

Self-Predication and the Third Man

Author(s): P. Schweizer

Source: *Erkenntnis* (1975-), Vol. 40, No. 1 (Jan., 1994), pp. 21-42

Published by: [Springer](#)

Stable URL: <http://www.jstor.org/stable/20012526>

Accessed: 29-08-2015 17:51 UTC

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Springer is collaborating with JSTOR to digitize, preserve and extend access to *Erkenntnis* (1975-).

<http://www.jstor.org>

P. SCHWEIZER

## SELF-PREDICATION AND THE THIRD MAN

**ABSTRACT.** The paper addresses the widely held position that the Third Man regress in the *Parmenides* is caused at least in part by the self-predicational aspect of Plato's Ideas. I offer a critique of the logic behind this type of interpretation, and argue that if the Ideas are construed as genuinely applying to themselves, then the regress is dissolved. Furthermore, such an interpretation can be made technically precise by modeling Platonic Universals as non-wellfounded sets. This provides a solution to the Third Man regress, and allows a consistent reading of both self-predication and the singularity of the respective Forms.

### 1. INTRODUCTION

The present paper will examine the structure of the Third Man Argument (TMA), which Plato briefly considers in the *Parmenides* dialogue, in an attempt to analyze the logical dynamics that underlie this celebrated objection to the theory of Ideas. A number of modern authors, beginning with Vlastos<sup>1</sup> and Wedberg<sup>2</sup>, have argued that Plato's notion of self-predication plays a crucial role in the infinite regress argument. The original contention of Vlastos and Wedberg is that the self-predicational aspect of the Ideas is inconsistent with other fundamental principles of Plato's theory, and that the Third Man is thus the result of mutually incompatible tenets. However, even those who construe the regress as the result of mutually consistent principles grant that self-predication is a key factor in the dynamics of the argument.<sup>3</sup> Accordingly, the Third Man is generally viewed as an unhappy consequence of the principle that an Idea must always be applicable to itself, and at the very least, the regress is seen to conflict with the (independent) view that a single, unique Form should preside over all the objects which share a relevant property (henceforth this will be referred to as the 'Uniqueness Principle' or UP).

However, in the following discussion, I offer a critique of this characterization of the logical dynamics underlying the regress, and argue that, contrary to the received view, the Third Man is a consequence of Plato's system only insofar as the Ideas are *not* genuinely self-predicational. I will maintain that if the Ideas are construed as applying

*Erkenntnis* 40: 21–42, 1994.

© 1994 Kluwer Academic Publishers. Printed in the Netherlands.

to themselves, then there is no regress, and that Plato's doctrine that the Ideas are self-applicable can in fact be interpreted in a fashion which is perfectly consistent with the uniqueness and singularity of each Form. The consistency of this approach will be illustrated by supplying a set-theoretical analogue of these Ideal semantical objects. Under the proposed formal interpretation, self-predication will be construed as self-membership, and thus Platonic Universals will be modeled as non-well-founded sets. This tactic provides an elegant solution to the regress, and it shows that, rather than contributing to a logical absurdity, or, at best, to an unwanted myriad of superfluous Forms, the Platonic principle of self-predication can instead be viewed as a technically coherent and ontologically parsimonious notion.

## 2. THE STRUCTURE OF THE TMA

Unfortunately, Plato does not offer a detailed analysis of the Third Man Argument, nor does he provide a definitive account of his theory of Ideas. Thus any particular diagnosis of the regress will involve substantial and unavoidable reconstruction. Parmenides' formulation of the argument is exceedingly terse, and in order to understand his objection, it is necessary to go beyond what is explicitly stated in the text. Suppressed premises must be supplied in order to fill in the gaps, and this in turn requires a semi-formal codification of Plato's doctrine, in order to provide a plausible and well defined source from which suppressed premises may be drawn. However, in the discussion below, textual difficulties concerning the process of reconstruction will not be addressed, but rather the focus will be on the *logical* motivations behind reconstructions which have already been proposed by modern interpreters of the Platonic corpus.

The argument as it appears in the *Parmenides* dialogue (132a1–b2) begins with a premise which expresses the basic principle through which the existence of an Idea is inferred:

I [Parmenides] imagine that your reason for assuming one idea of each kind is as follows:- Whenever a number of objects appear to you to be great there doubtless seems to you to be one and the same idea visible in them all; hence you conceive of greatness as one.<sup>4</sup>

In other words, given some collection of particular things which share

a common property, say *Y*-ness, there must be some single entity, the Idea of *Y*-ness, by virtue of which the particular things in the collection are all (a) *Y*. This will subsequently be referred to as the Abstraction Principle. Granted this principle, Parmenides moves to the next step in his argument:

But now, if you allow your mind in like manner to embrace in one view this real greatness and those other great things, will not one more greatness arise, being required to account for the semblance of greatness in all these?

In this step, a new collection is formed, which includes all of the members of the original set of great things, plus the Idea of Greatness which was inferred on the basis of the first set. Utilizing the Abstraction Principle, a new Idea of Greatness will be postulated in order to account for the common character of the members of this new collection. The crucial structure of this step is expressed by the following schema:

The Idea of *Y*-ness and all of the objects which are (a) *Y* form a new class, whose members all share the common property of being (a) *Y*.

This will subsequently be referred to as the Feedback Principle. Given these two principles, Parmenides combines them to generate an infinite regress:

Then another idea of greatness now comes into view over and above absolute greatness and the individuals which partake of it; and then another, over and above these, by virtue of which they will all be great, and so you will be left with not a single idea in every case, but with an infinite number.

So the regress appears to depend upon two basic assumptions. The Abstraction Principle supplies the procedure for inferring the existence of an Idea, and the Feedback Principle provides the justification for iterating this procedure indefinitely. Given this formulation of the regress, the first question to examine is why these two principles are thought to be components of Plato's metaphysical/semantical theory.

The role of the Abstraction Principle is comparatively clear, since it seems to enshrine the basic intuition upon which Plato's semantical realism depends. The fact that divergent objects can truly share a common property is accounted for by postulating a universal entity in

which all of the divergent objects 'partake.' This universal entity is viewed as ontologically independent of the particular objects which participate in it, and indeed, the Universal is seen as supremely real, while the ordinary objects which resemble it are relegated to an inferior realm of imperfection and impermanence.

Accordingly, transient and shadowy particulars are each (a) *Y* by virtue of their participation in or resemblance to the perfect and immutable Idea of *Y*-ness. And this immutable and perfect Idea is the referent of the predicate expression in the sentence '*b* is (a) *Y*'. So the Abstraction Principle is intimately linked to Plato's semantics for predication, as characterized by the following principle:

- (0) The sentence '*x* is (a) *Y*' is true iff *x* participates in the Idea of *Y*-ness.

This Platonic analogue of the Tarski biconditional schema is generally accepted as crucial to Plato's theory of truth for atomic predications. Any true sentence of this form will imply the existence of an Idea corresponding to the expression 'is (a) *Y*'. It is by virtue of participation in the relevant Idea that the linguistic phenomenon of predication is explicated, and thus it is the Idea which supplies the semantical/ontological foundation in terms of which property ascriptions are interpreted. So the Principle of Abstraction seems to constitute a central pillar of the Platonic edifice, and it supplies sufficient grounds for inferring the existence of an Ideal object as the semantical basis for true predications.

The theoretical status of the Feedback Principle is somewhat less clear. Wedberg, for example, argues that it results from the covert interaction of two mutually exclusive axioms of the theory of Ideas, which are expressed in his codification of the Platonic doctrine as:

- (1) An Idea is never one among the objects participating therein,  
and  
(2) The Idea of *Y*-ness is (a) *Y*.

These axioms are mutually exclusive given the above semantics for predication, and the contradiction can be made fully explicit in the following manner. Tenet (0), in combination with axiom (2), immediately yields the result that the Idea of *Y*-ness participates in the Idea of *Y*-ness. This notion can be expressed as:

(2') The Idea of *Y*-ness participates in itself.

Yet (1) is clearly equivalent to:

(1') An Idea can never participate in itself,

and tenets (1') and (2') are in flagrant opposition, which makes the collection of (0), (1) and (2) inconsistent. Nevertheless, Wedberg maintains that both (1) and (2) are genuine postulates of the theory of Ideas. He holds that (1) captures the intuition which is expressed metaphorically by Plato's phrase 'one over many.' This metaphor is used to describe the nature of the relationship between a Universal and its participants, and Wedberg interprets the phrase as implying that an Idea must be distinct from the many objects over which it applies. If an Idea were one of the objects which participated in that *same* Idea, then the 'one' would be one of the 'many,' and the metaphor would presumably be violated.

In turn, principle (2) is taken to capture the notion of self-predication adopted in Plato's system. It can be substantiated by various textual illustrations, such as the claim that 'Beauty is beautiful', (*Symposium* 210e, *Phaedo* 100c), and that 'Justice is just and Holiness holy' (*Protagoras* 330cd). Wedberg identifies (2) as the theoretical malady responsible for the Third Man, and as the notion which should be excised in order to repair the system. He states that, 'In contradistinction to Plato, Aristotle shows that he knew exactly how the infinite regress arises, *viz.* from thinking of the general Idea as a particular partaking of that Idea.'<sup>5</sup> Thus self-predication is diagnosed as the conceptual defect which underlies the Feedback principle and the Third Man regress. The Feedback Principle is seen as a (confused) attempt on Plato's part to interpret (1) and (2) consistently; first (2) is considered in isolation to group Ideas and particulars together, then (1) is considered in isolation to generate a 'new' Idea, and the Third Man emerges by switching back and forth between (1) and (2), while deftly never applying them both at the same time.

Other commentators, beginning with Sellars,<sup>6</sup> have supplied more charitable interpretations of the TMA. For example, Cohen<sup>7</sup> identifies the logical catalyst in the regress to be a version of Abstraction which embodies a mutually *consistent* reading of self-predication and non-identity. On this account, the one-over-many principle is expressed in terms of Forms which are 'immediately' over maximal sets of *Y* objects,

and an infinite regress is elegantly deduced from plausible and consistent premises. Although self-predication is not explicitly appealed to in the dynamics of the regress, it is built directly into the structure of the one-over-many principle, and hence plays a vital role in the argument. In this manner, the Third Man is saved while the uniqueness and singularity of the Forms is lost.

However, in the next section, I will argue that an infinite regress can be sustained only to the extent that principle (2') is false, and hence to the extent that the Ideas are *not* genuinely self-predicational. In order to derive the Third Man from a satisfiable premise set, it is necessary to abandon the principle of self-application, and the attempt to view the regress as incompatible with UP must either beg the question or else incorporate premises which are themselves mutually exclusive. In accordance with this critique of the logical dynamics of the TMA, I will advance a univocal reading of self-predication, which does not engender a regress and hence which preserves the feature that Parmenides' objection is designed to refute, namely, that a single Idea underlies the attribution of a single character or property.

### 3. SELF-PREDICATION VERSUS NON-IDENTITY

In an important sense, it is the interaction between self-predication and non-identity, in some guise or other, which furnishes the logical kernel of reconstructions which attempt to save the Third Man. It is purportedly the claim that the Idea of *Y*-ness is *itself* (a) *Y* which fuels the continual reapplication of the Abstraction Principle. And since, by the principle of non-identity, the Idea of *Y*-ness cannot participate in itself, it is said to follow that there must be *another* Form of *Y*-ness in which it does participate. But surely there is something peculiar about this line of reasoning. The most straightforward diagnosis, (and the one which was first offered by expositors of the role of self-predication) is that, when explicitly codified, the premises underlying the TMA are simply inconsistent. Yet this has the dissatisfying consequence that the infinite regress itself is then technically extraneous, since *anything* follows from an inconsistent premise set. And this diagnosis must therefore be inadequate as a *logical* reconstruction of the TMA, since there is nothing inherently absurd about a non-terminating series of Forms, and hence it is not logically necessary that the Third Man be derived from inconsistent assumptions.

This last consideration has motivated various attempts to derive the Third Man from logically consistent principles, and thereby to yield a formally non-trivial regress argument. The goal is to derive an infinite series of Forms from mutually compatible tenets, which then conflicts with the independently held Uniqueness Principle. This strategy thereby utilizes the TMA to identify an inconsistency between UP and the premises required to derive the Third Man. In order for a genuine regress to occur, a new and distinct Form must be deducible (from a satisfiable premise set) with each iteration of the Feedback Principle. So if the TMA is to be interpreted along these lines, then the previously stated principles must have been expressed in an equivocal fashion. In tenet (2), the same symbol 'Y' occurs in both subject and predicate position, and therefore the same Idea should apparently correspond to both occurrences. But for the Third Man to arise, 'Y' in the predicate position must correspond to a new Form. In view of this, principle (2) should be more perspicuously articulated as:

(2\*) The Idea of Y-ness is (a) Y\*,

where 'Y\*' refers to the new Idea in which the members of the original set and the original Idea of Y-ness all participate. And the Feedback Principle should properly be expressed as Feedback\*:

The Idea of Y-ness and all of the objects which are (a) Y form a new class whose members all share the common property of being (a) Y\*.

This fairly generic version of the argument is in conformity with tenet (1), is expressed in unequivocal fashion, is based upon a consistent set of premises, and it is structured in such a way that an infinite proliferation of Forms is induced. But what role does self-predication play in this disambiguated argument? The Ideas of 'Y-ness' which are abstracted from the successively expanded sets must be distinct in order for the regress to occur, and this distinctness is explicitly indicated by the \*-notation. But if each Idea is distinct, then, strictly speaking, it is not the case that the Idea of Y-ness is (a) Y. Indeed, if the symbolism is used unequivocally, then by tenet (1) it is literally the case that the Idea of Y-ness is *not* (a) Y. Tenet (2\*) merely asserts that the Idea of Y-ness is (a) Y\*, and thus Y-ness is not predicated of itself, but rather the new Idea of Y\*-ness is predicated of the Idea of Y-ness, and the new Idea of Y\*\*-ness is predicated of Y\*-ness, etc.



Principle (2'), viz. that every Idea participates in itself, does not follow from (2\*), and it plays no role in this version of the argument. But (2') is by far the most plausible characterization of self-predication in Platonic terms. Indeed, (2') is *derived* from (2) via (0), and it is tenet (0) which has been identified as the basic principle underlying the semantics of predication. If predication is explicated in terms of participation, then self-predication should presumably be explicated in terms of self-participation. Thus if the Ideas do not participate in themselves, then the sense in which they are self-predicational is far from clear. The notational similarity between  $Y$ ,  $Y^*$ ,  $Y^{**}$ , etc. is a purely syntactic phenomenon, and the claim that 'the Idea of Greatness is itself great in virtue of a new Idea of Greatness' appears to employ a notion of self-predication only because the same expression is used ambivalently with respect to two different Ideas.<sup>8</sup>

In response to this criticism, it could be countered that, in order to understand the TMA, it is necessary to distinguish properties *themselves* from the Forms that account for things having such properties.<sup>9</sup> On this view, the Form of  $Y$ -ness has the same relevant property as the set of  $Y$  particulars, but it has this property in virtue of a different Form, i.e. the Form of  $Y^*$ -ness. However, to hold that the various Ideas  $Y$ ,  $Y^*$ ,  $Y^{**}$ , . . . , generated by Feedback\* all correspond to the same property is itself tantamount to the claim that there is *more than one* Idea corresponding to the property in question. And this latter claim is in turn tantamount to the *conclusion* of the TMA (i.e. the negation of UP), and hence should not be incorporated in the argument as an implicit premise. In the foregoing generic version of the regress, all that has strictly been established is the existence of an infinite series of Ideas. This may be a curious state of affairs, but so far it does not contradict any of the explicit tenets of Platonic theory. Specifically, it does not yet contradict UP, which is the target of the polemic, unless it is also the case that these Forms correspond to precisely the same property. But this additional state of affairs is not logically entailed by the premises. It is quite possible that, since the Forms involved are in fact distinct, so too are the properties to which they correspond. There is no strictly logical mechanism in place which guarantees that identity of property is preserved over iterations of Feedback\*.<sup>10</sup> All that is required is that each successive set of objects to which Abstraction is applied have *some* property in common, and that this common property is accounted for through appeal to a new Idea. But, to hold that the

new Idea corresponds to exactly the same property as in the prior step in the regress, is simply to beg the question by assuming the negation of UP.

So, in order to defend the view that the TMA is based on a distinction between a character or property itself, and the Forms that account for things having such properties, it is necessary to start with this distinction as an explicit premise, and then deduce the Third Man as a logical consequence. When the argument is recast in these terms, the first assumption will be that the property had by the *Y* particulars is exactly the same as that had by the first Idea of *Y*-ness. It is then assumed (on the basis of non-identity) that the Idea of *Y*-ness does not have this property in virtue of itself. But then, according to the semantics of predication built into the Abstraction principle, the result is simply a contradiction. Tenet (0) is naturally expressed as a biconditional, since in order to supply a satisfactory account of predication, the tenet must provide necessary and sufficient conditions for truth. Therefore, as part of the very *definition* of the Idea of *Y*-ness obtained by the first application of Abstraction, it follows that if something is (a) *Y*, then it participates in the Idea of *Y*-ness. Thus the premise that the Idea of *Y*-ness does not participate in itself yields the result that it is *not* (a) *Y*, which simply contradicts the first premise; it does not generate any new Ideas to fuel a regress. In order to deduce the Third Man, it is necessary to weaken the definition of the abstract entity whose existence is inferred *via* the Abstraction principle. The Third Man can emerge only if Abstraction is specified as yielding a semantical object such that participation in that object constitutes a sufficient but *not* a necessary condition for being (a) *Y*. This means that tenet (0) must be replaced by the following constraint on the Ideal object inferred through Abstraction:

- (0\*) if *x* participates in the Idea of *Y*-ness, then the sentence '*x* is (a) *Y*' is true.

With this weakened definition of an Idea it is possible to satisfy both the premise that the Idea of *Y*-ness is (a) *Y*, and that the Idea of *Y*-ness does not participate in itself, by concluding that the two applications of Abstraction produce Ideas which are not numerically the same. In this manner, an infinite series of distinct Forms can be deduced from the premises, but then the regress argument itself becomes somewhat superfluous. If one is willing to replace (0) by (0\*) in order to avoid a

contradiction in the premise set, then it can hardly be surprising that the conclusion of the argument will then contradict UP. However, even though the conclusion may not be a surprise, it is still the case that merely adopting (0\*) does not entail that UP is false. The positive assumption of 'self-predication' embodied in the first premise is needed to force the conclusion that there must exist (at least) two distinct Forms of *Y*-ness.

So this reconstruction of the argument seems to satisfy the desideratum of producing a multiplicity of distinct Forms of the same property, from a consistent set of initial assumptions, one of which embodies a principle of (weak) self-predication. In this sense, it may appear to constitute a viable analysis of the logic underlying the TMA. For the sake of clarity, it is worth explicitly formulating the principles involved.

Abstraction\* (incorporating (0\*)): If any set of things share a given character, then there exists a unique Form corresponding to that character; and each of these things has that character by participating in that Form.

SP\*: whenever  $\Psi$  is a Form which accounts for a number of objects sharing character *Y* (by participating in  $\Psi$ ), then  $\Psi$  itself has character *Y*.

NI\*: Whenever a Form  $\Psi$  accounts for a number of objects sharing a character *Y* (again, by participating in  $\Psi$ ), then  $\Psi$  is not identical with any of these objects.<sup>11</sup>

But, despite the appearance of successfully deriving the Third Man from compatible assumptions, I will now argue that these three tenets are in fact inconsistent, and hence even these very attenuated Platonic principles are not able to cogently engender the Third Man. As a first step towards this result, it is important to note that on the initial reconstruction of the TMA given in this section, in terms of the disambiguated Feedback\* principle, a non-terminating series of Ideas is induced from consistent premises, but at each iteration, Feedback\* applies only to sets of *finitely* many Ideas generated by prior stages of the argument. This is because the principle merely asserts that, at each stage, all of the objects under consideration share a common character; but no claim is made that the identity of this character is preserved over iterations of Feedback\*. However, if it is asserted that *all* the Forms in the denumerable series are in fact Forms of the *same* property, as in the

regress generated by the combination of Abstraction\*, SP\* and NI\*, then the situation is significantly altered. In order for this latter state of affairs to be interpreted within a Platonic framework, there must be yet *another* (omega-level) Form in which all of the infinitely many Forms generated by the first regress jointly partake. By the very same reasoning already employed, this 'limit' Form of *Y*-ness, say  $\Psi_\omega$ , will be required to account for the common character shared by each member of the series  $\Psi_1, \Psi_2, \Psi_3, \dots$ . The TMA mechanisms can then be invoked again at this level to continue the series. Since by SP\*,  $\Psi_\omega$  is still a Form of *Y*-ness, then by NI\* there is yet another Form  $\Psi_{\omega+1}$  by virtue of which  $\Psi_1, \Psi_2, \Psi_3, \dots, \Psi_\omega$  are all Forms of *Y*-ness. The claim that the Forms in the regress correspond to literally the same character therefore leads to an augmented Transfinite Third Man Argument (TTMA).

In terms of the conflict with UP, a Transfinite Third Man is essentially no worse than *two* distinct Forms of *Y*-ness. But the mechanisms of the TTMA point to an internal contradiction embodied in the attempt to construe the different Forms as corresponding to a single character. This contradiction can be demonstrated by considering *all* the Forms of *Y*-ness (however many of them there may be) together with all the *Y* particulars. This totality should be ontologically well defined, since Plato is a realist, and (in the very next passage of the *Parmenides*) he explicitly denies the contention that the Ideas are mere 'cognitions'; rather they are objective, mind-independent entities, which exist prior to the operation of Abstraction. And, in any case, the present version of the TMA is tacitly committed to the existence of this totality, in order to supply a *necessary* condition for the truth of predications, and thereby fill the lacuna induced by (0\*). On the present account, an object is *not* (a) *Y* only if it does not participate in any of the Ideas belonging to this totality, and thus it must be possible to quantify over this totality in order to provide adequate truth conditions for simple predications.

So, we will now begin where the conclusion of the TMA has ended, and consider all the Forms of *Y*-ness, together with all the *Y* particulars. On the present reading, these objects must all share the common character of being (a) *Y*, and so by Abstraction\*, there is a unique Form, say  $\Psi_\infty$ , by virtue of which they all have this character. And by SP\* it follows that  $\Psi_\infty$  is itself (a) *Y*. But then either (i)  $\Psi_\infty$  is (a) *Y* in virtue of participating in  $\Psi_\infty$  which contradicts NI\*, or else (ii)  $\Psi_\infty$  was

not a member of the original totality, and hence either it is *not* a Form of *Y*-ness, which contradicts Abstraction\*, or else the original totality did not include all Forms of *Y*-ness, which contradicts the initial hypothesis. Therefore, the triad of Abstraction\*, SP\* and NI\* is inconsistent, and the preceding demonstration constitutes a *reductio* against the attempt to construe the various Forms in the regress as corresponding to a single character or property. It is also a *reductio* against the attempt to use even the weak self-predication of SP\* to generate an infinite series of Forms.

In summary, on the first reconstruction of the TMA given in the present section, where an endless series of Ideas *Y*, *Y*\*, *Y*\*\* , . . . , is derived from consistent premises, it does not follow that the various Ideas correspond to the same property, and hence the argument does not employ (even weak) self-predication, and the conclusion does not contravene UP. On the second reconstruction, where it is assumed as a starting premise that the property had by the *Y* particulars and the (first) Idea of *Y*-ness is numerically the same, then the initial premises are inconsistent, unless the principle of Abstraction is weakened so that an object can be (a) *Y* and *not* participate in the Idea inferred on the first application. This leads to an apparent regress based on the attenuated principles of Abstraction\*, SP\*, and NI\*. It was then shown that even this set yields a contradiction, and that consistency can be restored only by denying that Abstraction\* reveals pre-existing objects. Yet Socrates explicitly rejects such a move in the very next passage of the dialogue, and it would be quite ironic to attempt to salvage the Third Man by denying the objective status of the Forms. So either the regress does not appeal to self-predication, or else the premises are inconsistent, and in neither case does the logical reconstruction meet its intended goals.

However, if principle (2') is accepted, and a clear notion of self-predication in terms of self-participation is embraced, then there is no Third Man (nor Transfinite Third Man). If the Idea of Greatness is unequivocally great, then by tenet (0), it must participate in itself. But in that case another Idea of 'Greatness' is not required, and the Feedback Principle will not lead to the postulation of any new Ideas, nor to any new set of great entities. The Idea of Greatness and the great particulars are *all* great by virtue of participation in the single Idea corresponding to the single property. In this manner, the entire procedure will lead to closure after only one application of Abstraction.

Therefore, in contradistinction to Wedberg's analysis, self-predication characterized as '... thinking of the general Idea as a particular partaking of that Idea' is in fact a strategy for *eliminating* the Third Man.

If the Ideas are construed as ontologically autonomous objects, then the Form of Greatness exists independently of the set of great particulars. This implies that the original set of *Y* objects plus the Idea of *Y*-ness do not form a *new* set from an ontological perspective. So if it is held that the Idea of *Y*-ness is itself (a) *Y*, then this Idea must already have been a member of the set of *Y* entities, even before the existence of the Idea of *Y*-ness was inferred *via* Abstraction. Of course, as the Platonic theory of Ideas is presently formulated, there is still a contradiction with (1), but there is no longer a conflict with the Uniqueness Principle. In the ensuing discussion, I will therefore advocate the solution of embracing the consistent triad of (0), (2') and UP, while simply rejecting (1). A consistent set-theoretical picture of this version of self-predication will be given in Section 5 below. But Section 4 will first briefly examine a possible motivation for this central claim that the Idea of *Y*-ness is itself (a) *Y*.

#### 4. IDENTITY AS THE LIMIT CASE OF RESEMBLANCE

The previous section has shown that if the forms are ontologically autonomous and strictly self-predicational in the sense of tenet (2'), then the regress mechanism leads to closure after one iteration, and no Third Man emerges. Thus a natural move in response to the TMA is simply to accept tenet (2') and deny (1). The ensuing discussion will give some additional reasons in favor of this move. The textual evidence used in support of (1) consists largely in the 'one over many' metaphor, and so the foregoing criticism of the dynamics of the TMA naturally invites the question whether there is a plausible way to interpret the import of this metaphor which is compatible with (2').

And this can easily be done by restricting the application of the metaphor to the relation between Universals and *particulars*. The Ideas can be interpreted as 'one over many' with respect to ordinary objects, but this alone does not preclude them from participating in themselves. By their nature, ordinary particulars cannot have participants, and hence it is impossible for *them* to participate in themselves. So whenever a property is attributed to such an object, this property must be due to another object in which the particular thing partakes. But tenet (1)

assumes that, in a similar fashion, if a Universal is assigned a property, then this property must also be due to another Idea in which it partakes, and therefore that another Idea must always be 'over' the first one, as if the latter were simply a mundane particular.

Yet, since Ideas have participants, there is no immediate reason to suppose that a distinct object must always be invoked to account for each of their properties. If the Idea of *Y*-ness has the very same property that ordinary particulars have by virtue of their participation in this Form, then why should the property not be had unequivocally by the Idea through self-participation? And on the other hand, if the Form does not have the very same property, then it does not genuinely apply to itself. Ordinary objects must always invoke a distinct entity which is 'over' them, because this is precisely what it means to be a particular. Thus there is a metaphysical constraint which makes the metaphor applicable to the relation between Forms and mundane objects, but there is not a similar constraint in the case of the relation between a Form and itself.

If this restricted interpretation is accepted, then tenet (2') is left to cohabitate peacefully with the much weaker principle:

(1'') *A mundane particular can never participate in itself.*<sup>12</sup>

Under this reading, the metaphor is preserved for the first-order case obtaining between Universals and particulars, the theory of Ideas retains its consistency, the Idea of Greatness is unequivocally great, and there is no Third Man. An additional consideration in favor of this approach is supplied by the Platonic stand concerning 'degrees-of-reality.' As mentioned earlier, Plato holds that sensible particulars belong to an ontological realm inferior to that in which the Forms reside; ordinary objects are not absolutely real, as are the Ideas, but rather possess reality only to a limited degree. This ontological inferiority is intimately tied to the contention that statements of the form '*b* is (a) *Y*' (where *b* is a particular) are never completely true, but rather are only more or less true, depending on the degree to which the transient and mutable *b* participates in or resembles the perfect and eternal Idea.

If, as many have advocated,<sup>13</sup> the Ideas are construed as ultimate exemplars or paradigms (which intuitively is a very compelling version of property realism), and if predications are thus construed as relation-to-standard statements which are true to varying degrees, then it is quite natural to accept (2') and reject (1). Under this view, the Idea



of *Y*-ness is the ideal paradigm of *Y*-hood, and some object *x* is (a) *Y* only to the extent that it resembles this ideal standard in the relevant respect. If *x* is a sensible particular such as *b*, then the degree to which it resembles or participates in the perfect Idea of *Y*-ness will perforce be quite limited. However, if *x* is the Idea of *Y*-ness itself, then *x* will resemble *Y* to the highest degree possible. Identity is the limit case of resemblance, and so if the Ideas are conceived as paradigms, and predications are interpreted as relation-to-standard statements, then this supplies a *prima facie* motivation for thinking that the Ideas must be self-predicational to the strongest degree possible.<sup>14</sup>

An ideal paradigm version of property realism is thus able to provide a natural reading of tenet (2), under which it can be seen as an integral part of the theory of predication, rather than as a doctrinal foible leading to confusion and inconsistency. The Idea of *Y*-ness is superlatively and perfectly (a) *Y*, simply because it *is* the Idea of *Y*-ness, and anything which is not an Idea must necessarily resemble it to a less perfect degree. This is made entirely graphic by recasting 'participation' in (0) in terms of 'degree of resemblance':

- (0') The sentence '*x* is (a) *Y*' is true to the degree to which *x* resembles the Idea of *Y*-ness.

Given this reading, tenet (2) is no longer an independent claim, but rather follows as a direct consequence of (0'), since it is an obvious truth that every object resembles itself. However, this does not reduce the instances of principle (2) to vacuous self-identity statements. The sentence 'The Idea of *Y*-ness is (a) *Y*' has degree of truth equal to unity because the Idea of *Y*-ness is identical to itself, but not every self-identical object is thereby rendered self-applicable. Self-application is an additional claim which follows from self-identity only for a privileged set of entities, and only under very specific metaphysical and semantical assumptions.

It is also important to note that perfect resemblance is not used here to *define* the relation of identity; rather identity is taken as an intuitive primitive and used to place an upper bound on resemblance, in order to account for the central notion that the Idea of *Y*-ness is superlatively (a) *Y*. And this is not to say that identity is *required* for the unqualified truth of a predication, but only that identity (in certain cases) is sufficient to *guarantee* a true self-predication. For example, 'the Idea of Beauty is eternal' is a true statement, even though the Idea of Beauty



is not the same as the Idea of Eternality. An Idea may resemble another one perfectly (in the relevant respect), while a sensible particular never can. Thus mundane snow is never perfectly white, while the Idea of Whiteness *is* perfectly eternal. The point of considering identity in the light of the resemblance version of participation is simply that it can supply a conceptual foundation for certain instances of the claim that the Idea of *Y*-ness is perfectly (a) *Y*.

The resemblance interpretation enshrined in (0') therefore serves to provide a principled motivation for self-predicational claims. However, it is not without serious difficulties of its own, two of which will be briefly mentioned. First, tenet (0') does not have unrestricted scope, and thus tenet (2) is not universally true. As is well known, the Forms, *qua* ideal objects, are required to have certain general properties, and so any predicate which contradicts a general property enjoined by the ontological status of the Forms, will make reference to an Idea which cannot be consistently predicated of itself. As above, the Ideas are held to be eternal, while for any particular *b* it is the case that *b* is transient. This latter claim leads to the postulation of the Idea of Transience. But since *all* Ideas are eternal, it is false that

The Idea of Transience is transient

and this serves as a counterexample to (2). Similarly, the ideas of Mutability, Imperfection, Spatial Extension, etc., will not be self-predicational, and hence they will apparently not be ideal paradigms, either. This indicates that the intuitive content furnished by the notion of an ultimate standard or exemplar is not sufficient to serve as the basis for an unrestricted theory of predication. In the case of properties of particulars which contradict the ideal status of the Forms, a much more obscure account of 'participation' will be required. But this type of counter-instance does not eradicate the many clear cases of positive self-application, and it is the positive cases with which the TMA is concerned.

A second major difficulty involves the nature of the relation of 'resemblance' appealed to in the foregoing account of predication. This relation between an Idea, and an object of which it is predicated, cannot itself be explicated in terms of the theory of Ideas, because this will immediately induce the 'resemblance regress,' which is also considered in the *Parmenides*. If some object *x* is (a) *Y*, then according to the preceding account, it is true that *x* resembles the Idea of *Y*-ness. But

if this statement is in turn explicated in the same manner, then there must apparently be some further Idea in virtue of which the predication of resemblance can be explicated. And, in turn, the object  $x$  must bear the relation of resemblance to this new Idea, which will again require the mediation of yet another Idea, *ad infinitum*.

The resemblance regress, though structurally similar to the original Third Man, is not caused by self-predication. Instead, it is a far more general criticism, which is applicable to any attempt to explain the relation between an Idea, and an object of which it is predicated, in terms of another Idea. This regress shows that the relation of 'resemblance,' or any other proffered account of the relation between a Universal and a participant, must either be taken as a primitive at some level, or else it must be explained through some other semantical apparatus: it cannot itself be explained through appeal to the 'base level' account of predication.<sup>15</sup> However, the purpose of the present paper is not to attempt a general defense of the theory of Ideas, but rather to show that, whatever problems the theory may possess, the notion of self-predication *per se* is not one of them.

##### 5. PLATONIC UNIVERSALS AS NON-WELL-FOUNDED SETS

If the foregoing considerations are convincing, and the Third Man can arise only insofar as the Forms are not genuinely self-predicational, then why has some version of principle (2\*), wherein an equivocally designated new Form is introduced at each cycle of the argument, been so widely embraced by commentators on the TMA, including those who wish to defend the consistency of Plato's theoretical tenets? Why has tenet (2') not been retained as the most clear and compelling version of self-predication? I think that this is at least partially due to the fact that contemporary philosophers may tend to view tenet (2') from the perspective of standard set-theoretical semantics for predication. Under this technically well-defined and pervasive modern view, predications are interpreted as implicit set membership claims, so that ' $b$  is beautiful' is interpreted in the semantical metalanguage as asserting that the object denoted by ' $b$ ' is a member of the set which the interpretation function assigns to the expression 'is beautiful.' In this manner, the extension of the Idea of Beauty can be roughly equated with the set of beautiful objects, and the 'one over many' metaphor can serve to characterize the relation between a set and its members.

But then Russell's paradox immediately discredits any notion of direct self-predication. If the Idea of Beauty is indeed beautiful, and it is represented extensionally as the set of beautiful objects, then the set of beautiful objects must be a member of itself. And it is precisely this notion of self-membership which has been historically identified as the illicit principle leading to inconsistency in the foundations of mathematics. Various familiar strategies, from Russell's type theory to the cumulative hierarchy of sets, have been devised expressly to prohibit the formal realisation of the notion of self-application. Given this historical and conceptual heritage, it is then quite easy to see why tenet (2') should be viewed as egregious. The TMA purports to derive a contradiction within Plato's theory, and the mechanisms of the argument call upon a principle which is a notorious source of paradox in the foundations of mathematics. So it is extremely natural to perceive tenet (2') as fodder for a *reductio*, and to believe that any theory adopting this principle will be inconsistent. Hence even those who wish to absolve Plato from the charge of inconsistency are inclined to eliminate (2') and adopt (2\*) in its place, even at the cost of sustaining the Third Man, since a potentially benign regress is better than an outright absurdity.

Yet it is possible both to preserve (2') *and* logical consistency, while simultaneously dissolving the Third Man. Russell and others have stigmatized self-membership as a fundamentally flawed notion, but it is now well known that consistent systems of set theory which condone such non-well-founded entities can be constructed. Aczel<sup>16</sup> has explored a variation of ZFC set theory (Zermelo-Fraenkel plus the axiom of choice) in which the standard foundation axiom is replaced by the anti-foundation axiom (AFA). The AFA admits non-well-founded sets into the abstract universe, and thus dramatically expands post-Russellian sensibilities concerning the bounds of the coherent. For example, the AFA sanctions the set  $\Omega$ , such that

$$\Omega = \{\Omega\}.$$

This object induces an infinite descending chain of membership, but it is nonetheless hereditarily finite, since each member of the chain has only one element. ZFC, with the axiom of foundation replaced by the AFA (ZFC/AFA), is provably consistent relative to the original system. Thus circularity is formally absolved (while unrestricted Abstraction, *a la* Frege, is not), and the world of 'hypersets' is rendered just as axiomatically secure as the cumulative hierarchy.

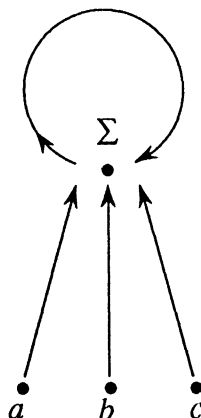


Figure 1.

Utilizing this liberated conception of hypersets, the self-predicational aspect of Plato's Ideas can be modeled quite nicely in terms of non-well-foundedness.<sup>17</sup> To illustrate this, let the Idea of *Y*-ness be modeled extensionally by the set  $\Sigma$ , such that for all objects  $x$ ,

if  $x$  is (a) *Y*, then  $x \in \Sigma$ .

Thus for all particular objects  $a, b, c, \dots$  such that the statement ' $b$  is (a) *Y*' is true (to some sufficient degree), it will be the case that  $\Sigma = \{a, b, c, \dots\}$ . If, in addition, the Idea of *Y*-ness can be used to instantiate (2), and hence if it is true that the Idea of *Y*-ness is itself (a) *Y*, then it will be the case that  $\Sigma \in \Sigma$ . In this manner,  $\Sigma$ , which is the formal representative of the Idea of *Y*-ness, will be such that

$$\Sigma = \{a, b, c, \dots, \Sigma\}$$

The hyperset  $\Sigma$  can be depicted by the labeled graph in Figure 1 above, where the arrows represent the membership relation, and where the particulars which participate in the Idea of *Y*-ness are the atoms corresponding to the terminal nodes. Under this scheme, ' $b$  is (a) *Y*' and 'the Idea of *Y*-ness is (a) *Y*' are both unequivocally true. The Idea of *Y*-ness is true of itself since it is a member of the set of entities which the interpretation function assigns to the predicate in the expression 'is (a) *Y*'. In this manner, Platonic self-predication is given a set-theoretical model, and tenet (2') is formally vindicated.<sup>18</sup>

If the Ideas are represented extensionally as hypersets, then the Third Man regress is deflated exactly as suggested in Section 3. Plato's relation of 'participation' is explicitly recast in terms of set membership, and on this reading, the great particulars and the Idea of Greatness are all great in exactly the same way. No 'new' Ideas of Greatness are generated, and the regress mechanism leads to closure after one step.<sup>19</sup> In this manner, the Forms are seen as truly *one* over many, rather than as one *immediately* over many. The singularity and uniqueness of each Form (exemplifying a particular property) is exactly the principle which Parmenides' argument is designed to refute, and thus it is presumably the principle which Plato wished to defend. So a possible solution to the Third Man is to retain the motivating principle of self-application, expressed in terms of (2') and (0), as well as the target of the polemic, *viz.*, the singularity of the Forms, and simply reject the overly restrictive (1). Therefore, if the relation of 'participation' is modeled in terms of set membership, then Platonic Universals can be modeled as non-well-founded sets, and the 'one over many' metaphor becomes technically compatible both with self-predication and with uniqueness.

## NOTES

<sup>1</sup> Vlastos, G., 'The Third Man Argument in the *Parmenides*', *Philosophical Review*, vol 63, (1954), pp. 319–349.

<sup>2</sup> Wedberg, A., 'The Theory of Ideas', Chapter 5 of *Plato's Philosophy of Mathematics*, Almqvist and Wiksell, Stockholm, 1955. Cited passages are from reprinted version appearing in G. Vlastos (ed), *Plato I*, University of Notre Dame Press, 1971.

<sup>3</sup> As, for example, Sellars, W., 'Vlastos and 'The Third Man'', *Philosophical Review*, vol 63, (1954), pp. 319–349; Strang, C., 'Plato and the Third Man', *Proceedings of the Aristotelian Society*, supp. vol XXXVII (1963), pp. 147–163; Cohen S. M., 'The Logic of the Third Man', *Philosophical Review*, vol 80 (1971), pp. 448–475.

<sup>4</sup> The translation used here is from B. Jowett, *The Dialogues of Plato*, Fourth Edition, Oxford University Press, 1953.

<sup>5</sup> *Op. cit.*, p.43.

<sup>6</sup> *Op. cit.*

<sup>7</sup> *Op. cit.*

<sup>8</sup> It is important to note that this point is not in conflict with the notion that the *specific* 'Y-nesses' realized in different objects must be numerically distinct. Rather, the claim is that different objects can possess numerically distinct instances of the *same* general property only by virtue of the same Idea. If different Ideas are invoked, then specific instances of *different* general properties would seem to be at issue. It is also important to note that this point is not affected by the distinction which is often drawn between *being* and *having* a certain property. If it is claimed that 'is' is not to be interpreted

univocally, and that particular objects merely have *Y*-ness while the respective Idea is *Y*-ness, then there is no immediate notion of self-predication and no threat of a regress. Self-predication and the Third Man enter the picture only if it is claimed that in addition to being *Y*-ness, the respective Idea also has *Y*-ness. And in this case, the present point becomes operative: the Idea of *Y*-ness should *have* the property of *Y*-ness by virtue of self-participation. Thus even on this reading which distinguishes between being and having, the 'is' of having should still be interpreted univocally.

<sup>9</sup> This is noted by Vlastos in 'Plato's 'Third Man' Argument (*Parm.* 132A1–B2): Text and Logic', *Philosophical Quarterly* 19 (1969), 289–301, [TMAII], note 15. I would like to thank an anonymous referee for bringing this to my attention.

<sup>10</sup> And this proposed distinction immediately invites the question 'On what theoretical basis is the relevant property said to be the *same*?' Sameness of Form provides a very clear criterion for sameness of property, especially if, on a realist model, having a property is ontologically grounded in terms of participating in an autonomous abstract object. But if things can have the same property in virtue of different Forms, then the Forms are reduced to the status of mere intermediaries, and are distinguishable solely on the basis of the things which participate in them. They do not then provide a foundation for semantics or cognition, since their primary role is determined by being Forms of the same *property*. Similarly, it is difficult to say what 'self-predication' amounts to, without a theoretical criterion in virtue of which the property being attributed is in fact the same. Since it is now properties rather than Forms which determine when the same character is being predicated of an object, genuine self-predication in this context would seem to require that the *property* of *Y*-ness is itself (a) *Y*.

<sup>11</sup> This reconstruction is articulated by Vlastos in TMAII.

<sup>12</sup> This is Vlastos' 'weak' non-identity principle, where the 'strong' version is tantamount to principle (1'). It is uncontroversial that Plato *at least* held (1'). Vlastos argues that Plato is forced to embrace (1') as well, because of his stand on 'degrees-of-reality', in combination with the univocality of 'is'. Vlastos argues that the perfect reality of the Forms is incompatible with their being the imperfect predicates of particulars, and thus there is a crucial 'Separation' between *Y*-ness itself and the salient property of *Y* particulars. But it is then unclear how Separation could contribute to the TMA. If the properties had by Ideas and particulars are systematically distinct, and this distinctness is made syntactically explicit, then it is false that any particular *b* is (a) *Y*\*. The disambiguated version of Parmenides' argument asserts that the Idea of Greatness and the great particulars are all great\*, but according to Separation this must be false, because Greatness\* applies to Ideas and *not* particulars. Therefore, the initial application of the principle Feedback\* in the regress mechanism is blocked. And furthermore, since Greatness\* applies to Ideas, Separation supplies no reason for supposing that Greatness\* cannot be unequivocally true of itself, in which case the regress mechanism would lead to closure after two applications of abstraction.

<sup>13</sup> See Geach, P., 'The Third Man Again', *Philosophical Review*, vol 55, (1956), pp. 72–82, for an early articulation of this view.

<sup>14</sup> Of course, there are contrasting interpretations of the texts which deny this, and I only wish to advocate the present view as a plausible reading which coheres with strong self-predication.

<sup>15</sup> This point is made in Passmore, J. A., *Philosophical Reasoning*, London, 1961, Chapter 2.

<sup>16</sup> Aczel, P., *Non-Well-Founded Sets*, Center for the Study of Language and Information, Stanford University, Lecture Notes Number 14, 1988.

<sup>17</sup> In Barwise, J. and J. Etchemendy, *The Liar*, Oxford University Press, 1987, the related, though semantically distinct phenomenon of *self-reference* is accounted for in terms of non-well-founded sets.

<sup>18</sup> This is not to say that every set definable in ZFC/AFA will correspond to some self-predicational Idea. Clearly not all sets in ZFC/AFA are non-well-founded, and not all Ideas are self-predicational. ZFC/AFA does still preserve the distinction between set and class, but there is now a *class* of sets which are not well-founded, in addition to those which are. This means that the Platonic Forms which are consistently self-predicational (i.e. which can serve to instantiate principle (2)) can be formally interpreted by certain members of the class of non-well-founded sets. That is, in the formal semantics, we can now define particular structures within a consistent set theory which can model self-predication via the set-theoretical trait of self-membership. This is not possible in classical set theory, precisely because it precludes non-well-founded objects. It is still possible to define a parameterized version of the Russell set in ZFC/AFA; let  $\Sigma_a = \{x \in a : \neg(x \in x)\}$ . This definition entails that  $\neg(\Sigma_a \in a)$  and hence  $\neg(\Sigma_a \in \Sigma_a)$ . From this it follows that the set  $\Sigma_a$  does not model any self-predicational Idea, and therefore  $\Sigma_a$  would not serve as a model theoretic analogue of, for example, the Platonic Form of Beauty.

<sup>19</sup> This interpretation does not capture the notions of resemblance to an Ideal paradigm and degrees of truth, since participation is interpreted directly as (non-fuzzy) set membership. It is perhaps technically possible to incorporate the notion of degrees of truth into a system of non-well-founded sets, but the present paper is concerned primarily with the explication of *self-predicational* claims, where such statements should have a degree of truth equal to unity. The main purpose for considering the Ideal paradigm reading is simply to show that there are additional reasons for taking self-predicational claims seriously.

Manuscript submitted 8 July 1992

Final version received 25 June 1993

University of Edinburgh  
Centre for Cognitive Science  
2 Buccleuch Place  
EH8 9LW Edinburgh, Scotland  
Great Britain