appeal to the fact that although the philosopher does what he realizes will make him more *εὐδαιμόνων* than any other course of action will, he does so not for that reason but rather because the Good-itself dictates that he should. So it is not as if we are presented with a clear counter-example to the egoistic interpretation, a case in which the philosopher chooses something other than what he supposes will be in his own best interest. His reasons may be different, but his choices are still extensionally equivalent to those he would make if psychological egoism were true. Hence, the success of the nonegoistic interpretation depends on the role of the Good-itself in providing the reasons for which the philosopher acts. Thus Cooper's second claim rests ultimately on the first, for everything now hinges on the role of the Good-itself in moral motivation. And that role is supposed to be a consequence of Plato's conception of the Good-itself.

The question whether Plato's conception of the Form of the Good has the consequences that Cooper alleges clearly depends, in large part, on the correct characterization of that Form. Here, alas, is where matters are murkiest. For while Plato likens its role in the intelligible realm to that of the sun in the visible realm, he does not have anything very illuminating to say about it. Yet how could he? The Form of the Good is the ultimate object of study of the science of dialectic, and if its nature could be revealed by Socrates in conversation, the arduous rigors of dialectic would be unnecessary. If we view the Socrates of the *Republic* as a spokesman for, but not (there, at any rate) a practitioner of, dialectic, his reticence on the topic of the Good-itself is perfectly comprehensible.

Still, the complaint that the Form of the Good does no real work for Plato is common enough. Thus, Popper writes: "Plato's Idea of the Good is practically empty. It gives us no indication of what is good, in a moral sense, i.e., what we ought to do ... (*The Open Society and its Enemies*, p. 274) ... purely formal information is all we get. Plato's Idea of the Good nowhere plays a more direct ethical or political role ... (*Ibid.*, p. 145)." If Cooper is right, this sort of view is entirely mistaken, for the Platonically just man, as a devotee of the Good-itself, will, according to Cooper, act in such a way as is "most likely to maximize

the total amount of rational order in the world as a whole” (p. 156). The Platonically just man becomes, on Cooper’s account, a kind of utilitarian with a somewhat austere mathematical and perhaps impersonal conception of what the good he is trying to maximize is like.

This seems to me to be a rather heavy burden for Plato’s Form of the Good to bear. How, on Cooper’s account, is it supposed to do it? Briefly, Cooper’s view is that the Good-itself has (1) the appropriate functional properties and (2) the sort of nature to bring it off.

(1) The Good-itself is good. It is, more importantly, not a good *anything*, but is just *good*. Not being (merely, and as opposed to everything else) a good something-or-other, it is purely and perfectly good. Any other good thing is good only to the extent that it approximates the goodness of the Good-itself. Hence (and here Cooper presupposes a good deal of the epistemology developed earlier in the *Republic* and, especially, in the *Phaedo*) one cannot know anything to be good except by knowing the Good-itself (and, presumably, making the requisite comparisons.)

(2) On the substantive side, Cooper is more tentative, but more suggestive. The Good-itself is to be thought of as “a perfect example of rational order, conceived in explicitly mathematical terms” (p. 155).

The evidence for (1) is quite straightforward: the paradigmatic role of the Forms in general, complete with traditional non-Pauline self-predication, and the epistemology of concept acquisition and deployment that is featured in the *Phaedo*’s defense of the doctrine of recollection. The evidence for (2) is much thinner: Cooper appeals to the mathematical nature of higher education in the *Republic*, the role of the Form of the Good in mathematics, and the utility of the study of mathematical harmonics for the investigation of the good.

Let us suppose, for now, that this characterization of the Good-itself is basically correct, and ask whether Cooper’s consequences for moral psychology are forthcoming. If the nature of the Good-itself is a mathematically articulable rational order, then clearly the just man, as a devotee of the Good Itself, will seek to impose such order on the world and will adopt a criterion of choice framed in terms of it. In short, the devotee of the Good-itself will choose whatever course of action most closely approximates the rational order which is the nature of that lofty Form.

But this is not quite the conclusion Cooper wants, for he sees the just man as trying to *maximize* the *total amount* of rational order in the world. Now the notion of the total amount of rational order is, on the face of it, quite a puzzling one. Indeed, it’s hard to tell what an amount of rational order would be. But, however this puzzle may be worked out, I think we will be left with the indisputable fact that approximating the ideal of rational order as closely as possible and maximizing the total amount of rational order are two different things.

Consider two alternative courses of action, one of which culminates in the production of a small, but nearly perfect bit of rational order, the other of which produces more total rational order, but spreads it thin. Imagine, if you will, a watchmaker who has the choice of making either a large number of watches each of which is marginally more accurate than those commonly in use, or a single watch which is nearly perfect in its accuracy. The watchmaker's devotion to accuracy in the measurement of time is analogous to the philosopher's devotion to rational order. I do not know how one goes about calculating either the total amount of rational order in the world or the net total of improvement in the accuracy of our measurement of time. But it does seem clear to me that, given a large enough number of watches, distributed to a large enough number of people, more total net improvement will be realized by the first course of action than by the second. But is it clear that our watchmaker, devoted as he is to perfection in time-keeping, will go in for mass production when he has the opportunity to produce one, but only one, nearly perfect watch? And is it clear that the devotee of the Good-itself would really be interested in the total amount of rational order in the world?

No, it seems to me that such a devotee would try to emulate rational order to as great a *degree* as he can, even if this prevents him from maximizing the total *amount* of rational order in the world. Cooper may be right in construing the Platonically just man as a consequentialist, one who chooses his actions on the basis of their consequences; and he may be right in supposing that consequences are measured in terms of rational order; but I do not see that he has made a case for saying that the Platonically just man is a utilitarian. It is true, of course, that Plato sees his guardians as trying to produce "happiness in the city as a whole" rather than in any special class (519e), but that is not the point. What is at issue is whether this sort of behavior can be derived from the role of the Good-itself as Cooper conceives of it. And it seems to me that it cannot; the devotee of the Good-itself is at least as likely to be the sort of (perhaps perverse) perfectionist that I have imagined as the utilitarian that emerges on Cooper's account.

There is another difficulty with the conception of the Good-itself that Cooper proposes that I would like to discuss. Here, however, I am not sure whether the difficulty is peculiar to Cooper's account or whether it is really Plato's problem. (The latter seems more likely to be the case.) The difficulty arises when we consider a bit further the metaphysical and epistemological roles of the Form of the Good as Plato presents them.

It is clear that Plato holds that a recognition of the Form of the Good, including, presumably, an awareness of its nature, is required if one is to know any other good (534b-c). Now this epistemological requirement, by itself, does not distinguish the Form of the Good from any other Platonic Form. If I cannot know that an action is good unless I have an adequate grasp of the Good-itself, neither can I identify something as a triangle unless I am similarly *en rapport* with the Triangle-itself. But it seems to me that the Form of the Good is intended to be set apart from the other Forms in terms of this epistemological role. The role that the Triangle-itself plays in our knowledge of triangles and our ability to identify them is played by the Good-itself in our knowledge of *any* Form, and hence in our

ability to identify anything at all. Nicholas White, in his recent book, puts the point this way: "... Plato appears to be saying that we cannot view other Forms, at least with full clarity, until we have viewed the Form of the Good" (*Plato on Knowledge and Reality*, p. 100). That the Form of the Good indeed plays this monumental epistemological role in Plato's thought seems to me to be strongly suggested by Plato's description of the Form of the Good as the "unhypothesized beginning" (510b7) of the downward path of dialectic, a process which "attempts to apprehend methodically, with regard to each thing, what each really is" (533b2-3, trans. Grube). There is further confirmation in Plato's characterization of the objects of knowledge (*τὰ γιγνωσκόμενα*) as owing not only their being known (*τὸ γιγνωσκέσθαι*) but also their very being (*τὸ εἶναι τε καὶ τὴν οὐσίαν*) to the Good-itself (509b6-9).

Why should the Form of the Good play this role for Plato? One common answer (cf. Ross, Hare, White) is that Plato is operating with the idea that one cannot know what it is to *be F* unless one knows what it is to be a *good F*. A somewhat fuller account of this line of thought would run like this:

One cannot apprehend something as being *F* unless one knows what it is to be *F*. To know what it is to be *F* is to recognize the *F* itself, a standard of comparison against which impure, defective, and imperfect *F* things can be compared. But the *F* itself is purely, non-defectively, and perfectly *F*; and nothing would count as recognition of the *F* itself if it were not recognition of the *F* itself as being purely, non-defectively, and perfectly *F*. Hence, to know what it is to be *F* requires a recognition of something as *perfectly F*. It is in this sense that the Form of the Good is the source not only of our knowledge of, but also of the being of, such objects of knowledge as the *F* itself. The perfection of the *F* itself is what is contributed by the Form of the Good.

There are two related consequences of this conception of the Form of the Good that must be noted. First, the notion of goodness involved is that of perfection—to be good, in this sense, is to be a perfect exemplification. Second, this notion of goodness is that of goodness *of a kind*—to be good is to be a good such-and-such. (Even if, as Cooper notes, the Good-itself cannot be said to be a good such-and-such, the sorts of thing judged to be good in light of that Form will always be judged to be good such-and-suches—good things of a kind.)

Now goodness of a kind may have some connection with moral goodness, but it is certainly not the same notion. There may be good men and good actions, but there are also good football players and good burglars. One might aspire to be perfectly kind, but, one might also aspire to commit the perfect crime. Presumably, Plato's devotee of the Good-itself will not aim at all of these kinds of perfection, but why not? It seems to me that Plato might indeed have a lot to say in answer to this question, but I doubt that in his answer he could make any significant or helpful appeal to the nature of the Good-itself. Certainly Cooper's rational order interpretation will not help him, for there may be more rational order in the perfect crime than in many pursuits more honorable but less orderly and less

thoroughly thought out. (One might always complain that the unhealthy desires of the master-criminal show that in him it is appetite and not reason that rules, thus disqualifying him from being Platonically just and hence showing him to be no devotee of the Form of the Good. But this is only relevant if the criminal is, for example, merely after the money; it does not touch the case of the man who is attracted by the sheer perfection of his plan, rather than by any possessions he will acquire by means of it.)

Of course, this interpretation of the Form of the Good, while widely held, is quite conjectural. And so one might attempt to come up with another interpretation according to which it is indeed moral goodness that is the nature of the Good-itself. But even if such an interpretation were possible (which seems unlikely) it would face problems elsewhere. For as soon as we manage to get Plato's metaphysics in line with his moral theory, we seem to sever the connections between his metaphysics and his epistemology. The contribution that the Form of the Good is supposed to make to the objects of knowledge makes sense only if the good is not restricted to moral goodness. Otherwise, how could the Good-itself play any role in my knowing what it is to be a burglar?

Plato, of course, only tentatively and infrequently supposes that there are "bad" Forms (cf. *Rep.* 476a, *Parm.* 130d-e) and we might urge that he was careless in so supposing and better advised not to admit such Forms. This would be bad advice, however. For one thing, even if there were no "bad" Forms, there would be enough morally (or aesthetically) neutral ones to introduce the same problem. But more importantly, to try in this way to splice the metaphysics and the moral theory would cause irreparable damage to Plato's semantic theory. For predicates expressing morally neutral and bad properties are meaningful and learnable parts of our language, and Plato's semantic theory requires that there be Forms corresponding to such predicates (cf. *Rep.* 596a).

It has been suggested (by Vlastos, "Degrees of Reality in Plato," pp. 7-8) in a similar context that Plato is operating with two senses of "real," in one of which it is a value-predicate and in the other means something like "cognitively reliable." The same might be said about his use of "good." But to recognize this is to recognize the central failure in Plato's attempt to provide a metaphysical and epistemological foundation for morality. For without a demonstrable connection between these two senses, the moral consequences Plato wishes to draw are founded on an equivocation and will not stand.

How fair are these criticisms of Plato's enterprise? In so far as Plato is advancing a single theory designed at once to solve problems in metaphysics, epistemology, and moral theory, they seem to me to be fair indeed. Cherniss has spoken in glowing terms of the "philosophical economy" of Plato's theory of Forms. Economy, perhaps; but I think it is only fair to add that (in the *Republic*, at any rate) what Plato is practicing is false economy.

S. MARC COHEN  
UNIVERSITY OF WASHINGTON