

VIII. Plato and the "Good Conscience": *Laws* 863E5–864B7

MICHAEL O'BRIEN

YALE UNIVERSITY

The ninth book of Plato's *Laws* has a long digression on the sources and types of wrongdoing. It starts in 860C, and it is meant to reconcile the Platonic doctrine that all injustice is involuntary with the practical need, in law, to distinguish between voluntary and involuntary crimes. The digression ends by summarizing, in a difficult passage (863E5–864B7), the argument of the preceding pages.¹

These lines are commonly taken to mean that justice consists in conforming to a purely internal standard, the voice of conscience. For example, Ritter finds in the passage that there are three sources of ἀμαρτήματα, viz. θυμός, ἡδονή, and ἄγνοια;² but of the three types of errors (ἀμαρτήματα), only those due to anger or pleasure are unjust (ἀδικήματα). An error caused by ignorance is never unjust. This is the meaning which Ritter finds in 864A1–8. That passage says that as long as the δόξα τοῦ ἀρίστου rules in a man's soul any action he performs is just. Ritter explains δόξα τοῦ ἀρίστου as "die ruhige Überlegung über das, was zu thun Recht und Pflicht sei" (283). Such a δόξα can be of

¹ I am obliged to Professors H. Cherniss, A. E. Raubitschek, and W. J. Oates for advice on several points in this article, and for certain references. The opinions given, of course, are my responsibility.

The following books and articles deal with the passage in question, and references to them are hereafter abbreviated: O. Apelt, *Platonische Aufsätze* (Leipzig 1912), and *Platons Gesetze* (Leipzig 1916); C. Badham, "Platonis De Legibus Liber Nonus," *Mnemosyne* 2nd ser. 12 (1884) 44–56; E. B. England, *The Laws of Plato* (Manchester 1921); L. Gernet, *Platon: Lois: Livre IX: Traduction et Commentaire* (Paris 1917); G. Grote, *Plato and the Other Companions of Sokrates*, vol. 2 (London 1885); G. M. A. Grube, *Plato's Thought* (London 1935); R. Hackforth, "Moral Evil and Ignorance in Plato's Ethics," *CQ* 40 (1946) 118–20; R. B. Levinson, "Moral Obligation in Platonic Ethics": 40th Meeting of the American Philosophical Association: résumé in *Journal of Philosophy* 37 (1940) 677–78, and *In Defense of Plato* (Cambridge, Mass. 1953); G. Müller, *Studien zu den platonischen Nomoi* (München 1951); D. A. Rees, "Bipartition of the Soul in the Early Academy," *JHS* 77 (1957) 112–18; C. Ritter, *Platos Gesetze: Kommentar zum griechischen Text* (Leipzig 1896); G. Stallbaum, *Platonis Leges et Epinomis* (Gothae et Erfordiae 1859–60); A. E. Taylor, *The Laws of Plato* (London 1934).

² Ritter 282–83. It is clear that 864B1–7 is a repetition (as δεδήλωται indicates) of 863B1–C2. The phrase ἐλπίδων δὲ καὶ δόξης τῆς ἀληθοῦς περὶ τὸ ἀρίστον ἐφεσις, then, must mean ἄγνοια, which is the name of the third cause of error in the earlier passage. To support this meaning, the phrase should be emended, according to Ritter. For his suggested emendations see note 15 *infra*.

two kinds, he adds: (1) ἀληθὲς δόξα (for which, of course, ἐπιστήμη can stand); (2) an erroneous δόξα τοῦ ἀρίστου, which is the equivalent of ἄγνοια. That a δόξα τοῦ ἀρίστου can be false Ritter sees implied in the phrase κἂν σφάλληται τι (864A4), which refers, he says, to "eine irrigte Vorstellung." He concludes that if justice consists in following one's own judgment, whether true or false, of the best, then justice, in our terms, is simply obedience to the "voice of conscience" (283–84).

Ritter finds the whole passage unclear, however, and implies that it is inconsistent with Plato's usual doctrine. This passage states, he says, that the worst kind of ἄγνοια can cause only βλάβαι, not ἀδικίαι. Yet it is also Platonic doctrine that the rule of θυμός and ἡδονή in the soul — which does cause ἀδικίαι — is possible only as a result of ἄγνοια concerning the true conditions of happiness (Ritter 286). In its main points, Ritter's interpretation of the passage is the common one.³

³ Stallbaum had already said much the same thing. He translated κἂν σφάλληται τι as "etiāmsi in quāpiam re a recto aberret" and added: "Itaque philosophus censet iustitiam in eo cerni, ut certi opinionem sequamur, eoque nomine illam ne tum quidem indignam videri, si quando iudicii errore a vero rectoque declinatum sit" (Stallbaum 3.53). Like Ritter, Stallbaum found the doctrine surprising in the light of other passages (he cites *Rep.* 440A), but he thought the disparity could be explained by saying that *Laws* 863E ff. is about justice in the ordinary man (3.53–54).

Grote, who reads τοῦτο for τούτων in 864A2, explains 863E–864A by saying that if a man's θυμός and ἡδονή "are subordinated to the opinion which he entertains about what is good and right — then everything which he does is just, even though he falls into error" (Grote 4.367). Grote finds that not all ἀμαρτήματα are ἀδικήματα, but Plato's "reasoning is not clear" (4.368), and the distinction is "perplexing" (4.369). But it means that if a man is "sincerely persuaded that he is acting for the best, he cannot commit injustice" (4.368).

Apelt refers to "unschuldigen Irrtume, der im ehrlichen Glauben (bona fides) das Beste zu wollen, fehlgeht — ein Fall, den Platon (864 A) fast mehr lobend als tadelnd beurteilt," and he calls this one type of action "was der durch irgendwelche Seelenbeschaffenheit getrübt, also irrende oder unwissende Verstand tut . . ." (*Platonische Aufsätze* 191).

Gernet translates τὴν δὲ τοῦ ἀρίστου δόξαν κτλ. as follows: "D'autre part, il faut appeler justes, dans une cité ou chez un individu, tous les actes conformes à l'opinion qu'ils ont du bien, quel qu'en soit l'objet et fût-elle erronée . . ." (36). As the translation "fût-elle erronée" (κἂν σφάλληται τι) shows, and as the commentary explains (109–10), ἡ τοῦ ἀρίστου δόξα can be in error, yet all acts which conform to it are just, in Gernet's view.

England, who reads τοῦτο γ' for τούτων in 864A2, takes 864A1 ff. to mean that "if the will to choose the best course — what we should call *conscience* — has supreme sway, all action so regulated is *right* . . ." (403, italics his). He argues, however, that κἂν σφάλληται τι means "even in case some damage is done" and connects the phrase with τοιαύτην βλάβην (864A7).

A. E. Taylor, who also reads τοῦτο γ' for τούτων, translates as follows: "But where there is the conviction that a course is *best* — wherever a society or private individuals may take that best to lie — where that conviction prevails in the soul and governs a

If the usual interpretation of the passage is right, and Plato here calls ignorance and injustice separate causes of error, then he has departed from his doctrine of the *Republic*, where all injustice is attended by ignorance (444A1–2) and ignorance is a form of injustice (444B6–8). Ritter is right in pointing out that there is a radical discrepancy between the apparent meaning of *Laws* 863E–864A and the rest of Platonic doctrine. There is no doubt, as he and others have complained, that the whole passage is very involved. I think that, in spite of this, it may be shown that the usual interpretation is wrong, and that Plato did not adopt the purely internal standard of "the voice of conscience" as his final criterion of justice.

The reason for the whole discourse about causes of crime is found near

man's conduct, even if unfortunate consequences should arise, all that is done from such a principle, and all obedience of individuals to it, must be pronounced *right* . . ." (252–53, italics his).

Grube, who keeps the manuscript reading *τούτων*, translates as follows: "But when a city or some individuals have any belief as to what is best, wherever it be among those emotions that they think to find it, and if that belief rules in their soul and regulates the whole man, then, even if they are mistaken to some extent, any action so performed is right" (228). Plato, Grube explains, had earlier distinguished unintentional injuries (*βλάβαι*) from intentional injuries (*ἀμαρτήματα*). The latter can be due to temper, pleasure, or ignorance. Only temper and pleasure, however, can cause *ἀδικία*. Plato here "emphatically asserts that the other class of errors, those due to ignorance, should never be called sin, *ἄδικον*, at all, though they do the greatest damage (863E). Whatever is done by a disciplined soul, under the influence of ignorance and wrong opinion, is not a sin at all" (229).

Levinson finds that "only in the Ninth book of the *Laws* (860c ff.) does Plato assert unequivocally that an ignorant and objectively wrong act, indeed 'a great and savage wrong,' [this apparently refers to 863c7] may be a just act, if only it proceed from an agent acting under the control of reason and in the light of his 'opinion of the best'" (*Journal of Philosophy* 37 [1940] 677–78). Levinson repeats this judgment and finds in the *Laws* passage a "certain noble inconsistency" with earlier doctrine, in his book *In Defense of Plato* (526–27).

Hackforth does not think that in this passage errors due to *ἄγνοια* are meant to be classified as *ἀδικία*. Plato "even goes so far as to say that any action done out of a conviction that it is good is just (*δίκαιον*) and 'best for the whole life of mankind' (864 A)" (*CQ* 40 [1946] 119).

Gerhard Müller finds that the possibility of *ἄγνοια* is implied in the phrase *ἡ τοῦ ἀρίστου δόξα*. The clause which follows the phrase he translates: "wie dies (*τοῦτο* für *τούτων* = τὸ ἀρίστον Cousin) auch nach dem Glauben einer Stadt oder einiger Einzelner sei" (58, note 1). But, he says, this *ἄγνοια*, which is the third type of *ἀμαρτήματα*, is merely ignorance of a technical character; "aus ihr ergeben sich Versehen und alles, was man landläufig *ἀκούσιος ἀδικία* nennt (864a), was aber Platon garnicht als Unrecht, sondern als Gesetzesübertretung (*ἀμάρτημα*) rechnet" (58).

More recently, D. A. Rees, who accepts Taylor's translation, finds in *Laws* 863E–864A an "internal" definition of justice," applied here to the individual soul, and adds: "Similar views have been adopted in recent centuries, as by Kant and T. H. Green (*cf.* also Rousseau)" (*JHS* 77 [1957] 115).

its beginning (861B1 ff.): the fact that in legal theory there is a customary distinction between voluntary and involuntary acts of injustice, whereas in Platonic moral theory all unjust acts are involuntary. The Athenian Stranger maintains that there is a distinction but that it is different from the customary one (861C2–3). His solution is to divide injuries into two aspects: (1) the injury proper, (2) the injustice, if any.⁴ The lawgiver must remedy the injury proper by compensation and by restoring amity between injured and injurer (862A–C). Where injustice exists, the lawgiver must cure the guilty person. He must not only be made to pay for the injury but also taught and compelled never or hardly ever to repeat his offense. Incurables should be executed (862C–E). Briefly, mere injuries (βλάβαι) are external and accidental; acts of injustice require that the offender be taught and compelled to change his ways.

Clinias then asks for a clearer explanation of the distinction between injustice and injury and of their confusion with the other pair of terms, “voluntary” and “involuntary” (863A3–6). It is in reply to this request that the Athenian develops the argument that follows; his argument, then, must refer to these earlier distinctions. He describes three causes of wrongs (ἀμαρτήματα) (863C1): anger, pleasure, and ignorance. Plato does not say whether ἀμαρτήματα means mere injuries or injustices or something else. There are several possible interpretations. The first is most clearly stated by Grube: that ἀμαρτήματα are intentional injuries, but of the three types of intentional injury wrongs due to anger and pleasure are injustices and those due to ignorance are mere injuries.⁵ This view would imply, however, that the whole discussion between 863A3 and 864A8 is about the distinction between types of intentional injuries, two of which are unjust, and the third just. Thus:

- | | |
|--------------------------|---------------------------------------|
| (1) unintentional injury | always just |
| (2) intentional injury | (a) unjust, if from anger or pleasure |
| | (b) just, if from ignorance |

But it is clear from the way in which the discussion begins and ends (863A3–6, 864A6–8) that it is an explanation of the distinctions already drawn between injury and injustice, and between voluntary and involuntary.⁶ What we must expect, therefore, at the end of this explanation

⁴ Injustice is not merely a subdivision of injury. Injustice need not produce injury (863E8); it may produce benefit (862A7–8, C6–7). But the main line of the argument concerns injustice only as one aspect of injury.

⁵ See note 3 *supra*.

⁶ This is confirmed by the verbal reminiscence *σαφῶς ἀν διορισαίμην οὐδὲν ποικίλλων* (863E6), which recalls *ἥδιον δ' ἂν ἔτι σαφέστερον ἀκουσαίμην . . . ὥς ἐν τοῦτοις διαπεποίκιλται* (863A4–6).

is a restatement of these distinctions, not a statement of a difference between types of intentional injuries. The end of the discussion confirms this by concluding, in 863E5-864A8, that there are two conditions which may cause offenses at law: (1) tyranny by the soul's lower impulses, (2) a condition which is the source of what is commonly called an "involuntary injustice" (864A7). Now it was the common use of this last phrase which provoked the whole digression (861B4-5)⁷ and for which Plato substituted the notion of "mere injury." The discussion ends, therefore, by reaffirming the division of the aspects of crime into injury and injustice. The phrase *κὰν σφάλληται τι* ("even if there is a slipup") must therefore bear the implication "even if some damage be done," as England argues, and not "even if some error of judgment be made."

One might still argue that 863A-864A is about the difference between crimes of anger and pleasure on the one hand and crimes of ignorance on the other, but that this is a restatement of the earlier distinction between injustice and injury. This appears to be Ritter's view (286). On this assumption, offenses due to ignorance are mere injuries and offenses due to anger and pleasure are injustices. But against this view there are objections. In the first place, if 864A1-8 refers to ignorance, it will be hard to explain why Plato calls every action performed in such a state of mind *ἐπὶ τὸν ἅπαντα ἀνθρώπων βίον ἄριστον* (864A6), especially when he has just said that ignorance induces us to act against our own wish (863E2-3). In the second place, mere injury is a purely external accident, for which only an external penalty, compensation or ritual purification, is prescribed. One example is the accidental killing of a man in the public games (865A).⁸ But ignorance is a radical fault of soul. The law's duty is to cure other faults of soul, contentiousness and love of pleasure (862D4-E1). It can be no less the law's duty to cure ignorance, and, if so, offenses due to ignorance will not be treated as mere injuries.

We must, therefore, take each of the three causes of *ἁμαρτήματα*, viz. anger, pleasure, and ignorance, as expressions of injustice, not the first two alone. This interpretation is confirmed by the lines which immediately precede Plato's final summary of the whole discussion. I refer to 863D6-E4, in which the argument is as follows. Men are said to get the better of pleasure and anger, or to be bested by them, but these terms are

⁷ Cf. also 862A2-7, where even the phrasing corresponds closely to 864A7-8 and shows that both passages make the same point. *ἀκούσιον ἀδίκημα . . . τὴν τοιαύτην βλάβην* is echoed in *ἀκούσιον ἀδικίαν . . . τὴν τοιαύτην βλάβην*.

⁸ In cases of accidental wounding the cause is explicitly said to be chance, not ignorance: *ὅς δ' ἂν ἄκων ἄλλος ἄλλον τρώσῃ, τὸ βλάβος ἀπλοῦν ἀποτινέτω—τύχης γὰρ νομοθέτης οὐδεὶς ἱκανὸς ἄρχειν . . .* (879B1-3).

not applied to ignorance. "But all of these, we say, often turn a man, when he is being drawn to his own wish, at the same time towards its opposite" (πάντα δέ γε προτρέπειν ταῦτά φαμεν εἰς τὴν αὐτοῦ βούλησιν ἐπισπόμενον ἕκαστον εἰς τὰναντία πολλακίς ἅμα).⁹ There follows immediately the definition of injustice as the tyranny of various feelings in the soul. The sequence of thought shows the meaning to be that these feelings can tyrannize the soul either directly, by their own violence, or indirectly, by the submission of an intellect unaware of its proper end to the lower impulses. The emphatic position of πάντα δέ γε (863E2) indicates that what is being stressed is a quality common to the three sources of error. The next lines confirm this: their common quality is that all three involve injustice. The state of soul described in 863E6–9, injustice, is the same as that described in 863E2–3; this is why the one passage leads directly into the other. The scheme is, therefore:

- (1) unintentional injury always just
- (2) intentional injury always unjust
 - (a) from anger and pleasure, which tyrannize the soul
 - (b) from ignorance, which allows such tyranny

The context, therefore, shows that the passage starting τὴν δὲ τοῦ ἀρίστου δόξαν κτλ., the description of the just soul, must not describe a condition of possible ignorance. The phrase ἡ τοῦ ἀρίστου δόξα, then, does not mean "the opinion as to what is best," but instead "the conviction which has the best for its object."¹⁰ The next clause means "in whatever way a city or private individuals may judge that this (the best) will come to pass."¹¹ This clause is not meant to concede that the judgment of cities and individuals may be in error, but rather that they will not all achieve "the best" in the same way. Circumstances will dictate a different course of action to each.¹² "If this conviction guides every

⁹ τὴν αὐτοῦ βούλησιν refers to a man's wish for his *own* good as opposed to the whim of his θυμός or ἡδονή, which are conceived as having their own βουλῆσεις (863B8–9). For the meaning of βούλησις in 863E2–3 cf. *Gorgias* 467A–468E.

¹⁰ The closest parallel to this use of δόξα is in *Phaedrus* 237D ff., where, as in *Laws* 863E–864A, the discussion concerns a contest for power within the soul. The two contenders are ἡ μὲν ἐμφυτος οὔσα ἐπιθυμία ἡδονῶν (237D7–8) and ἐπικτητος δόξα, ἐφίεμένη τοῦ ἀρίστου (237D8–9), also described as δόξης . . . ἐπὶ τὸ ἀρίστον λόγῳ ἀγούσης (237E2–3) and δόξης ἐπὶ τὸ ὀρθὸν ὁρμώσης (238B8). In these phrases δόξα is a movement of reason towards the best.

¹¹ I suggest this as a translation for this difficult clause. It requires the emendation τοῦτο γ' (Hermann, followed by Apelt, England, Taylor) for the manuscripts' τούτων (kept by Burnet, Grube). (There is a variant τούτων.) Cousin, Grote, and G. Müller read τοῦτο. Bury and Diès read δπηπερ ἂν ἐφείσθαι τούτου ἡγήσωνται. C. Badham (above, note 1) 49 reads δπηπερ ἂν ἔπεισθαι τούτῳ ἡγήσωνται πόλεις.

¹² The use of the future infinitive indicates that the object of the judgment

man and rules in his soul," Plato continues, "even if some injury should take place,¹³ we must say that every action is just which is done in accord with it and everything which takes place in obedience to its rule, in every individual, and best too for the whole life of men, but that an injury of such type is regarded by many as an involuntary injustice. Our argument does not now involve a contention about words;¹⁴ rather, since it has been shown that three types of errors occur, we should first recall these again. One type is that of pain, which we designate anger and fear. . . . The second is of pleasure and desires. A third type is that of expectations and an aiming at true opinion about the best."¹⁵

So interpreted, the passage does not introduce a new internal standard of justice, the "good conscience," but simply reiterates the earlier Platonic doctrine that to be just is to be wise.¹⁶

ἡγήσωνται is not "the nature of the best" but rather "how it will come to pass." *ἔσεσθαι* in the sense of "will happen" is unusual, but cf. Thucydides 2.13: *ὅτε ἡ ἐσβολὴ τὸ πρῶτον ἐμελλε Πελοποννησίων ἔσεσθαι*. Stallbaum translates: "quacunq[ue] demum ratione illud iudicaverint locum habiturum esse . . ." (3.53). Cf. also Apelt's translation: "die Rücksicht auf das Beste, wie sich auch immer der Staat oder der Einzelne die Verwirklichung desselben denken mögen . . ." (*Platons Gesetze* 367–68).

¹³ The acc. *τὴν . . . δόξαν* is an anacoluthon, like *ὠφελίαν* in 862A7 (cf. Stallbaum 3.51).

¹⁴ This responds to 860E3–4: *τίνα οὖν αὐτὸν τρόπον ἐγωγε συμφωνοίην ἂν τοῖς ἐμᾶτοῦ λόγοις*;

¹⁵ This keeps the reading of the manuscripts. The condition referred to is ignorance, so the *ἔφεσις* is one which misses the mark and fails of true opinion. Ast excises *τῆς ἀληθοῦς . . . ἔφεσις*. Others emend variously: *ἄφεσις* (Grou, Taylor, Diès); *ὑφεσις* (Jackson); *ἀμαθοῦς* for *ἀληθοῦς* (England); *δόξης τῆς μὴ ἀληθοῦς περὶ τὸ ἄριστον ἔφεσις*, or *δόξης τοῦ ἄριστου παρὰ τὸ ἀληθὲς ἔφεσις*, or *δόξης παρὰ τὸ ἄριστον ἔφεσις* (Ritter); *δόξης τῆς μὴ ἀληθοῦς περὶ τὸ τοῦ ἄριστου ἐφέσθαι* (Bury); *δόξης τῆς ἀμαθοῦς πρὸς τᾷριστον ἔφεσις* (Badham). Cf. Gernet 109, and England 2.403–4. For the meaning of the sentence which follows (864B8) see England 2.404.

¹⁶ *Rep.* 443E6–444B8. Aristotle's distinction between the terms *ἀτύχημα*, *ἀμάρτημα*, and *ἀδίκημα* can be found in *E. N.* 1135B12 ff. and, in a slightly different version, in *Rhet.* 1374B7. As usual, such distinctions are not Plato's.