Some 70 years ago, an exchange between M.B. Foster and J.D. Mabbott brought into focus a set of intriguing and important questions about Socrates’ defense of justice in the *Republic*.¹ These questions set the framework for much of the subsequent scholarly debate concerning the argument of the *Republic*; and while there has been a sharpening of the issues involved, scholars seem to be no closer to agreement on answers to the questions. One of the questions concerns the relationship between justice and happiness: (a₁) does Plato view happiness as a consequence of justice, or (a₂) does he see a closer relationship between the two, namely, that the two are related in such a way that showing that justice contributes to happiness is a way of showing that it is desirable for its own sake? A second, related, question is about Plato’s understanding of the claim that justice is valuable “for its own sake”: (b₁) does he mean that it is valuable apart from any desirable consequences it might have, or (b₂) that it is valuable insofar as it directly produces certain desirable effects or consequences? A third question has to do with the structure of the argument in Books II-IX: (c₁) is there one continuous argument designed to show that justice is valuable for itself by showing that it yields happiness, or (c₂) is there a two-part argument, the first part aimed at showing how justice is valuable for itself and the second at showing that it yields happiness (as a consequence)? How we answer these questions is of crucial importance for our understanding of why Plato thinks justice is worth pursuing and what sort of value he thinks it has. Finding the right answers is no easy task because of the mutual entanglement of the questions: attempting to answer one inevitably leads to the others, and it is easy to lose one’s way.

Scholars in recent years have staked out three main interpretations of Plato’s position on the value of justice based on their answers to the three questions.

I. Happiness as consequence of justice

(a₁) Plato views happiness as a consequence of justice.
(b₁) By “justice is valuable for its own sake” Plato means

¹ See Foster (1937), Mabbott (1937), and Foster (1938).
that it is valuable apart from any desirable consequences it might have, i.e. it is intrinsically valuable.²

c2) The Republic offers a two-part argument for the value of justice, the first part aimed at showing how it is valuable for itself and the second at showing that it yields happiness (as a consequence).³

II. Justice as constituent of happiness

(a2) Plato sees a particularly close relationship between justice and happiness (i.e. justice is a constituent of happiness) such that showing that justice contributes to happiness is a way of showing that it is valuable for its own sake.

(b1) By “justice is valuable for its own sake” Plato means that it is intrinsically valuable.

(c1) There is one continuous argument in Books II-IX designed to show that justice is valuable for itself by showing that it yields happiness.⁴

III. Happiness as direct effect of justice

(a2) Plato sees a particularly close relationship between justice and happiness (i.e. happiness is a direct effect of

² By “intrinsically valuable” I simply mean non-instrumentally valuable, i.e. X is “intrinsically valuable” if it has value that is not derived from beneficial effects it may produce. “Intrinsic value” might also be contrasted with “extrinsic value”: something might have value apart from any benefits it produces, but it might derive this value not from itself but from something else; for example, justice might be non-instrumentally valuable, but derive this value from the Form of the Good (see Aristotle, Nicomachean Ethics I 6, 1096b8-13; for helpful discussion of different kinds of value, see Korsgaard (1983)). Another point about terminology: in his tripartite classification of goods, Glaucon does not use a term like “valuable”—he speaks of “things we welcome” or “things we are glad to have” for their own sake or for their consequences (357b4-d2; cf. Aristotle’s use of “choiceworthy” [ὡ[:,:,]ι隳分会ς, e.g., Nicomachean Ethics I 7, 1097b30-1098a21); but since he clearly assumes that these things have genuine value, there is no harm in using the term “valuable” in describing his classification. It is worth keeping in mind, though, that all three types of goods are regarded as things that we can “choose” or “have” (357b5, 8, c8). If there are goods that we cannot be said to choose or have, e.g. Forms, they will fall outside of Glaucon’s classification.

³ See, e.g., Annas (1981), 168-9, 294 (cf. 318 for some second thoughts). Mabbott (1937) clearly sides with the first interpretation at certain points (e.g. 62), but he also seems to approve of (a2), which is incompatible with the first interpretation (63); cf. White (1984), 396, n.6.