
Are There Two Theories of Goodness in the Republic? A Response to Santas

Rachel Singpurwalla

It is well known that Plato sketches, through his similes of the sun, line and cave, an account of the good in the middle books of the *Republic* (504-535) and that this conception of the good relies heavily on his theory of forms. What is considerably less well noted is that Plato presents a distinct account of goodness — a functional account — in Book I of the *Republic* (352e-4a). Gerasimos Santas has, in numerous influential articles, offered an interpretation of Plato's two theories of goodness and their relationship to one another.¹ In this short discussion, I will argue that Santas' account of the two theories of goodness and their relation-

-
- 1 Santas provides an interpretation of the form of the good in 'The Form of the Good in Plato's *Republic*'. This article was originally published in *Philosophical Inquiry* (1980) 374-403. It has been reprinted in Anton, J.P. and Preus, A., eds., *Essays in Ancient Greek Philosophy*, Volume 2 (Albany: State University of New York Press 1983), 232-63; and Fine, G., ed., *Plato*, Volume 1 (Oxford: Oxford University Press 1999), 247-74. Santas provides an account of the two theories of goodness and their relationship to one another in 'Two Theories of Good in Plato's *Republic*', *Archiv für Geschichte der Philosophie* (1985) 223-45, but most recently in his *Goodness and Justice* (Oxford: Blackwell 2001), 58-193. All page references in this article refer to this latest version of his argument. Alternative recent accounts of the good are given in J. Annas, *Platonic Ethics, Old and New* (Ithaca, NY: Cornell University Press 1999), 96-116; John M. Cooper, 'The Psychology of Justice', *American Philosophical Quarterly* (1977) 151-7; G. Fine, 'Knowledge and Belief in *Republic* 5-7', in S. Everson, ed., *Cambridge Companions to Ancient Thought, I: Epistemology* (Cambridge: Cambridge University Press 1990), 85-115; D. Hitchcock, 'The Good in Plato's *Republic*', *Apeiron* 19 (1985) 65-93; and C.D.C. Reeve, *Philosopher-Kings* (Princeton, NJ: Princeton University Press 1988), 81-94.

APEIRON a journal for ancient philosophy and science
0003-6390/2006/3904 319-330 \$12.00 ©Academic Printing & Publishing

ship to one another cannot be correct. My critique of Santas' interpretation of Plato's theories of goodness has implications with regard to the issue of the relationship between form and function in Plato's thought, for the problem with Santas' account is due to his interpretation of Plato's conception of form and function. Thus, this discussion makes explicit the consequences of a certain interpretation of form and function for the relationship between the two in Plato's metaphysics of value.

The Functional Theory of the Good

According to Santas, a close examination of the *Republic* shows that there is a functional theory of goodness at work in the early and late books of the dialogue. According to Santas, Plato presents the functional theory of goodness in Socrates' final argument against Thrasymachus in *Republic* I (352e-4a). Moreover, Santas argues that this theory is the theoretical foundation upon which Plato builds up his ideals of the completely good city and the completely good person. Santas summarizes the functional theory in seven claims:²

1. Some things have a function.
2. The function of a thing is the work of a thing that it only or it better than anything else can perform.
3. Things with a function may perform their function well or poorly.
4. There is a virtue for everything that has a function.
5. A thing performs its function well by its own proper virtue, badly by its own proper vice.
6. A thing with a function is a good of its kind if it carries out well the function of things of its kind.

2 *Goodness and Justice*, 66-7

7. Functioning well and the virtue that enables a thing to function well are the good or part of the good of that thing.

According to Santas, these seven propositions state the theoretical backbone or the formal aspect of the functional theory of the good. They do not yet, however, tell us which things are good or which things are virtues. The theory implies that to discover what the good or the virtue of a particular thing is, one should follow a three-step procedure. First, find the functions of that thing. Second, isolate its performing these functions well. Third, find the conditions that enable it to function well; the conditions that enable it to function well are the virtues of that thing. Thus, when we conjoin the formal theory with factual propositions about what the function of specific things are, what it is for these things to perform that function well, and what enables these things to perform their function well, we can arrive at specific conclusions about goodness and the virtues. According to Santas, the functional theory thus conceived enables us to understand Plato's procedure for discovering the perfectly good city and what justice is in the city and the individual in Books II, III and IV of the *Republic*, for Plato begins by identifying the function of a city (to provide for the needs of the citizens), constructing a city which performs these functions well, and identifying the conditions that enable it to do so; these conditions are the virtues of a city. Plato then endeavors to show that this same account of the virtues applies to the individual.

The Metaphysical Theory of the Good

Although the functional theory of the good might enable us to understand Plato's procedure in the early books of the *Republic*, it must be admitted that the functional theory of the good is not the explicit theory of the good in the *Republic*. Rather, the explicit theory of the good is contained in the middle books of the *Republic*, where Plato discusses the form of the good in his similes of the sun, line and cave. Accordingly, Santas provides an interpretation of this characterization of the form of the good and its relation to the functional account of the good.

Santas derives his account of the form of the good from an account of the nature of the forms; and rightly so, since Plato says that the form of the good is that which gives the forms both their knowability and their being and essence (507b-11e). According to Santas, we should conceive

of the forms as ideal exemplars complete with self-predication.³ Thus, the form of the circle, for example, is a perfect circle, and the form of beauty is itself something beautiful, indeed, the most beautiful thing there is.

On this account of the forms we can say that the forms have two sorts of attributes: proper attributes and ideal attributes. The proper attributes of a form are those that it has in virtue of being the particular form that it is. Thus, the proper attribute of the form of the circle is being circular; and the proper attribute of the form of beauty is being beautiful. The ideal attributes of a form are those that it has simply in virtue of being a form; the ideal attributes of the forms are that they are perfectly, purely, and always whatever it is that they are. That is, the form F has all of the defining features of what it is to be F; it has all of these features to the highest possible degree; and the form F is not not-F in any respect, time, place or relational environment. Thus, the ideal attributes of the form of the circle are being perfectly, purely, and eternally a circle; and the ideal attributes of the form of beauty are being perfectly, purely, and eternally beautiful.

According to this account of the forms, then, we should think of the forms as the best objects of their kind, for their ideal attributes make them the best objects of their kind. Because all of the forms have the feature of being the best object of their kind, there must be some form in virtue of which they all have this feature. And this, according to Santas, is the form of the good. Since, the form of the good is the (formal) cause of all of the other forms having their ideal attributes, we must conceive of the form of the good as constituted by the very ideality common to all the other forms in virtue of which they are the best objects of their kind. The ideality that constitutes the form of the good must, however, be conceived very abstractly, since it does not contain the proper attributes of the forms. The form of the good, unlike the other forms, is not a good something or other; it is, rather, superlatively good.

This interpretation of the form of the good elegantly accounts for Plato's claims in *Republic* VI that the form of the good explains the knowability and the being and essence of the forms (507b-11e). For the forms, unlike sensible objects, are knowable because they are perfect,

3 My summary of Santas' metaphysical theory of the good is drawn from *Goodness and Justice*, Chapter 5.

pure and eternal. And the being or essence of a form qua form is perfection, purity, and eternity. And these features are features that the form has because they participate in or resemble the form of the good.

In addition, this interpretation of the form of the good can explain the goodness of sensible objects. According to Santas, the goodness of sensibles can be explained by their participation in the form of the good. To be a circle, for example, a sensible must participate in or resemble the form of the circle; it must, that is, participate in or resemble the proper attributes of the form. To be a good circle a sensible must participate in or resemble (to some degree) the ideal attributes of the form of the circle. But since the form of the circle has its ideal attributes in virtue of participating in the form of the good, a particular sensible circle is good in virtue of participating in or resembling, albeit indirectly, the form of the good.

The Relationship Between the Functional and Metaphysical Theories of Goodness

On Santas' interpretation of the *Republic*, then, the *Republic* contains two independently stated and apparently distinct theories of goodness. The early and late books contain the functional theory, which Plato uses to construct the ideal city and soul; and the middle books contain the metaphysical theory, the theory that Plato claims is prior to and more fundamental than every other theory in the *Republic*. But what, exactly, is the relationship between these two theories? And can a work that presents and makes use of two apparently distinct theories of value be a coherent piece of philosophy?

Santas embraces the idea that the two theories of goodness are indeed distinct and that the *Republic* is a coherent work. He argues that the two theories can be distinct and the *Republic* coherent provided that the following two criteria are met. First, the two theories of goodness in the *Republic* must be in agreement: when conjoined with the same factual propositions, the two theories must yield the same conclusions about which things are good. The second criterion that the two theories must meet for the *Republic* to be a coherent work is that the metaphysical theory must be prior to and more fundamental than the functional theory.

Are these criteria met? According to Santas, the criteria are met if Plato is making the following reasonable assumptions: (1) the function that a particular has depends on what it is (i.e., what form it resembles); and

(2) how well a particular carries out its function depends on the extent to which it resembles its form. If Plato is making these assumptions, then the two theories are in agreement, for if the function of a particular sensible object depends on its form, and its functioning well depends on its closely approximating the relevant form, then one should be able to judge the value of any particular thing using either of the two theories and get the same result. In addition, the metaphysical theory is prior to or more fundamental than the functional theory, for something's functioning well *is a result* of how closely it resembles its form, or the things that are perfectly good. And these in turn are perfectly good of their kind in virtue of participating in the form of the good.

An Initial Problem

In what follows, I will argue against Santos' account of the two theories of goodness in the *Republic*. I will begin by showing that there is a tension between Santos' two criteria for the relationship between the two theories of goodness; that is, I will show that each criterion makes a demand that suggests that the other criterion will not be met. I will then argue that the tension can be resolved if we revise the agreement criterion. The picture that results from this resolution, however, highlights deep problems for Santos' account of the two theories of goodness and their relationship to one another.

Let us begin by taking a closer look at the criteria that Santos thinks the two theories must meet in order for the *Republic* to be a coherent work. According to the first criterion the two theories of goodness must be in agreement; that is, when conjoined with the same factual propositions, the two theories yield the same conclusions about which things are good. *Prima facie*, it looks like we should interpret this as a demand for extensional equivalence: the two theories of goodness are in agreement if the class of good things is the same according to each theory.

According to the second criterion, the metaphysical theory must be prior to and more fundamental than the functional theory; that is, the function of a thing is due to its form and if a thing functions well it is due to the fact that it resembles its form. Now there are two ways of stating the relationship between form and function that accounts for these claims. The first way is to understand the form of a thing in terms of the function of that thing; in this case, a thing's form just is its function. The second way is to assert that form somehow determines function, but form is not identical with the function of a particular thing.

According to Santas, the priority criterion requires that we reject the claim that the form of a thing is its function, for if form and function were identical, it would be unclear why the metaphysical theory is prior to or more fundamental than the functional theory. Now, I must admit it is a bit unclear to me what Santas means when he says that the metaphysical theory must be prior to the functional theory, and so it is equally unclear to me why identifying form and function would make the metaphysical theory any less prior. There is, however, a more serious consideration in favor of the idea that form and function cannot be identical, and this is that Santas conceives of forms as perfect paradigms. If the form of a thing were identical with its function, then forms would be functional paradigms, or objects that carry out their functions perfectly. But Santas argues that the forms do not have functions in the relevant sense for they are 'at rest': they do not perform any activities, they do not do anything, and they may not be in time. Accordingly, Santas says that, 'the form of the eye does not see, the form knife does not cut, the form man does not deliberate, the form city does not protect anyone, and nobody can sleep in the perfect bed, not even the perfect man.'⁴ According to Santas, the fact that forms do not have functions prevents Plato from identifying form and function. He says, '... Plato would not be inclined to identify what a sensible is with its function, or to construct functional definitions of sensibles, as Aristotle sometimes did.'⁵ Thus, we cannot describe the relationship between form and function as one of identity. We should, then, account for the claim that the function of a thing is due to its form and that functioning well is due to approximating the relevant form with the claim that form determines function, but form is not identical with the function of a thing.⁶

4 *Goodness and Justice*, 188

5 *Ibid.*, 189

6 The claim that form determines function seems reasonable enough. As Santas notes, it seems that the function or functions that a particular has depends on its structure, organization, or characteristics. And whatever structure, organization, or characteristics a particular has, it has in virtue of resembling some form or forms. The examples bear this out: by changing the structure of a wheel, we can affect its rotation; by changing the structure of an eye, we can affect its seeing. In each case, the form of a thing affects what a thing can do and how well it does it.

To recapitulate, then, in order for the *Republic* to be a coherent work, the two theories of goodness must be in agreement, and the metaphysical theory must be prior to the functional theory. But now that we are clearer on how we should understand these criteria, we can see that they are in a bit of tension, for the agreement criterion seems to be a demand for extensional equivalence, and the priority criterion requires that form and function are not identical. But if form is not identical to function, then it surely seems *possible* that something could have a certain form and yet not have a function. And if this is the case, then why should we be so sure that the metaphysical and functional theories will pick out the same class of good objects?

This possibility becomes a reality when we consider the case of forms themselves. Recall that Santas thinks that (i) forms are the best objects of their kind, and (ii) forms do not have functions. According to the metaphysical theory of the good, then, forms are good; indeed, they are the best objects of their kind. The functional theory, however, would not conclude that the forms are good. According to the functional theory, what makes something good is that it functions well. If something does not have a function, then the functional theory is silent regarding its value. Thus, when it comes to the value or goodness of the forms the two theories are not in agreement. We will face this problem with any other object that we might call good, but that does not have an obvious function, for example, sensible triangles or circles, or elements such as gold.

The Problem Deepened

Now, if Santas is correct in arguing that there are two distinct and independent theories of goodness in the *Republic*, and if it is correct that each theory claims that different things are good, then Plato is in trouble, for then the *Republic* is not a coherent work. There is, however, a possible resolution to this problem. Since any object that functions well closely resembles its form, the functional and the metaphysical theory will be in complete agreement when it comes to objects with functions. Perhaps, then, we should understand the agreement criteria not as a demand for extensional equivalence over all of the possible objects of evaluation, but as a demand for extensional equivalence over a certain range of objects, namely, those objects with functions. Accordingly, we should understand the functional theory not as a complete theory of goodness, but as a partial or incomplete theory of goodness. It gives an account of what makes things with functions good, but it does not, indeed it cannot, give an account of what makes things without functions good.

On this picture, we have a theory of goodness, the metaphysical theory, which states that what makes something best is being perfectly, purely, and eternally what it is. In addition, this theory can account for the goodness of all objects of value. We also have another theory of goodness, the functional theory, which states that what makes something good is that it functions well; this theory can only account for the goodness of some sub-set of the objects of value, namely, those objects with functions. Finally, the metaphysical and the functional theory of goodness are in perfect agreement with regard to the sub-set of the objects of value that the functional theory can account for, namely, those objects with functions. Thus, we have a theory of the good that explains the goodness of all possible objects of value; in addition, we have another theory of the good that gives a different explanation of the goodness of some sub-set of the objects of value.

In addition to the theoretical strangeness of this picture, this account of Plato's theory of value is problematic for at least two reasons. On this account, two attributes make some sensible objects — those with functions — good: being perfect and pure and functioning well. Only one attribute, however, makes some other objects — those without functions — good: being perfectly, purely and eternally what they are. But if all this is true, then the sensibles can have a good-making attribute that the forms, since they do not have functions, can never hope to have. Thus, sensibles can have two good-making properties, while forms can only have one. There might be a sense, then, in which some sensibles have more goodness than the forms. And surely, this is not a result that Plato would endorse. Now we might escape this conclusion by arguing that even if the sensibles have some kind of goodness that the forms can never have, the perfection, purity and eternity of the forms so far exceeds the goodness that the particulars are capable of having that forms will always be better than sensibles. Still, though, if it remains true that the sensibles have a kind of goodness that the forms cannot have, then we might wonder why we should think of the forms as the best possible objects of their kind.

Finally, there is another worry with attempting to maintain that according to Plato there are two good-making properties: being perfectly and purely what you are and functioning well, and this is that Plato explicitly states that there is only one form of goodness and this seems incompatible with the idea that there are two non-identical properties that make things good. We ought to conclude, then, that there is in fact just one good-making property.

But which property is it? Is there any way of preserving Santas' idea that there are two theories of the good at play in the *Republic*? Now one option, of course, is to reconsider the claim that form and function are not identical; that is, we could attempt to defend the claim that being perfectly and purely what you are and functioning well are, in some objects, objects with functions, one and the same property. On this view, the functional theory of goodness and the metaphysical theory of goodness would simply be alternative ways of pointing to the same property. On this account of the relationship between the two accounts of goodness, we must assume that some objects, objects with functions, have a functional account; that is, their essence must be given in functional terms.

Can we understand the relationship between the functional and the metaphysical theory in this way? According to Santas, what prevents us from making this move is that forms do not have functions. Now I think that Santas is too quick to say that forms cannot have functions at all, at least not in the loose sense of function defined in *Republic* I. In Book I, Plato states that the function of a thing is that which it only or it can do best (352e). But on this account, forms could have functions, for surely one thing that only forms can do is cause or explain the particulars. But saying that forms could have the function of being the formal cause of the particulars, or even some other function, is not particularly helpful. For what we are interested in is the relationship between a particulars' functioning well and its resembling the relevant form, or being perfectly and purely what it is. If functioning well and being perfectly and purely what you are are to point to the same property, then the essence of some particulars must be defined in functional terms. But Santas' interpretation of Plato's claims about self-predication prevents us from defining form in terms of function. For recall that on Santas' account forms somehow exhibit the essences of particulars. If, then, the essence of particulars with functions is to have a certain function, then the form of those particulars must have that specific function. And admittedly, it is difficult to make sense of the idea of say, the form of a human, actually reasoning.

The source of Santas' problem is his interpretation of Plato's conception of the forms, specifically, his very literal interpretation of Plato's claims about self-predication. But it is not open to Santas to rethink his interpretation of Plato's claims about self-predication, for his interpretation of the form of the good rests on the idea that forms are the best objects of their kind because they literally have the characteristics of their kind perfectly, purely and eternally. It seems, then, that Santas is on the

horns of a dilemma: either functioning well and being perfectly and purely what you are all describe the same property, in which case the functional theory of goodness is saved, but the metaphysical theory crumbles, or functioning well and being perfect and pure are distinct properties, in which case the metaphysical theory is saved, but the functional theory flounders.^{7,8}

Department of Philosophy
Southern Illinois University
Edwardsville, IL 62026-1433
U.S.A.
rsingpu@siue.edu

7 Note that even if my criticism is accurate, Santas can still hold that functioning well plays an important role in our determinations of goodness, for, according to his account, if something functions well then it must resemble the relevant form and thus it must be good. So, while functioning well does not *make* something good, it can serve as a reliable indicator of whether or not something is good.

8 I would like to thank Eric Brown, Jim Hankinson, William Larkin, Christopher Kelly, Gerasimos Santas, Christopher Shields, Daniel Sturgis, Shelley Wilcox, audience members at the 'Author Meets Critic: *Goodness and Justice*' session of the Pacific APA, and at the University of Missouri in St. Louis philosophy department for helpful comments and discussion.