Companions to ancient thought 4

Ethics

EDITED BY
STEPHEN EVERSON

CAMBRIDGE UNIVERSITY PRESS

3
Platonic ethics

C. C. W. TAYLOR

The fundamental question of Platonic ethics is 'How should one live?' (*Republic* 1.352d, *Gorgias* 500c). That question is not to be understood as 'What is the morally best way to live?', as is shown by the fact that in *Rep.* 1 an appropriate, though in Plato's view false, answer to it is that given by Thrasymachus, namely that one should live by emancipating oneself to the best of one's ability from the restraints of morality with a view to the furtherance of one's own interest. Rather it is to be understood as 'How may one achieve the life which is, objectively, but from the point of view of one's own interest, the most worth living?' (*Rep.* 1.344e). The Greek term for the achievement of such a life is *eudaimonia* (literally 'having a favourable guardian spirit' (*daimōn*)), conventionally translated 'happiness', but in view of its objective character better rendered 'blessedness' or 'well-being'. According to Aristotle (*Nicomachean Ethics* (*EN*) 1095a18–20) it was universally acknowledged (a) that *eudaimonia* was the supreme good and (b) that the term meant 'living well' and 'doing well'; nothing in the texts of Plato suggests that his use of the term conflicts with these claims. In the same passage Aristotle tells us that there were substantive disputes about what living well amounted to, some holding, for example, that it consists in acquiring wealth, others that it consists in a life of honour or of intellectual achievement; Plato depicts such substantive disputes in Socrates' confrontations with Thrasymachus in *Rep.* 1 and Callicles in the *Gorgias*.

The agreement on 'How is one to live well?' as the basic question of ethics forecloses certain ethical disputes while leaving others open. Most fundamentally, ethical questions are approached from the standpoint of the individual's interest, the promotion of which is assumed to be the primary function of the individual's practical rationality. On that assumption one has adequate reason to undertake any action if and only if so doing will contribute to one's living well, i.e. to one's having an objectively worthwhile life. The conception of an objectively worthwhile life should not be construed in a

narrowly egoistic way, since it may be part of an objectively worthwhile life that one cares for the good of others, not merely instrumentally, with a view to the benefits one may expect to gain from such benevolence, but for its own sake. Nevertheless, it is broadly egoistic,[1] in that it is assumed that the value to the carer of that selfless care lies primarily in its contribution to the *life of the carer*, and only secondarily in its contribution to the life of the person cared for. Since this broadly egoistic conception of the role of practical reason is assumed from the outset, there is no room in Platonic thought for theories of a Kantian type, which seek to identify moral principles as imperatives binding unconditionally on any rational agent in total independence of any considerations of the interest of that or any other agent.[2] It is therefore unsurprising that our texts provide no hint that the very conception of that type of theory had so much as occurred to Plato.

The broadly egoistic starting-point of Plato's ethical enquiries is not, then, open to question. By contrast, the status of morality is an eagerly debated question in some of his major dialogues. By morality I understand a socially regulated system of norms imposing restraints on the pursuit of self-interest with the general aim of furthering social co-operation, for which the nearest Greek terms are *to dikaion* and *dikaiosunē*, conventionally 'the just' and 'justice'. The questions 'What is the morally right thing to do?' and 'Which is the morally best way to live?' were certainly not unintelligible to Plato, or to Greeks of his time generally. They were, however, both distinct from and posterior to the fundamental question which we have already identified, 'How should one live?', i.e. 'How should one achieve the best life for oneself?' The former questions were posterior in that, whereas the rationality of the pursuit of the best life for oneself was unquestioned (indeed one could go so far as to say that that pursuit constituted practical rationality for the Greeks), the rationality of the individual's observance of the dictates of morality required to be established by showing that the acceptance was

[1] For a useful discussion of these two varieties of egoism, labelled respectively 'moral solipsism' and 'moral egocentrism', see Irwin [293], 255.

[2] This contrast between Platonic and Kantian theory leaves open the question whether, in the former, the agent's interest may itself be seen as consisting in the acquisition of states of the personality (i.e. virtues) which have value independently of their contribution to the agent's *eudaimonia*. There are theoretically at least three possible views on the relationship of virtue to *eudaimonia*: (i) virtue is valuable purely instrumentally, as a means to *eudaimonia*, (ii) virtue is at least partly constitutive of *eudaimonia*, and is intrinsically valuable *qua* constitutive of *eudaimonia*, (iii) virtue is valuable both in its own right and as either a means to or as a constituent of *eudaimonia*. I argue in note 21 below that (ii) is closer than (i) to giving an account of Plato's view in the early dialogues. I know of no evidence to suggest that Plato was aware of (iii) as an alternative to (i) and (ii). (*EN* 1097b2–5 indicates that Aristotle may have been.)

necessary for the achievement of the individual good life. Plato attempts to meet the challenge to provide this justification of morality in the *Gorgias* and, in a much more elaborate and extended form, in the *Republic*.

He accepts, then, that morality requires justification, in the form of a defence of its rationality, that this justification must be in terms of a broad conception of individual interest, and that such a justification can be provided. The requirement of justification seems an inevitable response to the fact that morality is essentially cooperative, requiring sacrifices from the individual for the common good. By Plato's time there had been developed theories of the social nature of morality, which attempted to ground morality in self-interest (in a similar fashion to the theories of Hobbes and Hume[3]) by showing how norms of self-restraint and social co-operation would naturally develop in primitive societies as a device for mutual protection against the onslaughts of wild animals or (more plausibly) of unsocialised individuals.[4] But the success of such theories seems limited. They show convincingly why self-interested individuals have reasons to prefer the existence of such institutions to a Hobbesian war of all against all, 'since we benefit from one another's justice and goodness' (*Protagoras* 327b). But they are unable to show that, given that the institutions exist, each individual benefits more from the sacrifices which he or she is required by the norms of the institution to make than he would do by taking advantage of the sacrifices of others to promote his own interest. The sacrifices might be regarded as one's subscription to the mutual security club, which it is in one's interest to pay, since the other members would not accept that one should enjoy the benefits of membership without paying one's dues. But if one can get away without paying, as one fairly clearly can *now and again* (though not, doubtless, always), why pay *in those circumstances?* Of course, it is unfair not to pay, which is a perfectly good reason for someone who is already committed to being fair, but the theory was supposed to generate a purely self-interested reason for undertaking that commitment, which it clearly fails to do. As Glaucon points out (*Rep.* 362a), the most that the theory can provide is a self-interested reason for making other people believe that one always deals fairly with them, which falls short of a reason for always actually doing so; but the latter, not the former, is what is required for the justification of morality.

These problems reflect a crisis in traditional Greek morality, to which

[3] Hobbes, *Leviathan*, chs. 13, 17; Hume, *Treatise* III.ii.1–2.

[4] Plato provides examples in the myth in the *Protagoras* (320c–322d) and in the theory proposed by Glaucon in *Rep.* II (358e–359b); an example independent of Plato is the so-called 'Anonymus Iamblichi', on which see Guthrie [3], 71–4 and 314–15. For a valuable discussion of the historical sources of this tradition see Kahn [106].

Plato's ethics presents a sustained response. Traditional morality recognised certain states of character, principally courage, self-control, justice or fairness and piety, as the principal qualities which made their possessor an outstanding and admirable person.[5] The young were brought up to regard possession of these qualities as fine and admirable (*kalon*) and the lack of them as disgraceful (*aischron*), and their inculcation was the principal aim of education.[6] These were the most important among the excellences (*aretai*), i.e. those qualities stable possession of which, together with such external goods as wealth, position in society and physical health, constituted success in life (*to eu zēn* or *eudaimonia*), which, we have seen, was universally acknowledged as the supreme good. But the arguments which we have just glanced at show that the claims of certain *aretai*, notably justice, to be constitutive of the agent's good are at odds with the other-regarding character of those qualities. A further difficulty arises from the fact that some of these qualities are no less indeterminate in character than *eudaimonia* itself. Thus even if it is granted that success in life requires piety, i.e. a proper attitude to the gods, including respect for those obligations which the gods impose on us, there can be apparently irresoluble disputes about what kind of conduct really is required by the gods, as in the famous example in Herodotus (III.38) of the diverse customs of different nations in the disposal of the dead, a central case of religious obligation. Both difficulties may be seen to have prompted an emphasis on the distinction between the nature or reality of things (*phusis*) and convention (*nomos*).[7] The lack of coincidence between the agent's interest and the demands of morality leads to the claim that while nature prompts us to seek our own interest, the demands of morality spring from nothing but convention (with the implication that the latter, unlike the promptings of nature, lack any authority).[8] The indeterminacy of some of the conventional excellences is similarly attributed to the facts that

[5] Pindar, *Isthmian* 8.24–7; Aeschylus, *Seven Against Thebes* 610; Euripides, fr. 282.23–7; Xenophon, *Memorabilia* III.ix.1–5; Plato, *Prot.* 329c, *Meno* 73e–74a, *Rep.* 427e, etc. (See Irwin [293], p. 287, n. 1.)

'Self-control' renders *sōphrosunē*, a term which lacks a precise English equivalent. It connotes primarily a proper sense of oneself and one's limitations in relation to others, and derivatively various applications of that sense, especially control of the bodily appetites. Hence in many contexts, including those provided by Plato's tri-partite psychology, it is appropriately rendered 'self-control'. Where in this chapter I use the terms 'self-control' and 'self-controlled' they correspond to the Greek *sōphrosunē* and its cognate adjective *sōphrōn*, but I have not attempted total uniformity of usage, preferring sometimes to use the Greek term. (In my commentary on the *Protagoras* ([237]), I prefer the rendering 'soundness of mind', for reasons explained there on pp. 122–4.)

[6] *Prot.* 324d–326e.

[7] Principal discussions of the *nomos–phusis* distinction include Heinimann [54]; Guthrie [3], ch. 4; Kerferd [86], ch. 10.

[8] *Rep.* 359c. Cf. *Prot.* 337c–d.

(a) their requirements arise from nothing more than convention, and (b) different societies have different conventions.[9]

These criticisms of conventional morality should not be assimilated to modern attacks on the objectivity of values. Critics who, like Glaucon or the sophist Antiphon,[10] denigrate conventional morality by contrast with the promptings of nature assume that the latter is a locus of objective value. The extreme form of this position is upheld by Callicles, who maintains that unrestrained self-assertion is naturally *right* (*phusei dikaion*), and that the conventional morality (*nomōi dikaion*) which opposes it is, therefore, naturally wrong, i.e. really or objectively *immoral* (*Gorg.* 483c–484c).[11] While Thrasymachus is not prepared to go that far, he asserts that injustice is a form of wisdom and an excellence, in that it gives the agent a worthwhile life, while justice is weakness and folly, in that it harms the agent and promotes the good of others (*Rep.* 343b–344c, 348c–e). We have no reason to interpret these claims otherwise than as statements of fact. It is significant that the earliest application of the nature–convention contrast to morality, attributed to Archelaus, who is said to have been a pupil of Anaxagoras and a teacher of Socrates, states simply that 'the just and the disgraceful are by convention, not by nature' (Diogenes Laertius II.16; D–K 60A1). Here, what is purely a matter of convention is what is morally right and wrong, not what is good or bad; for example there is no suggestion that it is purely (or at all) a matter of convention that health is a good state and illness a bad one. Later, when the sceptics applied their universal strategy of suspension of judgement to the special case of claims about value, they did not confine their critique to moral value, but applied it to value generally;[12] but that seems to have been a post-Platonic development. At *Theaetetus* 172a–b Socrates asserts that 'in matters of what is just and unjust and holy and unholy [people] are willing to maintain that none of these things is so in reality (*phusei*) or has its own nature (*ousian*), but what is agreed on [sc. by each community] is the truth for as long as it is agreed', but contrasts that conventionalism about morals with the position about what is advantageous, where 'no one would dare to say that what a community lays down as advantageous for itself is so in fact'. That assertion is confirmed by our other evidence: both sides of the dispute about nature and convention accepted that genuine values were part of nature, the critics of conventional morality attacking its values as spurious because they are *merely* conventional and therefore not part of nature, its defenders urging that on the contrary moral values are natural and therefore genuine. We have already noticed one

[9] Herodotus III.38.
[10] Diels–Kranz 87B44. See Guthrie [3], chs. 4 (a) (ii) and 11 (5); Saunders [107].
[11] I discuss the point more fully in my [68].
[12] See Julia Annas' chapter below.

defensive move, the theory (outlined in Protagoras' myth) that moral conventions are themselves natural, in that they are strategies for cooperation developed by human beings struggling for survival in a hostile environment. But that defence was insufficient, since it failed to show that moral value passed the primary test for being natural, namely that of promoting the individual well-being of the agent. If his defence of traditional morality was to pass that test, Plato had to develop a better theory of the nature of morality and of human nature, in the hope of demonstrating the objective goodness of the traditional virtues via their contribution to the perfection of that nature, and therefore to the objectively worthwhile life for the agent.

In what follows I shall set out what I take to be the main lines of Plato's attempt to develop such a theory. I shall distinguish three stages in this process:

- i the theory of the early dialogues[13]
- ii the theory of the *Republic*
- iii developments subsequent to the *Republic*

i The early dialogues

A central preoccupation of the early dialogues is the search for definitions, whether of individual excellences (courage in *Laches*, *sōphrosunē* in *Charmides*, piety in *Euthyphro*), of excellence in general (*Meno*) or of friendship (*Lysis*), an aspect of life intimately related to excellence and the good. In order to understand the prominence of definition in these dialogues, and its connection with the theory of the nature of the virtues which emerges in them, it is necessary to consider what Plato's Socrates is looking for when he looks for an ethical definition. In outline, the project of the early dialogues is to give accounts of the traditional virtues which will exhibit them as natural goods; to investigate how that project was carried out would require close examination of the relevant texts, for which space is lacking here.[14] I must confine myself in this chapter to a bald statement of the results of that investigation, focusing primarily on the *Meno*, which, though probably one of the latest of the dialogues which I here count as early, and in its introduction of the theory of recollection transitional to the metaphysical dialogues of the middle period, is of all the early dialogues the richest in evidence for the Platonic/Socratic theory and practice of definition.

A Socratic definition is not, in the first instance, a definition of any of the items that might spring to the mind of the modern reader, such as a term,

[13] For the purpose of this chapter I count the following as early dialogues: *Apology*, *Crito*, *Euthyphro*, *Charmides*, *Laches*, *Lysis*, *Euthydemus*, *Protagoras*, *Meno*, *Gorgias*.

[14] For fuller discussion see my 'Socratic Ethics' in [130], 137–52.

the meaning of a word, or a concept. Ideally, a Socratic definition answers the question 'What is . . . ?', where the blank is filled in by a word designating some quality or feature of agents, such as courage or excellence. So, literally, what are to be defined are those qualities or features themselves, not anything standing for them, as words or perhaps concepts might be thought to do. But, of course, we cannot in general draw a sharp distinction between specifying what something is and defining or elucidating the concept of that thing. The concept of F (where 'F' is some general term) is what we understand or possess when we use the term 'F' with understanding, and in some cases saying what F is precisely is defining or elucidating the concept of F. Thus if I answer the question 'What is justice?' by saying that justice is giving everyone their due, I have thereby attempted (however inadequately) to elucidate the concept of justice, in that the answer is intended to make explicit what is standardly conveyed by our talk of justice. The ultimate authority for the correctness of that sort of definition is the competent speaker of the language in which the elucidation is expressed, and the ultimate test which that authority applies is conformity with his or her linguistic intuitions. In other cases, however, the question 'What is F?' is aimed to elicit, not an elucidation of the ordinary concept of F, but an account of the phenomenon couched in terms of the best available scientific theory. For example, 'light is a stream of photons' is not an elucidation of the ordinary concept of light, i.e. of what the standard speaker of English understands by the word 'light'. It is an account of what light is, i.e. of what science has discovered light to be, and that account presupposes, but is not exhausted by, the grasp of the concept which is available to the competent but pre-theoretical speaker of the language.[15] Hence the ultimate test of its adequacy is not its fit with that speaker's linguistic intuitions, but its explanatory power, empirical testability, or whatever else constitutes the test of a good scientific theory. The form of words 'What is . . . ?' may express the search for a definition of either kind (and in any case the distinction between the two is less sharp than the foregoing over-simplification has suggested[16]). Precisely what kind of search is afoot has to be determined by the context, which cannot be guaranteed to provide an unambiguous result.

We learn from the *Euthyphro* that a Socratic definition of a given quality should (a) specify what is common to all and only those things to which the name of the quality applies (5d), (b) specify that in virtue of which the name applies to them (i.e. give the nature of the quality, not a mere distinguishing mark of its presence) (6d, 11a), and (c) provide a criterion by reference to

[15] In certain cases the scientific account may even demand revision of the pre-existing concept, as in cases where the latter itself carries connotations of a superseded scientific theory, for example hysteria.

[16] See Putnam [964].

which disputed cases may be determined (6e). Requirements (a) and (b) are explicitly endorsed at *Meno* 72c, where Meno is invited to specify 'the single nature they [i.e. the various types of human excellence] all have in virtue of which (*di'ho*) they are excellences', and though the *Meno* has nothing to say about disputed cases of *aretē*, we have no reason to suppose that the third requirement has been abandoned. These three requirements are satisfied alike by conceptual elucidations and by that kind of account which we have contrasted with those, and which we might call, traditionally, 'real definitions' or, perhaps more informatively, 'substantive or scientific accounts'.

In the course of the discussion Socrates gives two model definitions, of shape and colour respectively. The first of these, 'Shape is the limit of a solid' (76a7), is a conceptual elucidation, whereas the second (76d4–5) is a substantive account of colour in terms of a scientific theory, namely the physiology of Empedocles.[17] He says that the latter is inferior to the definition of shape, but he does not say why it is inferior, and it would therefore be rash to conclude that Socrates' preference for the definition of shape indicates a preference for one *type* of definition, conceptual elucidation, over another, a causally explanatory account. Indeed, the text thus far gives no indication that Plato is even aware of the distinction between those types of definition. So if we are to answer the question 'What kind of definition is Socrates looking for in the *Meno*?', we must take into consideration the rest of the dialogue, where, though nothing more is said about the methodology of definition, the question 'What is excellence?' is answered. Excellence is first (87c–89a) argued to be knowledge, then another argument leads to the revision of that account in favour of the answer 'Excellence is true opinion' (99b–c), an outcome which is further qualified by a strong hint at 100a that the former answer gives the true account of genuine excellence, while the latter gives an account of what passes for excellence by ordinary standards.

What these answers share is a conception of excellence as a cognitive state, or more portentously, a grasp of truth. We are not concerned with the details of the distinction between knowledge and true opinion, but its essence is the firmness or reliability of the grasp; knowledge is a reliable (because systematic) grasp of truth, while true opinion is an unreliable (because unsystematic) grasp of truth. What sort of account of *excellence*

[17] 'Colour is a flowing out [sc. from the coloured object] of shapes [i.e. physical particles of various shapes] symmetrical to vision [i.e. of such shapes and sizes as enable them to penetrate the channels in the eye] and perceptible' (76d4–5). (The modern analogue would be an account of colour in terms of light waves of different lengths.)

is provided by the conception of it as a cognitive state? Note that Socrates and Meno are not discussing excellence in some theoretical sphere, such as excellence at mathematics; the paradigms of excellence in this final stage of the discussion are individuals such as Pericles who embody the same ideal of success in public and private affairs as Meno had assumed from the outset. That is to say, cognitive states are presented as giving an account of all-embracing social and political merit, of the state of the totally well-rounded, successful and admirable person (= man, by this stage of the discussion). In what sense is that person's admirable state a cognitive one? The sort of excellence in question is above all practical, manifested in action and a whole style of behaviour; it does not seem that specification of a cognitive state could give an account of *the manifestation* of excellence. In other words, 'knowledge or true opinion' does not offer the same kind of account of excellence as 'having a good service, ground strokes and volley' does of excellence at tennis, since the latter account does precisely specify the kind of actions which manifest that excellence, whereas the former does not. Rather, it gives an account of *what is manifested in* excellent performance, as 'co-ordination, stamina, courtcraft, etc.' does of what is manifested in excellent tennis-playing. And as that account expresses not an elucidation of the concept of excellent tennis-playing but an empirical theory of the causes of the type of play which we count as excellent, so that cognitive account of overall human excellence expresses, not an elucidation of that concept, but a causally explanatory theory, what we may call the Cognitive Theory of Excellence. We find, then, that despite Socrates' expressed preference for the conceptual definition of shape over the causally explanatory account of colour, the account of excellence which he endorses in the concluding section of the dialogue is of a type represented, not by the former, but by the latter.

The aim of our enquiry into the nature of definition in the early dialogues was the elucidation of the nature of the theory of virtue to which that practice of definition is preliminary. That aim has now been achieved in part, with the identification of that theory as the Cognitive Theory of Excellence (more familiar as the first of the so-called 'Socratic paradoxes', the thesis that virtue is knowledge). But that elucidation lacks content without further exploration of just what the Cognitive Theory claims. That exploration will also, I hope, indicate the connection between the Cognitive Theory and the second paradox, the thesis that no one does wrong willingly,[18] and also with the much-discussed question of the unity of virtue.

[18] On the paradoxes see Santas [161].

The Cognitive Theory is a theory to the effect that overall success in human life is guaranteed by the possession of certain cognitive states.[19] This theory in turn rests on a theory of the explanation of intentional action, which combines to a remarkable degree a staggering audacity and simplicity with a high degree of plausibility. It states that provided that the agent has a conception of what is overall best for the agent, or (equivalently) what is maximally productive of *eudaimonia* (for the agent), that conception is sufficient to motivate action with a view to its own realisation. That is emphatically not to say that motivation does not require desire as well as belief. On the contrary, Socrates makes clear his view (77c1–2, 78b4–6) that everyone desires good things, which in context has to be interpreted as the strong thesis that the desire for good is an invariable motive. That desire is then conceived as a standing motive, which requires to be focused in one direction or another via a conception of the overall good. Given that focus, desire is as it were locked on to the target which is picked out by the conception, without the possibility of interference by conflicting desires. Hence, given the standing desire, all that is required for the correct conduct, i.e. for the manifestations of excellence, is the correct focus. And that focus has to be a correct conception of the good for the agent, i.e. a correct conception of *eudaimonia*.

From this theory it follows immediately, given the conception of excellence accepted in the *Meno* as what is manifested in excellent conduct, that

[19] There has been a lively debate among recent commentators on whether the Socratic claim is the extreme claim that virtue (i.e. knowledge) is sufficient for success in life (i.e. *eudaimonia*) *by itself*, or the less ambitious claim that it is sufficient given a (modest) sufficiency of those goods which it is beyond the power of the agent to procure, for instance good health. The former view is maintained by Irwin [293], 100 and [180], the latter by Vlastos [184], and by Brickhouse and Smith [179]. The issue seems to me not to be addressed very clearly in the dialogues. While the Socratic claim that the good man cannot be harmed (*Apol.* 30c, 41c–d), *if taken literally*, implies the former view, that view is clearly inconsistent with the assertion at *Crito* 47e that it is not worth living with a sickly body, unless we attribute to Socrates the implausible view that virtue guarantees good health. He might indeed have held that virtue will minimise the causes of ill-health by eliminating those which spring from lack of self-control, but it seems very implausible that he should have thought that such factors as climate or accident could have no effect on one's health. As Irwin points out ([180], 92–4), Socrates does argue at *Euthyd.* 279–80 that wisdom always makes those who possess it *eutuchein*, so that the wise do not need good luck, but in the context that seems to amount just to the claim that knowledgeable practitioners of any skill (doctors, etc.) are more successful than the unskilled, and that they do not need to rely on luck, as the unskilled do. In general, Socrates' insistence on the paramount importance of the state of one's soul (assumed to be within one's control) and the comparative worthlessness of other reputed goods (*Apol.* 28b, *Prot.* 313a) has the outcome that the virtuous agent will need very little from fortune to give him or her a totally worthwhile life, but it is pressing the literal reading of the texts too far to take him to claim that he needs literally nothing.

the traditional virtues are in fact one and the same state of the agent. Courage is that stable state which is manifested in the proper handling of fearful situations, justice the stable state which is manifested in one's proper dealings with others, piety the stable state which is manifested in one's proper dealings with the gods. And the stable state in question is the same state in every case, namely the agent's grasp of the correct conception of the good. Indeed, that same state is designated by the non-synonymous names of the traditional virtues, whose distinct connotations pick out the distinct manifestations of that single state: for example the connotation of the name 'justice', i.e. 'what is manifested in one's proper dealings with others', is distinct from the connotation of the name 'piety', i.e. 'what is manifested in one's proper dealings with the gods'. The point of retaining these various names is to do justice to the fact that for the achievement of overall success in life the same cognitive state has to be manifested in various ways, whether in various types of conduct which overlap only partially (for example if all courageous actions are self-controlled but not vice versa) or in different aspects of co-extensive action-types (for example if all and only just actions are pious, their being pious, i.e. appropriate to dealings with the gods, is a specific modification of that attribute of the action-type which is their being just, i.e. appropriate to dealings with others). But the retention of the names of the traditional virtues should not disguise the essence of the theory, that what is manifested in all these different ways is identical, namely the agent's grasp of his overall good. The 'parts' of total excellence are not distinct motive-forces or tendencies to action, as on the traditional conception, which allowed that they might be separable from one another. Nor are they distinct cognitive states, for example kinds of knowledge, as knowledge of history is a distinct type of knowledge from knowledge of geometry; knowledge of how to treat others and knowledge of how to control one's passions are not distinct types of knowledge, but rather different aspects of a comprehensive knowledge of how to live, which is what controls one's activity in all areas. That is precisely the account of the virtues which the cognitive theory of the *Meno* would lead us to expect, and it is the theory which we find explicitly argued for in the *Protagoras* and implicit in the accounts of individual virtues in the *Euthyphro*, *Laches* and *Charmides*.[20]

The central claim of the Cognitive Theory, as so far elucidated, is not one from which Thrasymachus or Callicles need dissent; that theory claims that a reliable grasp of the good is sufficient for overall success in life, but is of itself neutral between incompatible conceptions of the good. In order to realise

[20] I discuss the unity of virtue more fully in my commentary on the *Protagoras* ([237]) and in [219]. Vlastos has vigorously contested this account of the doctrine in his [264]. For a reply to his main points see the revised edition of my commentary.

Socrates' aim of vindicating the traditional conception of virtue it has to be supplemented by an account of the good, which will show either that traditional virtue (or a sufficiently close approximation to it) is instrumentally necessary for success in life, or that it is as least partially constitutive of it.[21] In advance of that account the Cognitive Theory entails the self-interested version of the second Socratic paradox 'No one acts intentionally against his overall interest.' The supplementation of the Cognitive Theory with that account yields the moral version of the paradox 'No one does intentionally what is morally wrong.' But when we look for arguments in favour of that account, the dialogues provide us with very little. *Crito* 47e states, but does not argue for, the analogy between health of the body and justice in the soul, asserting that injustice damages the soul as sickness does the body, and that as it is not worth living with a sickly and diseased body, even less is it worth living with a corrupted soul. On the strength of that analogy it is agreed (48b) that living well (the universally acknowledged good) is the same thing as living creditably (*kalōs*) and justly (i.e. that the good life is identical with the moral life), but Thrasymachus could properly point out that no argument has been given for the crucial claims that injustice harms the soul and that justice benefits it. *Gorg.* 504–5 merely gives a more extended version of the analogy: all types of craftsmen aim to produce a good product, and in

[21] Irwin argues in [293] (for example on pp. 84–5) that the craft analogy commits the Socrates of the early (= Socratic) dialogues to an exclusively instrumental view of the relation of virtue to *eudaimonia*, and treats acceptance or rejection of this instrumental view as the criterion for distinguishing between the views of 'Socrates' and those of 'Plato'. But first, not all crafts (*technai*) are exclusively instrumental in character. A *technē* is any skilled activity which can be systematically taught, a description which embraces the performing arts as well as the productive: see, for instance, *Symp.* 187b (music), *Laws* 816a (dancing). The craft analogy itself, therefore, need not prevent Socrates from claiming both that virtue is a *technē* and that it is at least partly constitutive of *eudaimonia*, while *Crito* 47e shows him firmly committed to the latter claim. So unless Irwin is to attribute the view of the *Crito* to 'Plato' rather than to 'Socrates', he must abandon the view that the latter's conception of the relation of virtue to *eudaimonia* is *exclusively* that of an instrumental means. Irwin's view is criticised by Zeyl in [186].

The above is not to say that the position of the early dialogues on the relation of virtue to *eudaimonia* is consistent. As pointed out in the text (p. 63 below), in different passages Socrates is represented as maintaining both that virtue is knowledge (sc. *of* the good) and that it is itself the health of the soul, and therefore the good itself. Those theses are inconsistent, since knowledge of the good requires that the content of that knowledge should be independently specifiable, which is impossible if the good of which one has knowledge is the very state of having knowledge of the good. The fault, however, lies not with the craft analogy, which can be stated without inconsistency as the thesis that success in life is a skilled activity analogous to skill in the performing arts, but with the conjunction of the claims that excellence is a state of knowledge of how to achieve a goal and that it is that goal itself.

every case the goodness of the product consists in its order and arrangement. So a well-made boat has all its parts properly fitted together, and in a healthy individual all the bodily constituents are properly ordered. As health is the name for the proper ordering of bodily components, the name for the proper ordering of psychic components is justice and *sōphrosunē*, which expresses itself in proper conduct towards gods and men (piety and justice) and in the proper control of pleasure and pain (courage) (507a–c). Once again, the crucial identification of the proper order of the soul with conventional justice and *sōphrosunē* is unargued. It is an appropriate *ad hominem* rejoinder to Callicles that his ideal of the unrestrained satisfaction of the pleasure of the moment does not provide an adequate rule for the long-term planning of one's life, and that without some such rule one's life will collapse into a chaos of conflicting desires.[22] To be satisfying one's life must be coherent, and to be coherent it must contain some discrimination of pleasures into the more and less significant, and some circumscription of the pursuit of the latter with a view to the greater enjoyment of the former. But that element of rational planning and self-control might be exercised in the pursuit of a life of injustice and self-indulgence; Don Juan could satisfy that requirement by making the seduction of as many women as possible his paramount aim, and by refusing to be distracted from it by the momentary attractions of a life of scholarship or of quiet domesticity.

Earlier in the *Gorgias* (473–5) Socrates argues against Polus that injustice is against the agent's interest, and hence (by the second paradox) that no one acts unjustly intentionally (509e). The argument does not contain any positive account of the agent's good, but proceeds directly to the conclusion that injustice is bad for the agent, relying on Polus' admission that injustice, while more advantageous to the agent than justice, is more disgraceful than justice; Socrates gets Polus to agree that if x is more disgraceful than y, then x is either more unpleasant than y or worse (i.e. more disadvantageous) than y, and then concludes that injustice is worse than justice via the claim (agreed by Polus) that injustice is not more unpleasant than justice. The weakness of the argument is obvious; as Callicles points out (482d–e), a tougher-minded opponent (such as Thrasymachus) would not have made the initial concession that injustice is more disgraceful than justice. Further,

[22] Of course, one's only rule for planning one's life might be 'Don't plan: take every pleasure as it comes.' Such a rule is not formally self-contradictory, nor is the description of someone as attempting to live by it. It is, however, practically self-defeating, in the sense that anyone who tried to put it into practice would find that the attempt to adhere to it consistently required him to break it. The reason is that some pleasures, as they come, require that one should plan, for example the pleasure of getting started on a career.

even given that concession, Socrates' argument requires the general principle that if action-type *a* is more disgraceful to the agent than action-type *b*, it must be either more unpleasant *to the agent* or worse *for the agent*, but there is no ground to think that that principle is true; an action-type might be disgraceful to an agent (i.e. such as to bring him into justified disrepute) which was neither unpleasant to that agent nor bad for him, but either unpleasant to or bad for others.

The only other early dialogue which contains any account of the good is the *Protagoras*, where Socrates is represented as arguing for the conclusion that courage is a kind of knowledge from the premise that the good is pleasure. Without renewing the controversy as to whether this premise is (as I believe) presented in the dialogue as Socrates' own view or merely that of the majority of ordinary people,[23] it has once again to be observed that it is (a) not itself supported by any argument and (b) insufficient to provide the desired vindication of conventional morality.

Taken as a whole, then, the early dialogues fail to realise Plato's project of providing that vindication. The theory of motivation which is outlined in

[23] Commentators are divided on the question of whether Socrates is represented as seriously espousing hedonism. I append a list of some writings on either side of the dispute.
For the thesis that he is serious see:

Grote [289], vol. 2, pp. 87–9
J. and A. M. Adam [236], xxix–xxxiii
Hackforth [252]
Vlastos [239], xl (note)
Dodds [271], 21–2
Crombie [287], vol. 1, pp. 232–45 (with reservations)
Irwin [293], ch. 4
Taylor [237], 208–9
Gosling and Taylor [53], ch. 3
Nussbaum [60], ch. 4
Cronquist [242]

Against:

Taylor [299], 260–1
Sullivan [263]
Raven [291], 44–9
Gulley [117], 110–18
Vlastos [164]
Manuwald [258]
Kahn [175]
Dyson [245]
Duncan [244]
Zeyl [185]
Stokes [119], 358–439
Kahn [233] and [253].

them requires to be complemented by an account of *eudaimonia*, but no such account is provided. Instead, the whole theory rests on the analogy between conventional virtue and bodily health, which begs the crucial question of the value of conventional virtue to its possessor. Moreover, that analogy threatens the cognitive account of virtue itself; by that account virtue is knowledge, i.e. knowledge of how to achieve *eudaimonia*, whereas on the analogy virtue is *eudaimonia* itself. If the theory were to fit the analogy, knowledge of how to achieve *eudaimonia* should be analogous to knowledge of how to achieve health, i.e. to medicine rather than to health itself. But then virtue would be of purely instrumental value, whereas the analogy with health represents it as having intrinsic value. A further difficulty is this, that the cognitive account of virtue depends on the thesis that a cognitive grasp of the good is sufficient to motivate the agent to achieve it, a thesis which, though notoriously beset by the counter-evidence of ordinary experience, is defended by nothing more than a single unsound argument (*Meno* 77–8). When we turn to the *Republic* we find Plato developing a more elaborate psychology which enables him at the least to make a serious effort to remedy these deficiencies, in that it not only provides the material for the necessary account of *eudaimonia* but also allows him to abandon the counter-intuitive claims that virtue is knowledge and that it is impossible to act contrary to one's conception of one's overall good.

ii The *Republic*

Among the many ways in which the *Republic* is innovative is its attempt at a comprehensive integration of individual psychology with political theory. That there was some connection was not, of course, a novel idea; naturalistic theories such as those of the *Protagoras* myth had shown how social institutions including morality would naturally have developed in response to the needs of individuals for protection and cooperation. According to these theories morality was therefore both social, i.e. a set of social norms, and natural, i.e. grounded in individual human needs. We saw, however, that these theories failed to show that the observance of morality by any particular individual is an intrinsic good to that individual. Plato's innovation in the *Republic* may be seen as an attempt to bridge the gap between the individual's good and that of the community by internalising the social nature of morality, in that the individual personality is itself organised on a social model, and its best state, which is the supreme good for the individual, consists in a certain social organisation.

However counter-intuitive, the social conception of individual morality is not an arbitrary construction, designed to fill a gap left by earlier theory, but has a firm theoretical base. This may be set out as follows.

1 Key evaluative predicates such as 'good', 'just', 'courageous' and 'self-controlled' are applicable to communities as well as to individuals (368e–369a).
2 Any predicate which applies both to an individual and to a community applies to the one in virtue of the same feature or features as those in virtue of which it applies to the other (435a–b).
3 Since the perfectly good individual is wise, self-controlled, courageous and just, by 2 the perfectly good community is wise, self-controlled, courageous and just (427e).
4 The perfectly good community is just in virtue of the fact that the members of the three functionally defined classes into which it is divided (rulers, military auxiliaries and economic producers) stick to the social function which defines their respective class, and to which they are fitted by their natural abilities, developed by appropriate education (433a–434c).
5 The psychology of every individual comprises a tri-partite structure of intellect, self-assertive motivation and bodily appetite corresponding to the political structure of the perfectly good community (435e–441c).
6 Therefore, by 2, 4 and 5, the perfectly good individual is just in virtue of a relation between the three elements of his or her personality corresponding to that between the classes in the perfectly good community which constitutes the justice of that community (see 4) (443c–444a).

In this derivation premise 2 has a pivotal role, mediating the inferences from the character of the individual to that of the community (step 3) and conversely (step 6). It is an *a priori* thesis, which applies to the case of the community and the individual[24] the Socratic thesis (see above, p. 55) that all the things to which a single predicate 'F' applies share a single common nature in virtue of which they are all Fs. Unfortunately for Plato the thesis is false; even leaving aside cases of simple equivocation like 'pen' or 'cape', a predicate may apply to things of different kinds, not in virtue of the fact that all the things to which it applies share a common nature, but in virtue of the fact that each kind of application relates differently to a central notion.[25] Ironically, 'just' provides a clear example. A political community is just if

[24] The crucial sentence (435a5–7) is 'So, if things larger and smaller are called the same, are they alike in the respect in which they are called the same, or unlike?' It is assumed without argument that the community and the individual are examples of 'things larger and smaller'.

[25] As is well known, Aristotle discovered this kind of application. See *Met.* 1003a33–b12, *EE* 1236a16–23: for discussion see Owen [928].

either it is internally organised according to just principles, or in its relations with other communities it acts according to just principles. An individual, however, is just only in the latter way, not in the former, since the notion of just principles has no application to the psychological organisation of an individual. That is because just principles assign rights and obligations to the individuals composing a community, whereas the elements in an individual's psychological organisation are not themselves individuals, and are therefore not subjects of rights and obligations. Even this very crude sketch shows that the justice of a principle is the primary application of the predicate 'just', and that communities and individuals are derivatively just in virtue of different relations to the primary application. Plato's error is twofold, first in assuming the univocity of the predicate, and secondly in applying to the individual member of a community an internal, structural model of justice appropriate to the complex social entity, not to its individual components.

The identification in step 4 of adherence to one's generic social role as the organisational principle of justice in a community may also seem very contrived, but this too has some theoretical backing. First, it depends on step 3, and therefore on premise 2, on which 3 relies to identify the traditional list of *individual* virtues (see above, p. 52) as the virtues of the perfect community. It therefore succumbs to the refutation of 2 in the preceding paragraph. But waiving that objection for the sake of argument and granting that the excellence of the perfect community consists in wisdom, courage, self-control and justice, what grounds the identification of justice with adherence to one's generic social role? Here Plato appeals to the traditional conception of social justice as each one's having his own and doing his own, i.e. that each individual should be secure in the possession of what he or she is entitled to and should not encroach on the entitlement of another (433e–434a). Traditionally this expresses an individualistic principle of ownership, but Plato transforms it into a collectivist principle of service to the community: what belongs to one is above all the contribution one makes to the common good, and to be treated unjustly is to be deprived of that contribution, and thereby of the good itself, which can be realised (for all) only if each makes (and *a fortiori* is allowed to make) his or her specific contribution to it (434a–c).

Social justice is thus redefined, via the 'doing one's own and having one's own' principle, as adherence to optimum social organisation. I shall here assume the conclusion for which I have argued elsewhere,[26] that the criterion of the optimum social organisation is the maximisation

[26] In my [368].

of *eudaimonia*, understood as the provision for every member of the community of the conditions either for the realisation of *eudaimonia* or for as close an approximation to it as the limitations of the individual's psychological capacities allow. Given that account of social justice, it follows by premise 2 that individual justice is optimal psychological organisation, which is that very state of *eudaimonia* which it is the function of social organisation to make as widely available as possible. But the claim that the good for the individual, which it is the aim of social organisation to realise, is optimum psychological organisation advances us little beyond the truism that the good is living well and doing well. It makes some advance, locating doing well in the possession of a certain psychological state, rather than in the possession of external goods, but Heraclitus (D–K 22B119) and Democritus (D–K 68B170–1) had already said as much. If the theory is to say more, and in particular if it is to provide the vindication of morality which eluded the early dialogues, it must (a) give an account of optimum psychological organisation which is both informative and acceptable and (b) justify the claim that that account is an account of individual *justice*.

Since space is lacking for even the most summary account of the psychology of the *Republic*,[27] I must be content with dogmatic statement. There are three principal[28] elements in the personality, the intellect, the bodily appetites, above all those for food, drink and sex, and a loosely defined cluster of motivations which Plato calls 'spirit' (*thumos* or *to thumoeides*), including anger, shame, ambition and a sense of honour or self-respect, all of which may be understood as aspects of a fundamental impulse of self-regard and self-assertiveness.[29] The intellect is not a purely ratiocinative faculty, but has its own motivations; hence the tri-partition is at least in part a distinction between three kinds of motivation, towards intellectual activity, self-realisation and bodily satisfaction respectively. But in addition to providing its own specific motivations the intellect has the function of directing and co-ordinating the activity of all three kinds of motivation with a view to the realisation of the agent's overall good, since only the intellect is capable of the grasp of the good presupposed by that direction and co-ordination. There are therefore two ways in which the intellect is supreme in the state of optimal psychological organisation. First, all the agent's specific desires are directed by the intellect with a view to the agent's overall good, and secondly, that good consists in a life in which the satisfaction of the specific desires of the intellect (i.e. desires for

[27] See, for instance, Woods [353].

[28] *Rep.* 443d7 explicitly leaves open the possibility that there may be others.

[29] See Gosling [288], ch. 3. Cf. Rawls [966].

intellectual activity) takes priority over the satisfaction of the other kinds of desire. The satisfaction of intellectual desire should be understood as neither purely theoretical nor (because purely theoretical) exclusively egoistic; according to the metaphysical system of the *Republic*, the supreme object of understanding is the Form of the good, and someone who grasps what goodness itself is is thereby motivated to realise it not merely in his own life but (by the theory of love of the *Symposium* and *Phaedrus*) in the lives of those he loves and in the community of which he is a member.

A central feature of this psychology is Plato's abandonment of the theory of uniform motivation which was presupposed in the early dialogues, and with it the Cognitive Theory of Excellence and the strong version of the unity of virtue which that theory implied. Plato no longer accepts either the strong thesis that every intentional action is aimed at the realisation of the agent's conception of his or her overall good nor the weaker thesis that whenever that conception is present it motivates, since the bodily and 'spirited' appetites motivate independently of, and even in opposition to, the conception of the overall good. Hence what makes the difference between the virtuous and the non-virtuous agent is not simply the possession by the former of a cognitive state which the latter lacks, but the possession by the former of a psychological structure lacking in the latter, in which the specific desires are appropriately responsive to the direction of the intellect. And that responsiveness is not guaranteed by the content of the intellect's direction, but requires that the specific desires should have been conditioned by the process of education described in books II–III to respond instinctively to the guidance of the intellect by loving what the intellect reveals as good and hating what it reveals as bad. Since what is manifested in the various types of virtuous conduct is no longer a single cognitive state, the former version of the unity thesis has to go. But since by the new theory it is the same psychological structure which is manifested in those types of conduct, it might seem that the strong theory could survive the shift from cognitive state to total structure, with 'courage', 'wisdom', etc. functioning as non-synonymous names of that total structure.

In fact the shift is more substantial, for two reasons. First, it seems that those names apply, not to the same total structure under different aspects, but to different aspects of the structure; thus self-control is 'a certain ordering and mastery of pleasures and appetites' (430e) while justice is the state in which each element plays its proper part in the optimal structure, wisdom is the care of the intellect for the whole (441d–e), and courage is the retention by the spirit, despite pleasures and pains, of the instructions of the intellect about what is to be feared and what is not (442c). While justice and self-control are hard to distinguish on this account (except that the terms are

non-synonymous), courage and wisdom are not the total structure itself but aspects of it, ascribed in the first instance to particular elements of that structure. Secondly, the possibility of disorderly appetite indicates that while self-control and justice are not only mutually necessitating but severally impossible without wisdom, the possession of wisdom does not guarantee self-control. It appears, then, that the psychology of the *Republic* requires the abandonment of the unity of virtue doctrine, in that even the weakest form of that doctrine, the thesis that anyone who possesses any of the virtues necessarily possesses them all, has to be replaced by a still weaker thesis, namely that courage, self-control and justice all require wisdom (and perhaps require one another also), but wisdom does not guarantee the presence of the other virtues.

Does Plato give us adequate reason to accept that this psychological organisation is in fact optimal? To do so he has to show that each of the elements in the psychological structure functions at its best when co-ordinated by the intellect so as to make the appropriate contribution to a life where the highest priority is given to the pursuit of intellectual satisfactions. We may accept on the basis of the argument against Callicles (see above) that a satisfactory life is possible for the individual only if his or her potentially conflicting motivations are intentionally co-ordinated, which is a rational process requiring the identification of priorities and long-term goals. But we need further argument to show that the supreme long-term goal of the optimal life must be the theoretical understanding of reality (of which the primary object of understanding is goodness), and the realisation of that understanding in practical, including political, life. Plato attempts to meet this challenge by arguing in book IX that the life devoted to those goals is the pleasantest life possible, but his arguments are unsatisfactory. He gives two main arguments: the first (famously recalled by Mill in *Utilitarianism*) is that the devotee of intellectual pleasures is the appropriate judge of which life is the pleasantest, since he has experience of the pleasures of appetite and ambition (the dominant goals of the rival lives), whereas the adherents of those pleasures lack experience of intellectual pleasures. This argument fails because experience of the pleasures of a life requires commitment to the activities and the values constituting that life, but the intellectual is as remote from immersion in the rival lives as the rivals are from the intellectual life. The second argument depends on the conception of pleasure as the making good of a deficiency in the organism (for example hunger is a state of bodily depletion and ignorance a state of intellectual depletion, and the pleasures of satisfying hunger and of discovery are the processes of making good the respective depletions). Plato uses this model to make two basic points, whose relation to one another is

obscure; the first is that since the state of bodily depletion is painful, what we think of as bodily pleasures are in fact mostly episodes of getting rid of pain, not genuine pleasure. The second is that whereas bodily deficiencies cannot be properly or genuinely made good, intellectual deficiencies can be. Leaving aside the question of the adequacy of the depletion model in general, its application in this argument is highly obscure, since it is unclear whether the inferiority of bodily pleasures is supposed to lie in the fact that they require to be repeated (since, for instance, one gets hungry again a few hours after having eaten) or in the alleged insatiability of the desires which they presuppose (so that one can never get enough food) or simply in the alleged confusion between bodily pleasure and the getting rid of bodily distress. If the point is the latter, Plato's diagnosis of the alleged confusion is itself confused, since the depletion model yields the result that the process of making good the depletion is pleasant, irrespective of whether the depletion is painful. Hence whenever the depletion *is* painful, getting rid of that painful lack will be genuinely pleasant, and there will be no confusion of genuine pleasure with something else, namely the getting rid of pain. If, on the other hand, the target is the insatiability of bodily desires, the alleged fact should be denied; normal bodily desires are not insatiable, unless 'insatiable' is reinterpreted as 'recurrent', in which case the point is after all the first, i.e. that bodily desires are recurrent whereas intellectual desires are not. But with respect to that point it is not clear that the need for recurrent satisfaction differentiates a life devoted to bodily satisfactions from one devoted to intellectual; no doubt a truth once discovered does not have to be rediscovered, but a meal once eaten does not have to be eaten over again, and an intellectual *life* will require repeated acts of thought (whether new discoveries or the recapitulation of truths already known) no less than a life of bodily satisfactions will require repeated episodes of bodily pleasure. (The point also applies to the pleasures of the life of ambition, which is for the purposes of the argument required to share the defects of the life of bodily pleasure, but which is in fact barely mentioned.)[30]

Plato does not, then, succeed in establishing the optimality of his preferred psychological structure. Does he fare any better in showing that that structure captures the nature of justice as a virtue of the agent? Since his account of justice is avowedly revisionary, he cannot be held to the requirement to show that the presence of that structure in an agent is necessary and sufficient for that person's being just by ordinary standards.

[30] For fuller discussion of the arguments of *Rep.* IX see Annas [325], ch. 12; Gosling and Taylor [53], chs. 6 and 17.2; Stokes [119].

Someone who conforms consistently to ordinary morality as it is depicted by Glaucon[31] is just by ordinary standards but not by Plato's, since his commitment to justice is conditioned, not by acceptance of the value of justice for itself, but by the belief that he could not succeed in doing what he would really like to do, viz. to promote his own interests at the expense of others. The most that can be required is that Plato should show that the presence of the structure is sufficient for the performance of a sufficient range of central cases of just dealing and the corresponding avoidance of unjust dealing. At 442e–443a Socrates asserts that the Platonically just agent will never commit any major crime such as theft, sacrilege, treason or adultery, and while he gives no argument we may concede that the control of appetite which characterises such an agent will make him proof against the standard temptations to such wickedness. Moreover, as we saw, his love for the good will make him concerned for the good of others and for that of the community. How, then, could he fail to be just by conventional standards?[32]

The flaw in the theory is that the structure itself defines the good for the agent; hence concern for the good of others and for that of the community is concern to maintain that structure in others and the corresponding structure in the community. Consequently just actions are redefined as whatever actions serve to create and preserve that structure, and unjust as whatever destroy it (443e). That redefinition clearly licenses substantial interference with the autonomy of others, with a view to the promotion of their own good (as redefined) or the good of the community; indeed, if the setting up of the ideal Platonic state required extermination and enslavement of whole populations, by this account such acts would be just (not merely permitted, but required). This is explicitly acknowledged in the text: at 540e–541a Socrates recommends the foundation of the ideal state by the process of expelling the whole population of an existing city over the age of ten, taking the children from their parents and bringing them up in the educational system which he has just described. Plainly the forced evacuation of an entire city and the enforced separation of a complete generation are acts of extreme violence, which could not in practice be perpetrated without considerable loss of life. As regards loss of autonomy, the subjection of the lowest class in the ideal state is complete enough for their state to be

[31] This assumes that ordinary morality is not so confused as to make consistent adherence to it impossible. It is not clear that Plato would accept that charitable assumption: the arguments against the accounts of justice proposed in *Republic* I by Cephalus and Polemarchus suggest that he thinks that ordinary moral beliefs are thoroughly confused.

[32] See Sachs [363]; Demos [357]; Vlastos [365]; Irwin [293], ch. 7, secs. 10–11 and 19.

described as one of slavery (590c–d); it is, of course, paternalistic slavery, since it is better for the lower classes to be enslaved to those who have their good at heart and who know what that good is than to be enslaved to their own lower nature and to mistaken conceptions of what is good for them. Yet it is slavery none the less, since in the last resort the direction of their lives rests not with their own intellect and will, but with those of the rulers. The Platonically just agent will not, therefore, be unjust from the vulgar motives of private gain or personal lust, but he will not be just for all that. The form of injustice to which he will be prone is something much more terrible, the enforcement of an ideology which, in virtue of its comprehensiveness and its redefinition of benevolence, admits no limitation in the name of individual liberty, and is therefore liable to press its claims to the extremes of tyranny.[33]

iii Developments subsequent to the *Republic*

The ethical theory of the *Republic* represents Plato's most sustained attempt to vindicate the claims of morality. Subsequent developments in his ethical thought narrowly conceived (i.e. as distinct from political theory, which requires separate treatment) amount to modifications of detail, not to any radical shift of view. They may therefore be dealt with briefly.

The second of the two stages which we have distinguished in the development of Plato's ethical thought was marked off from the earlier by the abandonment of the 'Socratic' theory of uniform motivation and hence by the rejection of the Cognitive Theory of Excellence and the thesis of the unity of virtue. In the dialogues subsequent to the *Republic* the non-uniform character of motivation is even more strongly emphasised, while in the *Statesman* he insists on the *dis*unity of virtue (see below, p. 74). These developments may be attributed in part to an increasing sense of the dichotomy, already present but not dominant in the *Republic* (611a–d), between the rational element in the personality, motivated by the good and only contingently and temporarily embodied, and the

[33] Plato is quite explicit in drawing the implication at *Statesman* 293c–d:

> [T]he only constitution worthy of the name . . . must be the one in which the rulers are . . . men really possessed of the scientific understanding of the art of government. Then we must not take into consideration whether their rule be by laws or without them over willing or unwilling subjects or whether they themselves be rich or poor men.
> No.
> They may purge the city for its better health by putting some of the citizens to death or banishing others. (tr. Skemp from [385])

A similar view is expressed at *Laws* 735b–736c.

non-rational spirit and appetites, which spring from the body and are motivated independently of the good.[34] To the extent that the reason is identified with the real self, spirit and appetite come to be seen rather as alien forces requiring to be kept in subjection by reason than as manifestations of a uniformly rational agency. This allows Plato, despite having abandoned the thesis that knowledge of what is best is sufficient for doing what is best, to continue to maintain the second of the Socratic paradoxes, that no one voluntarily acts wrongly. Contrary to the Socratic position, action against one's better judgement is possible; but such action is not voluntary, since the agent (= the rational self) is overwhelmed by external forces, i.e. the non-rational passions (*Timaeus* 86e, *Laws* 734b, 860–3). On the Socratic model, all purported cases of action against one's better judgement had to be explained as cases of intellectual error; Plato now recognises that some cases have to be explained by a mismatch between intellectual judgement and passion, but saves the doctrine by counting such mismatches as sources of involuntary action.[35] That saving move has its cost, in that it breaks up the unity of the agent, but Plato's dichotomy of rational and non-rational elements in the personality, itself a reflection of the more fundamental dichotomy of rationally apprehended reality and the imperfectly rational material world, already encourages that split. The split is, however, fatal to the project of vindicating morality by showing it to be constitutive of the best life for the agent. For morality concerns the embodied agent, whereas the best life for the agent is the discarnate life of pure thought (*Tim.* 90). The most that Plato could hope to provide by way of vindication of morality would therefore be the claim that immorality as ordinarily conceived hinders the achievement of that life (*Tim.* 90).

This fundamental dichotomy tends to drive Plato further from the doctrine of the unity of virtue: if wisdom is in the last analysis the activity of the immortal 'real self' and courage, self-control and justice different

[34] In the *Phaedo* the appetitive and spirited motivations are attributed to the body (66b–c, 68b–c), which is sharply distinguished from the (rational) soul, whose task is to master and control the body (80a). The rational soul is implicitly identified with the self: the survival of the self is the survival of the soul, and the task of the philosopher is to prepare for what is ordinarily thought of as death, but which is in fact the fullness of life, free from the distractions of the body (64c–69e). While it might seem obvious that on this model also the motivations of the body are independent of the good, it is not clear that that is so, since the body is also described as a source of illusions, i.e. false beliefs (81b, 83d). Hence the spurious morality of the non-philosopher, guided by a calculation of bodily pleasures (68b–69c), is, although undoubtedly a state of enslavement of the soul to the body, nevertheless not incompatible with the Socratic theory of motivation.

[35] R. M. Hare makes the same move in [948], ch. 5 (modified in [949], 23–4, 58–60).

aspects of the subjection of the mortal and non-rational by that self, then the connection between the virtues has been loosened to the extent that, so far from being in any sense the same virtue, they are no longer even virtues of literally the same subject. Rather, the perfection of the real self requires, contingently and temporarily, the co-operation of the mortal self. This co-operation consists in the exercise of the four cardinal virtues, wisdom, *sōphrosunē*, justice and courage (*Laws* 631c), as is brought out in a later passage (653a–c) describing the general aim of education. This is essentially that of the primary education of the *Republic*, so to train the motivational impulses by means of the basic stimuli of pleasure and pain that the child comes to like what reason dictates and to hate what reason forbids. The affective responses are formed before the rational judgements, but when the child is mature enough to form those judgements they agree with the content of the affective responses 'and this agreement as a whole is *aretē*' (653b6). This formula may plausibly suggest an account of what courage, *sōphrosunē* and justice have in common; each is an agreement between a specific motivational impulse and a rational judgement,[36] or (perhaps closer to the theory of the *Republic*) each is an aspect of a state of agreement between the agent's motivations taken globally and the deliverances of his or her reason. There is, however, still the problem of what wisdom itself has in common with those virtues whose essence consists in agreement between the non-rational and wisdom. It may be a sense of this difficulty which prompts Plato to say at the end of the work (963) that while it is not hard to see how the virtues differ from one another, it is a real problem to explain how states as different as courage and wisdom are one, i.e. to determine what they have in common; this problem he leaves unanswered.

I suggest that this problem reflects Plato's difficulties in fitting his account of virtue to his sense of the dichotomy between the mortal and immortal elements in the soul. That same difficulty is also, I think, reflected in another feature of the treatment of the virtues in the later dialogues, namely that Plato sometimes reverts to the earlier 'Presocratic' tradition of treating them as separable components of excellence, as when he asserts that the state in which justice, wisdom (*phronēsis*) and *sōphrosunē* are 'unified together with courage' is better than courage alone (630a–b), citing as an example of the latter, as Protagoras had done in opposition to Socrates (*Prot.* 349d), the

[36] Compare Aristotle's account of the conditions of correct choice in *EN* 1139a22–6: 'since excellence of character is a disposition concerned with choice, and choice is deliberative desire, for this reason the judgement (*logos*) must be true and the desire right, if the choice is to be good, and the one must say the same thing as the other pursues'.

courage of wicked and licentious soldiers (cf. 696b–e). The conception of courage as a non-rational impulse combating fear, to be found even in animals (cf. *Lach.* 197a–b), recurs at *Laws* 963e and at *Statesman* 306. In the latter passage the Eleatic Stranger goes out of his way to emphasise the *un*homogeneity of the specific virtues, by first stating the conventional view of courage and *sōphrosunē* as parts of total excellence, and then urging the 'unfamiliar' thesis (306b13) that they are hostile and opposed to each other, in the sense that courage, understood as an aggressive impulse, is opposed to *sōphrosunē*, understood as an impulse to quiet and unassertive behaviour. It is the task of the statesman to devise forms of education and political institutions which will harmonise these opposed impulses for the benefit of the individual and the community. Here once again we see Plato apparently reverting to conceptions of the specific virtues which were rejected in the early dialogues, courage as aggressiveness in *Laches* (197a–b), *sōphrosunē* as quietness in *Charmides* (159b–160b).[37] In so doing he ignores the crucial distinction between motivational drives and the proper organisation of those in the integrated personality which was one of the main achievements of his mature theory, and thereby generates a spurious paradox. Aggressiveness may indeed be opposed to quietness, but courage is not opposed to *sōphrosunē*, since both are aspects of a structure of motivations organised under the direction of the intellect.

[37] There is even a trace of these conceptions in *Meno* 88a–c, where Socrates, arguing that knowledge is the only unconditional good, includes the virtues of *sōphrosunē*, courage and justice (as well as learning, good memory, personal splendour 'and everything of that kind') among the things which are good only on condition that they are directed by knowledge (i.e. directed aright), but which are harmful if misdirected. But in fact this passage seems to draw the very distinction between motivational drive and virtue proper which is blurred in the later dialogues. For Socrates asks Meno whether, if it is not the case that the virtues are sometimes beneficial and sometimes harmful, they can be anything other than knowledge. For instance, if courage is not intelligence (*phronēsis*) but a sort of boldness, is it not the case that when someone is bold without thought (*nous*) he is harmed, but when he is bold with thought he is benefited? After citing the examples of *sōphrosunē* and learning Socrates concludes as follows (c1–5):

> So to sum up, all the undertakings and endurances of the soul result in *eudaimonia* if they are directed by intelligence, but in the opposite if they are directed by folly.
>
> So it seems.
>
> If excellence, then, is one of the things in the soul and it is necessary that it should be beneficial, it must be intelligence . . .

'Undertakings and endurances of the soul', i.e. the acts which one is prompted to by motivational drives, are clearly distinguished from excellence, which is identified as knowledge. It follows that courage, for example, which is acknowledged to be a part of excellence (see above), is not itself a motivational drive such as boldness, but is itself knowledge.

I confess to being puzzled as to why Plato should have blurred this distinction (which is central to Aristotle's theory of virtue as well as to Plato's own) in his later writings; all that I can suggest is that his sense of the discontinuity between the rational and immortal elements in the personality on the one hand and the non-rational and mortal on the other may have made him uncomfortable with accounts of the specific virtues as modifications of the latter by the directive activity of the former. According to the theory of the *Republic*, the specific virtues, while no longer a single virtue as in the early dialogues, are still virtues of a single subject; yet the argument of *Rep.* x.611b–d that since the soul is immortal and what is immortal cannot be composite the tri-partite soul is not the true soul leads to the conclusion that the tri-partite soul is not a genuine unity, but rather an adventitious agglomerate of disparate elements, like the sea-god Glaucus overgrown with shells and weed. If Plato takes that conclusion seriously, then there is no single subject for all the specific virtues; wisdom is an attribute of the immortal soul, and the other virtues attributes of the mortal elements, which may have made it easier to revert sometimes to the traditional view of them as non-rational impulses, requiring the direction of reason to attain the status of true virtues. Of course, one *need* not conceive of them in that way, since the dualistic conception of the soul still allows the alternative (closer to the Aristotelian view) that, for example, 'courage' is not the name of a non-rational impulse requiring to be modified by reason, but the name of the state of having that impulse properly modified by reason, which seems to be the view which predominates in the *Laws*. There seems, then, to be evidence of some vacillation in the conception of the virtues in the later dialogues, which may perhaps be explained by the increased influence in this period of the dichotomy between the rational and non-rational in Plato's view of the soul.

The above account of the 'moral' virtues as consisting in agreement between affective responses and rational judgement, produced by the pre-rational training of the non-rational elements in the personality, has close and obvious affinities with Aristotle's theory, on which it was doubtless an influence. Another similarity between Plato's later theory and Aristotle's is found in the section of the *Laws* (660e–663d) where the Athenian Stranger discusses the requirement in an adequate code of legislation to show that the good man will be *eudaimōn* and the wicked wretched. Since the point of this provision in a code of legislation is to motivate people to obey the law, it is assumed, in line with the generally hedonistic account of motivation which is taken for granted in the *Laws* (see, for example, 636d–e) that the appropriate way to show that the good agent will be *eudaimōn* is to show that his or her life will be pleasant, and the rival lives unpleasant. The strategy is that of

Rep. IX, but the arguments have none of the metaphysical elaboration of those employed there. Instead, the Stranger argues simply that everyone sees his or her own preferred life as pleasantest from its own perspective, and that the correct perspective from which the assessment should be made is that of the virtuous agent (663b–d). No attempt is made to support this principle; Plato does not, for instance, employ the analogy, to which Aristotle sometimes appeals (see, for instance, *EN* 1173b22–5, 1176a8–22), with the perception of the healthy as the criterion of correctness in judgements of sensible qualities such as colour or taste.

The central theme of this chapter has been Plato's attempt to anchor morality on the 'natural' side of the nature–convention dichotomy by grounding it in an adequate theory of human nature. That theory develops from the optimistic over-simplification of the early dialogues to the more complex psychology of the middle and later periods, leading to a picture of the virtuous agent as one who achieves through intellectual and emotional training the right fit between intellectual judgement and affective response. That picture of the ideally developed agent was perhaps Plato's most important legacy to his successors, from Aristotle through the Stoics to Christian theorists and their post-Christian followers. But the theory underlying the picture faced the problems (a) of giving an account of the element of intellectual judgement which would justify its claim to truth and (b) of defending the integrity of the personality against the threat of dualism, in which the intellect arrogates value to itself, and the affective elements in the personality are correspondingly devalued along with the body. Plato himself solved neither problem: both are visible in Aristotle, Stoicism and Christianity, while at least the former remains to trouble their 'realist' successors.[38]

[38] The final version of this chapter was submitted to the editor in December 1990.

© Cambridge University Press, 1998