MERE APPEARANCES

Appearance, Belief, & Desire in Plato's *Protagoras*, *Gorgias*, & *Republic*

DAMIEN STOREY

A THESIS SUBMITTED FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY AT THE
UNIVERSITY OF OXFORD
MICHAELMAS
2012

For my father, David Storey

Contents

Abstract

'Mere Appearances: Appearance, Belief, & Desire in Plato's *Protagoras*, *Gorgias*, & *Republic*'

Damien Storey, Somerville College
Submitted for the degree of DPhil, Michaelmas 2012

This thesis examines the role appearances play, with notable continuity, in the psychology and ethics of Plato's *Protagoras*, *Gorgias*, and *Republic*. Common to these dialogues is the claim that evaluative appearances are almost invariably false: what *appears* to be good or bad is typically not in fact so and what *is* good or bad typically does not appear so. I argue that this disparity between apparent and real value forms the basis of Plato's diagnoses of a wide range of practical errors: psychological phenomena like *akrasia*, mistaken conceptions of the good like hedonism, and the influence of cultural sources of corruption like oratory, sophistry, and poetry. It also, relatedly, forms the basis of his account of lower passions like appetite, anger, or fear. Such passions are especially prone to lead us astray because their objects—appetitive pleasures like food, drink, or sex, for example—present especially deceptive appearances. One of the principle aims of this thesis is to show that this presents a significant point of agreement between the psychologies of the *Protagoras*, *Gorgias*, and *Republic*. In all three dialogues, I argue, motivational errors result from a specific kind of cognitive error: the uncritical acceptance of appearances. Plato's early and middle psychologies differ in their account of the subject of this error—in the *Protagoras* and *Gorgias*, the whole person; in the *Republic*, the appetitive or spirited part of a person's soul—but not in their basic theory of how our passions arise or, crucially, why they are liable to motivate us towards harmful ends.

Acknowledgments

This thesis has benefited greatly from discussion with many students and faculty members at Oxford. For comments on drafts of chapters, I would like to thank David Charles, Lindsay Judson, and my supervisor, Terence Irwin. I would especially like to thank Jessica Moss: her work, her guidance, and her endlessly kindness have had a profound affect on both the shape of this thesis and the quality of my years as a DPhil student. I thank my mother, Mary Hanley, for her enormous support throughout my studies. Finally, I am very grateful to my brother Aonghus Storey for his help dotting the i's and crossing the t's in the final version. Any errors that remain are, of course, entirely my responsibility.

Introduction

Plato has a pessimistic assessment of our ability to live virtuous and happy lives. For the most part we err, pursuing what harms rather than benefits us and vice rather than virtue. Why are we so prone to error? What might immediately come to mind is the extremely demanding requirements that Plato places on virtue and happiness, requirements that most if not all people fail to meet: living well requires both ethical knowledge, which is largely the preserve of philosophers, and, by the time he writes the *Republic*, a carefully orchestrated education, from early youth, of the non-rational aspects of one's soul. This, however, is Plato's account of the cure for error, not his diagnosis of its cause. That is, it does not tell us why Plato believed that without this demanding education an ordinary person's beliefs and desires fare so badly, leading them to unhappiness and vice. This thesis aims to show that Plato does have an answer, an answer that stems from a simple but highly explanatory claim: when it comes to questions of value, we live in an inherently misleading world—what appears good is typically not so, and what is good typically doesn't appear so.

In one sense, few would disagree that 'appearances can be deceiving': a cursory, *prima facie* assessment can of course lead us to believe something that turns out to be false on closer inspection. Plato's concern, however, is with a more literal and systematic disparity between appearance and reality. Plato believes that, without knowledge, we are prey to a kind of evaluative illusion, one that he often compares to optical illusions: certain harmful things appear to us to be good—are presented to us as good—just as a straight, submerged stick appears to us to be bent. Like optical illusions, they do not appear good because we judge them to

be so, but are rather a cause of misjudgement, inviting us to believe that they are as they appear and so leading us to mistake bad for good things. Only by going against this deceptive evaluative grain, questioning what very much seems to be the case, can we hope to live well.

The paradigm case for Plato is pleasure and pain. More specifically, the kind of sensory or bodily pleasures and pains that are typically associated with our basic appetites and aversions, the pleasures, for example, of food, drink, and sex. Such pleasures are naturally deceptive, appearing good—and highly persuasively so—even when they are harmful. Since pleasures or pains attend or result from almost every action we might perform, they are also a pervasive influence on our evaluation of what we ought to do. Consequently, such appearances are the primary cause of human error—not, in the first instance, our *desire* for pleasure, but pleasure's ability to deceive us (or a part of our soul) into thinking that that it is worthy of pursuit.

I examine the role appearances play, with notable continuity, in the psychology and ethics of Plato's *Protagoras*, *Gorgias*, and *Republic*. I argue that this disparity between apparent and real value forms the basis of Plato's diagnoses of a wide range of practical errors: psychological phenomena like *akrasia*, mistaken conceptions of the good like hedonism, and the influence of cultural sources of corruption like oratory, sophistry, and poetry. It also, relatedly, forms the basis of his account of lower passions like appetite, anger, or fear. Such passions are especially prone to lead us astray because their objects—appetitive pleasures like food, drink, or sex, for example—present especially deceptive appearances. One of the principal aims of this thesis is to show that this presents a significant point of agreement between the psychologies of the *Protagoras*, *Gorgias*, and *Republic*. In all three dialogues, I argue, motivational errors result from a specific kind of cognitive error: the uncritical acceptance of appearances. Plato's early and middle psychologies differ in their account of the subject of this error—in the *Protagoras* and *Gorgias*, the whole person; in the *Republic*, the appetitive or spirited part of a person's soul—but not in their basic theory of how our passions arise or, cru-

cially, why they are liable to motivate us towards harmful ends.

My first two chapters examine the role appearances play in two dialogues, the *Protagoras* and the *Gorgias*, that defend—or so I argue—the psychological theory known as Socratic 'intellectualism', a theory that denies that desire has any independence from belief: we desire something if and only if we believe it to be good; if we desire something bad, this is only because we have mistaken it for something good. Intellectualism must, then, reject a popular and very plausible explanation for error, that sometimes our actions are guided not by our beliefs but by unruly passions. I argue that in the *Protagoras* and *Gorgias* Plato develops an alternative account of error that gives appearances, specifically the appearance of pleasure and pain, the explanatory responsibilities that are popularly given to putatively belief-independent, non-rational passions.

Chapter 1 looks at Socrates' account of the cause of the experience of *akrasia*: the 'power of appearance' that immediate pleasures and pains possess, the power to appear better than they really are and, therefore, to cause us to misestimate their value. I present arguments for two claims: that the 'power of appearance' is held by only by short-term bodily pleasures like food, drink, and sex and that Socrates also recognises a second kind of pleasure, the pleasure of virtue, that, despite being superior, lacks bodily pleasure's more obvious and immediate appeal. This distinction allows Socrates to explain why the akratic error is, as the many assume, caused by the attraction of specifically appetitive pleasures. Unlike the many, however, he locates their attraction in the deceptive nature of pleasure's appearance rather than the non-rational nature of our desires for pleasure.

Chapter 2 offers an interpretation of the *Gorgias* as a dialogue that examines what we are liable to believe if we lack knowledge of what is good and bad and, therefore, are compelled to judge by appearances. Socrates answer is pessimistic: without knowledge, we are liable to be persuaded by the apparent value of pleasure and pain and unmoved by the real values of virtue and vice. The danger this poses is illustrated by Polus and Callicles: their refusal to countenance a difference between what seems and what is good leads them to endorse a life that Soc-

rates deems 'by nature the greatest and first of all bad things' (479D4–6), namely the hedonistic, unjust life of the tyrant. I argue that, despite passages that seem to anticipate the *Republic*'s psychology, the *Gorgias* is in fact consistently intellectualist, developing further the account of appetite that was suggested in the *Protagoras*: appetites arise through an uncritical acceptance of the appearance of appetitive pleasures like food, drink, and sex.

The remaining chapters turn to the *Republic*. While Plato's treatment of appearances in the *Republic* inherits the complexity of its psychology and metaphysics, it remains remarkably continuous with the account found in the *Protagoras* and *Gorgias*. Most surprisingly, despite Plato's rejection of intellectualism in the *Republic*, his account of lower passions like appetite, anger, and fear remains in one very significant respect the same: lower, unruly passions are those that arise from an uncritical acceptance of mere appearances. The appetitive and spirited parts of the soul have their distinctive, often reason-opposing passions precisely because they are incapable of reasoning and are, therefore, compelled to judge by appearances. Hence, Plato's basic theory of practical error remains the same.

Chapter 3 defends the account of non-rational cognition that forms the basis of my interpretation. I offer a careful reading of a passage in book 10, 602C–603A, that presents an argument for the partition of the soul. In structure this argument is very similar to the better-known arguments for partition in book 4, but it begins from cases of cognitive rather than motivational conflict. When we encounter an optical illusion, Socrates claims, a part of us uses 'reasoning' or 'calculation' (*logismos*) to conclude that the appearance is false but another part of us that is incapable of employing calculation believes, at the same time, that the appearance is true. My interpretation of this argument centres on two claims. With a number of recent commentators, I argue that this is a division not, as has been suggested, within the rational part of the soul but, as in book 4, between the rational part and one or both of the appetitive and spirited parts. Second, I address the following interpretive difficulty: the argument of 602C–

603A relies on an unexplained inference from an opposition between belief and appearance—between belief and how an optical illusion looks—to the presence of two opposing beliefs. I argue that this inference is valid, and plausible, only if Socrates treats the opposing appearance as itself a belief-like state: just as beliefs represent the world to the rational part of the soul, appearances represent the world to the non-rational parts.

Chapter 4 looks at another passage from book 10, 595C–598D. Appearances are, it seems, part of the basic furniture of Plato's metaphysics in the *Republic* (thus they have a place on book 6's Divided Line) and in 595C–598D Socrates explains their position in this metaphysics: just as ordinary objects and properties are a kind of unfaithful copy of the Forms, appearances are an unfaithful copy of ordinary objects and properties. They are, then, 'at three removes from the truth' (597E6): they are copies of copies of Forms and they suffer a degradation of 'truth and reality' at each removal. I argue, against some popular readings of this passage, that Socrates' conclusion is that appearances are by their nature deceptive, a claim that is essential to his account of why poetry's imitative appearances are corrupting.

Chapter 5 turns to the role appearances play in the *Republic*'s ethics. Chapter 3 establishes an account of non-rational cognition: the non-rational parts of the soul, though incapable of reasoning, acquire information about the world through appearances. Chapter 4 explains why the information imparted by appearances is by its nature unreliable. These results lead us, I argue, to the following connection between non-rational cognition and non-rational conation: the appetitive and spirited parts of the soul have their distinctive passions, passions that are often at odds with the calculated aims of the rational part, precisely because they are compelled to judge by unreliable appearances. Hence, the cognitive conflict described in 602C–603A explains the motivational conflict described in book 4: the passions of the spirited or appetitive part can conflict with those of the rational part because the former follow appearance and the latter follows calculation. This chapter develops and substantiates this account through an examin-

ation of two of the central ethical discussions of the *Republic*: the description of lowest level of ethical education in the Allegory of the Cave and the criticism of imitative poetry in book 10. I argue that these two discussions are, respectively, an allegorical and a literal account of appearance's corrupting effect, an effect exacerbated by 'image-makers' (*eidôlou dêmiourgoi*)—'puppeteers' in book 7 and poets in book 10—who exploit our non-rational parts' susceptibility to appearances. I argue, first, that the cognitive level of *eikasia* in books 6 and 7 is the *Republic*'s account of judging by appearances, the same limited form of cognition that we find in poetry's audience in book 10. I turn then to a problem shared by books 6–7 and book 10: what is the relation between sensory appearances like shadows and optical illusions and evaluative appearances like 'shadows of justice' (517D8–9) or 'images of virtue' (605E5)? Finally, I offer an account of Socrates' 'greatest charge' against poetry, that it can corrupt even the 'decent man' (*epieikês*), despite his recognition of a poem's unethical content. Through a careful analysis of 605C–606B, the description of the decent man's conflict when he encounters a poem, I argue that deceptive ethical appearances corrupt from the bottom up: affecting first the lower, appearance-sensitive parts of a person's soul and then, through their corrupting influence, persuading the whole person of ethical falsehoods.

CHAPTER 1

Appearances & *Akrasia* in the *Protagoras*

In this chapter I offer an interpretation of the claim that immediate pleasures and pains possess a certain 'power of appearance' (*tou phainomenou dunamis*), a power that makes them deceptive, appearing more pleasant or painful than they really are. I argue that Socrates' appeal to the 'power of appearance' is not just an appeal to the idea that appearance can be deceiving—that an unreflective, *prima facie* assessment can be mistaken—but the more specific claim that there is a certain consistent, recalcitrant disparity between what appears good and bad and what is good and bad. Central to my interpretation is the claim that the power of appearance is held by only one of two kinds of pleasure that Socrates recognises in the *Protagoras*: sensory or bodily pleasures, epitomised by the pleasures of food, drink, and sex. I argue that he also recognises a second kind of pleasure, the pleasure of virtue, that lacks the power of appearance. This distinction between kinds of pleasures—some that do and some that do not exert the power of appearance—allows Socrates to address, in a precise way, one of the central beliefs in the popular conception of *akrasia*, namely that we have a special kind of desire that leads us astray in such circumstances: non-rational and especially appetitive desires for pleasures like food, drink, or sex (and the corresponding passions related to pains: aversions and emotions like fear). While it is commonly thought that such pleasures lead us astray by being the objects of non-rational passions, Socrates argues that they in fact lead us astray by appearing more attractive than they really are. In short, Socrates replaces the motivational push of

non-rational desires with the epistemic pull of false appearances of the good.

I begin in §1.1 with two related preliminary tasks. First, I offer an explicit statement of the psychological theory known as Socratic 'intellectualism' and, against some common objections, I argue that this is the psychology that Socrates defends in the *Protagoras*. Second, I try to give a clear account of Socrates' general strategy in his treatment of *akrasia*. Motivated by his commitment to intellectualism, Socrates' aim is in part to deny that *akrasia* occurs, at least as it is commonly understood. I emphasise that there is also a constructive side of his argument that aims to give intuitively satisfying alternatives for the unruly appetites and motivational conflict that seem to be a part of *akrasia*.

In §1.2 I argue that Socrates' appeal to the 'power of appearance' should be seen as his attempt to give an alternative explanation for the unruly appetites that intellectualism precludes. I begin by looking at Socrates' alternative account of the akratic experience, in which he argues that it is a miscalculation caused by a certain illusion: immediate pleasures falsely appear larger than equal but later pleasures (356C8–E4). I argue that this passage should be understood in relation to two other passages. First, Socrates' discussion of so-called 'bad pleasures' (353C3–354E2), which gives us a clearer idea of what immediate and later pleasures are, and, second, his claim that courageously going to war is pleasant (359E1–360A2), which reveals a distinction between kinds of pleasures: bodily pleasure and the pleasure of virtue.

1.1 INTELLECTUALISM & *AKRASIA*

1.1.1 *Socratic intellectualism*

The central claim in intellectualism is: (1) that we desire something if and only if we find it good and desire it because we find it good (the 'good' here is prudential, that is, what is good for the desirer). We can add to this a further claim that gets considerable emphasis in the *Protagoras* (e.g. 356A5–C3): (2) that our desires motivate us towards what we find good optimally, that is, towards what

seems to be the *best* possible outcome. This second claim ensures that we don't desire equal good things unequally, pursuing certain kinds of goods more vigorously than others and, therefore, potentially choosing a lesser good at the price of a greater good (one kind of *akrasia*). Both (1) and (2) are expressed in terms of 'good' and 'desire' for the sake of simplicity; they should be taken to apply *mutatis mutandis* to the bad and to all motivational states, such as aversions and emotions (all of which I will at times bundle under the term 'passions').

These two claims get us very close to intellectualism. They are the claims that are explicitly defended in the early dialogues and often they, and even (1) alone, are thought to sufficiently characterise Socratic psychology. It is, however, instructive to notice that unless we exclude some further possibilities—although perhaps unusual or sophisticated ones—(1) and (2) fall short of intellectualism. By themselves, they don't necessarily preclude motivational conflict.

Since the good is an empty goal without some idea of what particular things are good—this glass of wine, this virtuous action, and so forth—desires need some way of having good things represented to them as such. They need to be 'told' by one or other cognitive faculty what counts as good. In (1) and (2) I have used the loose word 'find', where Socrates would use 'believe', to leave open the possibility that our desires might be responsive to finding something good through some faculty other than belief. It could be claimed—and in chapters 3–5, I argue that this is precisely what Plato claims in the *Republic*—that something can be represented to us as good through perception alone and, therefore, that our desires can be responsive to the mere sight of an apparent good, whether or not we believe what we see. Importantly, this is not a claim that Socrates' countenances in the *Protagoras*. Whatever 'appearances' are in the *Protagoras*, and however rich their contents may be, they cannot motivate desires independently of beliefs. Intellectualism, then, has a third claim: (3) that belief is the only faculty through which we can find something good. (A not unreasonable, but certainly notable assumption here is that we cannot knowingly hold beliefs that contradict. We cannot, then, experience conflicting desires as a result of holding,

simultaneously, opposing beliefs about what is best.)

If claims (1)–(3) are true, then motivational conflict is impossible. Desire and intellect invariably agree because desire aims, and can only aim, at what the intellect judges best. Before I turn to the *Protagoras*, one terminological distinction: throughout this thesis, I will call a desire 'rational' if it meets (1)–(3) (that is, after making the changes necessary for each to apply to an individual desire) and 'non-rational' if it fails to meet at least one of (1)–(3). Socratic intellectualism, then, can be reasonably summarised as the claim that all desires are rational or that non-rational desires are psychologically impossible.

In the remainder of this section I will argue that in the *Protagoras* Socrates defends—or at least takes himself to have defended—the psychological theory described by (1)–(3). To see this, we need to look at the end of our passage, where Socrates offers a carefully stated summary of the conclusion he has reached:

> Then if the pleasant is the good, no one who either knows or thinks (*oute eidôs oute oiomenos*) that there is something else better than what he does, something possible, will continue to do it, when he could be doing what is better. Giving in to oneself (*hêttô … hautou*) is nothing other than ignorance, and mastering oneself (*kreittô heautou*) is nothing other than wisdom. (358B6–C3)

> (a) Now, no one willingly (*hekôn*) goes toward the bad or to what he believes to be bad; (b) neither is it in human nature, so it seems, to want to go towards what one thinks bad instead of to the good. (c) And when he is forced to choose between one of two bad things, no one will choose the greater if he is able to choose the lesser. (358C6–D3)[1]

These two passages indisputably state at least that we *act* as we would expect if (1)–(3) were true. The first passage tells us that whatever action we do, it will be the action we know or believe to be best; if we know or believe that another action is better, we will do this better action.[2] In line (a) we are told that no

[1] Unless otherwise noted, all translations are taken, with greater or lesser alteration, from J.M. Cooper (ed.) *Plato: Complete Works* (Indianapolis: Hackett Publishing, 1997).

[2] While it seems fairly certain that this is what Socrates is illustrating, his example could be clearer: if a person believes there is something better than what he is doing, why was he doing it in the first place? Surely one can not start doing something that one can not continue doing? Socrates might mean to illustrate precisely this, namely the absurdity of his first clause: since one would always switch to the better action, one could never even begin the worse action. Alternatively, he might have in mind an action that is interrupted by a new belief: anyone doing one

one willingly goes towards either what is bad (what is in fact bad) or towards what he thinks bad (but might be good). This is a statement of the famous Socratic paradox that no one errs willingly.[3] This is usually understood to be the claim that no one errs *knowingly*. Much of Socrates' argument has been devoted to showing that while we often do bad things, we only do so in ignorance and therefore, he claims here, such actions are not done willingly. Line (c), however, shows that an additional requirement is needed: while knowledge is necessary for a willing action, it is not sufficient, since given a money-or-your-life choice we will knowingly but unwillingly choose something bad—albeit the best available option—because we are 'forced' (*anagkazesthai*) to choose between two bad options. In light of this, line (b) is best characterised as the claim that no one knowingly and freely goes towards either the bad or what they believe to be bad.

The first passage and line (a) and (c) of the second passage give us unambiguous statements of an intellectualism at the level of actions, but they tell us nothing explicit about our desires. Line (b) is the putative exception, but it is also a contentious one. The verb *ethelien* can be translated either 'to want' or—mirroring the *hekôn* in (a)—'to be willing'. As we have seen, an action that is done willingly is an action that is done knowingly and freely. A natural explanation for why we never knowingly and freely pursue bad things is that we never *desire* bad things, but this does not follow necessarily: strong desires that cause us to act against our better judgement might be thought of as a kind of force or we might have desires for bad things that we never directly act on.[4] Consequently,

---

thing who comes to believe that doing something else is better will immediately switch to the better action. In any case, unclear as the details are, on either interpretation the conclusion turns out to be the same.

[3] This was introduced earlier, in Socrates' analysis of Simonides' poem: 'For Simonides was not so uneducated as to say that he praised all who did nothing bad willingly (*hekôn*), as if there were anyone who willingly did bad things. I am pretty sure that none of the wise men thinks that any human being willingly makes a mistake or willingly does anything shameful or bad (*aischra te kai kaka*). They know very well that anyone who does anything wrong does so unwillingly (*aekôn*).' (345D6–E4)

[4] Santas (1966) defends the former. Devereux (1995), Brickhouse and Smith (2006) and (2007), and Singpurwalla (2006) defend the latter. See also Reshotko (2006).

only the former translation, 'to want', gives us an explicit claim about the nature of human motivation.

To decide between these two translations we need to look at the relation between claims (a) and (b). While these claims look quite similar, the latter is introduced by a 'neither' (*oute*) that tells us that it adds something substantial. If we choose 'to be willing' then any distinction between (a) and (b) will come down to either the latter's appeal to 'human nature' (*anthrôpou phusei*) or the fact that it adds the proviso that we don't go towards the bad 'instead of to the good'. The second option is unpromising: not only does this follow trivially from the fact that no one willingly goes towards the bad, but its purpose would seem to be to make it clear that the circumstances of the choice in *both* (a) and (b)—a choice between good and bad—is different from the circumstances of the choice in (c), where there are only two bad options. What about the appeal to human nature? Certainly this adds force to Socrates' claim, but on the 'to be willing' reading the force is purely rhetorical. That *no one* willing goes towards the bad should be enough to show that this is a fact of human nature—what else besides human nature might explain something that is universally true of the way people act?

On the other hand, if *ethelien* is translated 'to want', (b) is not only a new, substantial claim, but one that serves to explain claim (a): no one goes willingly towards the bad *because* it is simply not in human nature to want what is bad. As we would expect, our actions and choices are explained straightforwardly by the desires that motivate them and beliefs that guide them. Indeed such an explanation is sorely needed; without this claim—that is, if we allow that we sometimes desire what we believe to be bad—it seems entirely mysterious that we *never* act on these desires. After all, motivating actions is surely desire's *raison d'être*.

There are, then, good linguistic and philosophical reasons for taking (b) to be a claim about the nature of human desire. Nonetheless, it would be precarious to rest our understanding of Socratic psychology on the translation of just one

line (one line in some eight relevant Stephanus pages). However, immediately following the quoted passages there is a further, often missed piece of evidence that corroborates this reading. Socrates offers a definition of fear. If Socrates thought that there are any passions that can conflict with our better judgement, it seems certain that he would consider fear to be among them. In psychologies that allow non-rational passions, passions that can make us act against our better judgement, fear is typically, and rightly, a paradigmatic example. In fact, however, Socrates defines fear as a 'kind of expectation of the bad' (358D6–7): a *prosdokia*, a kind of forward-looking belief.[5] This clearly entails that fear will follow our evaluative judgements. This is confirmed a few lines later:

> If what I have said up to now is true, then would anyone be willing to go toward what he dreads, when he can go toward what he does not? Or is this impossible (*adunatos*) from what we have agreed? For it was agreed that *what one fears one holds to be bad;* no one goes towards those things which he holds to be bad, or chooses those things willingly. (358E2–6; my emphasis)

This shows decisively that Socrates believes that we fear only what we judge to be bad. Excepting the unlikely possibility that fear holds a unique place among the passions, it also shows decisively that Socrates takes all passions, all desires and emotions, to follow our judgements about what is good and bad. His intellectualism, then, is both desiderative and (as a result) behavioural: intellectualism as it has been traditionally understood.

[5] This might seem to reduce the apparently conative state of fear to a purely cognitive state—the affective, heart-thumping feeling of fear to a mere view about one's future fortune—but note that an almost identical account is found in the *Laws*, where it would be more surprising to find such a crudely cognitivist view of the passions: we have 'opinions about the future, whose general name is "expectation" (*elpis*). Specifically "fear", when it is an expectation of pain …' (*Laws*, 644C9–D1; cf. Aristotle, NE 1115A9). In both cases, it seems more likely that in defining fear as a kind of expectation Plato is not ignoring or denying fear's affective side, but simply pointing to a necessary, defining constituent of fear, the constituent that tells us why it is fear rather than some other passion. See also Segvic, who argues that the *tis*—fear is a *kind of* expectation—leaves the definition enough latitude to fit fear's affective side (2000, pp. 34–35). For a defence of the crudely cognitivist reading of passions in the *Protagoras*, see Kahn (1998).

1.1.2 *What Socrates does and does not deny*

Intellectualism entails that any form of motivational conflict is impossible. This makes it a highly controversial thesis. The strongest evidence against it is simply that it very much seems to be the case that we do experience motivational conflict, most notably in the common experience that we call *akrasia*. Socrates must deny that *akrasia*, so conceived, can occur. Many have felt that this is denying something that is obviously true. Vlastos puts this criticism forcefully: Socrates, he says, was 'too fascinated by the patterns into which he could organise his propositions to reflect with the needed sensitiveness and humility on matters that can only be learned from the facts themselves'; he attempts to 'excogitate by pure deduction a fact of human nature … the kind of matter of fact that can only be found out by observation' (1956, pp. xliv; xxxix). Vlastos takes Socrates to be denying that a common psychological experience occurs. If this is what Socrates is doing, then he is certainly mistaken. Someone can doubt that the experience we call *akrasia* has been correctly understood—indeed, it is still a live subject of debate—or even that it ought to be called *akrasia*, but that there is an experience that requires explanation can not be plausibly denied. To argue on philosophical grounds that it simply doesn't occur is to pit 'patterns of propositions' in a losing battle against observable fact.[6]

The argument I present in this chapter relies on a very different understand-

[6] Many have shared Vlastos's opinion. Aristotle described Socrates' denial of *akrasia* as an 'argument that contradicts things as they appear manifestly (*tois phainomenois enargôs*)' (NE 1145b27–28). Lemmon gives the following unequivocal assessment: 'It is so notorious a fact about human agents that they are often subject to acrasia that any ethical position that makes this seem queer or paradoxical is automatically suspect for just this reason. Of Socrates we can say that as a plain matter of fact he was just wrong—acrasia does occur' (1962, p. 144). More recently, authors have targeted the 'traditional' interpretation, i.e. any reading that takes Socrates to be committed to intellectualism. For example, Ferrari says that 'on the traditional interpretation, Socrates is simply spiriting away the phenomenology of *akrasia*' (1990, 136, n. 31). Similarly, for Devereux what Socrates defends on the traditional interpretation 'contains nothing that even *seems* like acting against knowledge or one's better judgement.… Simply denying that agents ever experience desires that conflict with their rational decisions would be completely unsatisfactory' (1995, p. 389). The most notable exception to this trend is Terry Penner, in many ways intellectualism's modern apologist. He has defended Socrates' treatment of *akrasia* in a series of articles: (1990), (1996), and (1997).

ing of Socrates' strategy in the *Protagoras*. In the next few pages I aim to show that a careful look at the structure and stated aims of Socrates' argument makes it clear that he fully appreciates why *akrasia* presents a *prima facie* difficulty for his position and that his argument is designed to address the empirical evidence that Vlastos believed he ignores. Whether or not we find his answer satisfactory, it is nonetheless an honest attempt to answer, rather than side-step, the doubts that many people will have about a psychology that denies that our desires can conflict. Most importantly, it is as much an attempt to show what *akrasia* is as it is an attempt to show what it is not.

What Socrates is specifically rejecting is the many's claim that 'often people who know the best action are not willing to do it; while it is possible for them, they do otherwise … [and] those who act that way do so because they are overcome by pleasure (*hêdonês hêttasthai*) or pain or are being ruled by one of those things I referred to just now [sc. anger (*thumos*), love, or fear]' (352D6–E2).[7] The many believe that people sometimes desire pleasure simply *qua* pleasant, whether or not they believe it to be good, and that such desires can cause us to act contrary to what we know to be best. There are two parts to Socrates' response, a negative and a positive part. I will look at the details of the positive part in the next section, but here I want to look at why Socrates finds it important. First, very briefly, the negative part:

By arguing for the identity of the pleasant and the good, Socrates reduces the possible objects of desire from two to one. The many think that they disagree with this on the grounds that they 'call some pleasant things bad and some painful things good' (351C3). However, by examining a number of examples of what they call 'good' pains and 'bad' pleasures, Socrates shows them that when they call something pleasant 'bad' it is only because 'it deprives us of greater pleasures than it itself provides or brings about greater pains than the pleasure it contains' (354C7–D1) and likewise, *mutatis mutandis*, for what they call 'good' pains. The

[7] Given that fear is later defined in terms of the bad, and the bad in terms of pain, we might reasonably expect a similar analysis of anger and love. All three, then, would be additional cases of being overcome by pleasure or pain.

many, it seems, reason as far as 'cautery is good because it leads to health', which appears to contradict hedonism, but fail to ask themselves why they value health. On examination, Socrates argues, they will find that they think health is good only because of the pains it relieves and pleasures it promotes. What compels the many to concede this point is the fact that they are simply *unable* to find any other end they have in mind when they call something good or bad other than the overall pleasure or pain that results: 'yet if you call enjoyment itself bad for some other reason and with some other end in view than the one I have suggested, you could tell us what it is—but you won't be able to' (354D1–3; see also 354D7–E2 and 354E8–355A2). Socrates draws the conclusion that 'good' and 'pleasant' are two names for the same thing and likewise 'bad' and 'painful' (355C3–B1). Since we can substitute co-referring terms *salva veritate*, 'being overcome by good' can be substituted for 'being overcome by pleasure' and, therefore, we see that the many's claim 'is ridiculous (*geloion*)—someone does what is bad, knowing that it is bad, when it is not necessary to do it, having been overcome by good things' (355D1–3).[8]

This is a bare outline of the argument, glossing over many of its subtleties and difficulties, but it is sufficient for our purposes. We can see that it is a purely deductive refutation and one that might well leave the many unsatisfied—this may be where the argument has brought them, but it can't spirit away the fact that they've *experienced* what very much seems like being overcome by pleasure. Socrates is well aware of this and he does not in fact think this argument alone can persuade them. Consider Socrates' own description of his aim and its rationale. He and Protagoras, he says, must try:

[8] A further point is clarified, albeit quietly, during the course of this argument. They do not desire pleasure simply *qua* pleasant, but because they believe it is good: 'so then you pursue the pleasant *as being good* (*hôs agathon on*) and flee pain *as bad*' (354C3–5; my emphasis). In other word, while the good and the pleasant are extensionally identical—every good thing is also a pleasant thing and *vice versa*—the good has intentional priority: we desire the pleasant because it is good but not *vice versa*. This does not necessarily follow from the identity of the pleasant and the good, but it would certainly be odd if we desired pleasure twice, once simply because it is pleasant and a second time because it is good. That we desire it for one reason only, because we believe it is good, is an inference to the best explanation. The many, then, implicitly accept (1): we desire something if and only if we believe it to be good and because we believe it to be good.

> … to persuade and teach the many what this experience (*pathos*) is, which they call being overcome by pleasures and on account of which they say one fails to do what's best, when knowing what it is. For if we said 'You do not speak correctly, people, but falsely', they would probably ask us: 'Protagoras and Socrates: if this experience is not being overcome by pleasure, then what on earth is it? What do you say it is? Tell us.' (352E6–553A6)

The negative, deductive part of Socrates' argument is the part that says 'you do not speak correctly, but falsely', but this is only one of two necessary steps. In addition to showing them what it is not, he also needs to teach the many what this experience *is*. Note that Socrates refers here, and throughout the argument, to the 'experience' (a *pathos*, something one suffers or undergoes) that the many call 'being overcome by pleasure'. Socrates never questions whether or not this experience occurs. By calling this experience 'being overcome by pleasure' the many propose a theory to explain what is happening when it occurs. The experience is the *explanandum* and 'being overcome by pleasure' is one possible *explanans*. The existence of the *explanandum* is an empirical fact. We know it occurs for the simple and certain reason that we experience it. If a deductive argument concludes, contrary to our experience, that there is nothing in need of explanation, then all the worse for the argument. What Socrates rejects is only the *explanans* that the many propose; he does not deny, but rather aims to explain, the *explanandum*.[9]

Socrates sets the agenda with the many's question: 'if the experience is not being overcome by pleasure, then what on earth is it?' (this question is repeated verbatim, and declared answered, at the conclusion of what I've described as the positive side of his argument, 357C6–D1). Given the tone of the question, it seems to both express incredulity and pose a challenge: '… what on earth is

[9] Should Socrates be described as denying that *akrasia* occurs? How one answers this question will depend largely on whether one thinks that the the word '*akrasia*' should apply to the *explanadum* (what we experience) or to a specific *explanans* (a theory of what this experience is) along the lines the many propose. A reasonable case can be made for both choices. I choose the former and will at times refer to Socrates' account of the *explanadum* as his account of *akrasia*. First, this exegetically simpler: if we do not call the *explanadum akrasia* it has no accepted name and we must resort to a cumbersome description or neologism. Second, I believe Socrates' description of the *explanadum*, discussed in the next section, retains enough of what we typically associate with *akrasia* to be considered a theory of what *akrasia* is.

it (*ti pot' estin*)? What do you say it is? Tell us'. To the extent that it expresses incredulity—they believe it must be being overcome by pleasure, that it could hardly be anything else—part of an appropriate answer will be to show them that their seemingly obvious explanation fails, which the negative side of Socrates' argument achieves. But the challenge posed by the question—to *answer* the question—plainly implies that the many will not be persuaded until Socrates also gives them a plausible alternative, an account of the experience that can satisfy their sense of what it feels like to undergo it. Socrates clearly thinks this is an important and difficult challenge: 'it is not easy to show you what this experience which you call being overcome by pleasure is *and yet it is on this that all the arguments depend*' (354E5–8; my emphasis). He also understands that it will take a careful argument to persuade the many. If he and Protagoras, in answer to the many's question, 'had straight away said 'ignorance', you would have laughed at us' (357D1–2). He believes, then, that prior to his argument his position seems laughable or absurd (*geloion*: the same description is applied, after his argument, to the many's position), presumably because it contradicts something that seems obvious to most people.

Indeed, Socrates' unusual choice of interlocutor, the collective 'many', is itself an indication that he has been careful to pose himself a fitting challenge. The many are a natural mouthpiece for common sense and Socrates' stance on *akrasia* particularly offends common sense. More specifically, Socrates' stance on *akrasia* not only contradicts what is commonly believed, as many of his claims do, but it is in danger of contradicting what is commonly *experienced*, which is considerably less acceptable.[10] If Socrates believes that any explanation of this experience must capture its phenomenology—which is really just to say that it is an explanation that fits the *explanandum* in question—then the many, who are guided by what experiencing *akrasia* feels like, are an ideal yardstick to meas-

[10] This answers an objection by Raphael Woolf (2002). Woolf argues that since Socrates' stance on *akrasia* is one of a number of Socratic paradoxes, so-called precisely because they go against what is commonly believed, the fact that it offends common sense is insufficient to explain his unusual choice of interlocutor. It is unique, however, in also going against what is commonly experienced.

ure his theory against. If he can plausibly be said to persuade them, by giving them an acceptable answer to the question they pose, he can be confident that his account is true to the familiar experience of *akrasia*—not contradicting 'the kind of matter of fact that can only be found out by observation' but fitting a very large set of people's observations, the observations of 'the many'.

These considerations make it clear that Socrates wishes to give an account of what *akrasia* is, an account that, however revisionary in other ways, is still recognisably an explanation of the experience we are familiar with. The most obvious challenge he faces, and the one that has received most scholarly attention, is to explain the fact that *akrasia* is an experience of inner conflict. I will focus on a different, but importantly related challenge: to explain the fact that *akrasia* seems to involve a conflict between two *kinds* of motivations, an opposition between motivations that support our considered view of what is best, on the one hand, and unruly passions like anger, fear, or appetite, on the other.

1.2 THE POWER OF APPEARANCE

Having listened to Socrates argue that, given hedonism, it is absurd to suppose that we act against our better judgement because we are 'overcome by pleasure', the many raise the following objection: 'But Socrates, immediate pleasure (*to parachrêma hêdu*) is very much different from pleasure or pain at a later time' (356A5–7). This might be the highly implausible objection that immediate pleasure is in fact more pleasurable than later pleasure, simply by being immediate, but it is unclear why the many would wish to make this claim: if their objection is to be relevant to the *error* of being overcome by pleasure, it shouldn't offer a reason to think that it is *right* to choose immediate over later pleasure. Read more charitably, the many intend the following, more powerful objection. We can be overcome by the pleasant even if the pleasant and the good are identical: a lesser but immediate pleasure can, despite our knowledge that it is lesser, stir a stronger desire than a greater but distant one, leading us to pursue the lesser

pleasure at the price of the greater pleasure, contrary to what we know to be best.[11]

Socrates' responses caters for both of these possible objections. He first asserts that the only important difference between pleasures or pains is the difference revealed by a calculation as impartial and objective as weighing: 'you put the pleasures together and the pains together on the balance scale, both the near and the remote, and say which of the two is more. If you weigh pleasant things against pleasant, the greater and more must be taken; if painful against painful, the fewer and smaller' (356B1–5). Socrates' next step is to explain why, despite there being no real difference, some people do indeed desire immediate pleasures more than later ones. If one lacks an art as exact as weighing, one will be prey to a certain illusion:

> Do things of the same size appear to you larger when seen near at hand and smaller when seen from a distance? … And equal sounds seem louder when near at hand, softer when farther away? … If our well being depended on this, doing and choosing large things, fleeing and not doing the small ones, what would we see as our salvation in life: the art of measurement or the power of appearance? Appearances cause us to wander in confusion and make us shift back and forth many times about the same things (*ê hautê men hemas eplana kai epoiei anô te kai katô pollakis metalambanein tauta*) and regret our actions and choices between the large and the small. Measurement makes the appearances powerless (*akuron*), by revealing the truth, and makes the soul stable, abiding by the truth, and saves our life. (356C8–E4)

This passage has deservedly received a great deal of attention. It is generally agreed to give us, albeit in a very brief form, Socrates' own account of the experience the many call 'being overcome by pleasure'.

Socrates contrasts two ways in which we can choose between the 'the large and the small': allowing ourselves to be influenced by the 'power of appearance' (*tou phainomenou dunamis*) and applying the 'art of measurement' (*metrêtikê technê*). These two possibilities are presented as exhaustive: if we lack the art of measurement, then we will be prey to the power of appearance. The art of meas-

[11] This objection rejects (2), that desires motivate us towards the good optimally, and possibly (3), that desires are responsive only to beliefs. But note that it doesn't reject (1): the immediate pleasures that we desire more strongly than distant pleasures can still be desired *qua* good.

urement is described as a kind of knowledge (357A1–2), so even on Socrates' own account of what it is to act against our better judgement—if we allow that it is a kind of 'acting against'—it is something that only affects beliefs: beliefs but not knowledge are affected by the power of appearance. While our beliefs are liable to waver, and potentially overturn, when we encounter a misleading appearance, knowledge is 'strong' enough to remain unchanged: it 'makes the soul stable, abiding by the truth' and renders misleading appearances 'powerless' (*akuron*).[12]

What causes one's beliefs to waver is immediate pleasure's deceptive appearance. Like a near object that looks larger than an equal but distant one, an immediate pleasure appears more pleasurable—and so more desirable—than an equal but distant pleasure. This appearance causes an event with three stages: first, it causes the akratic to 'wander' and shift 'back and forth' about the same thing; second, implicitly, he mistakenly indulges in the pleasure; and finally he 'regrets' or 'rues' (*metamelien*) his action. This is highly condensed, but with some reconstruction it lends itself to a plausible intellectualist account of *akrasia*. An influential, and I think largely correct, reading locates the conflict of the experience in the agent's unstable assessment of the tempting pleasure.[13] Its appearances causes him to 'shift back and forth many times about the same thing'. Plato regularly uses the phrase 'shift back and forth'—literally 'up and down' (*anô kai katô*)—to describe instability or vacillation in belief.[14] The suggestion, then, is

[12] Socrates initially declares his aim to be to show that 'if someone were to know what is good and bad, then he would not be forced by anything to act otherwise than knowledge dictates—wisdom is sufficient to save a person' (352C3–7). However, he also claims that no one acts contrary to what they *believe* best (e.g. 358B6–D3, discussed on p. 4 above), while still maintaining that only knowledge, not belief, is 'sufficient to save a person' (356E5–357A4). This suggests that the relevant difference between true belief and knowledge is that only the latter remains stable in challenging epistemic situations. While we never act contrary to a true belief while we still hold it, the power of appearances can make us act otherwise by causing us to retract our belief. For defences of this reading see Penner (1996) and (1997), Carone (2001), and Reshotko (2006).

[13] For a detailed defence of this reading see, for example, Penner (1990). It is also interestingly close to the view of *akrasia* that Plutarch attributes to Chrysippus: *On Moral Virtue*, 446F–447A.

[14] For example, at the aporetic ending of the *Hippias Minor*: 'about these matters I wander back and forth (*anô kai katô planômai*) and never believe the same thing' (376C1–3). See also *Gorg.* 481D7–E3, *Rep.* 508D4–9, and *Theat.* 195C1–4.

that since the akratic agent lacks any reliably means—the art of measurement—to assess the veracity of the appearance, he experiences a kind of 'will I, won't I' conflict, vacillating between opposing assessments of, and desire and aversion for, the tempting pleasure: one that bids him to indulge in the pleasure, based on the pleasure's appearance, and one that bids him to refrain from it, based on less immediate considerations like, for example, his future health. As Penner describes it, this is a 'diachronic' conflict: a conflict within a period of time but without conflict at any particular instant of time (without 'synchronic' conflict). While this is a weaker form of conflict than the many proposed, it is certainly a *bona fide* conflict. It seems perfectly accurate, for example, to described it as an inner struggle or as the experience of being pulled in opposite directions by opposing desires.[15]

What I wish to look at more closely is this experience's cause, the 'power of appearance' that immediate pleasures possess. My principal claim is that the power of appearance addresses the many's belief that a distinctive kind of desire causes *akrasia*: non-rational, unruly passions, and especially appetites for pleasures like food, drink, or sex. While the many believe that these pleasures lead us astray by being the objects of strong non-rational appetites, Socrates believes that they lead us astray by appearing more attractive than they really are.

The most controversial implication of this reading is that some but not all pleasures and pains exert the power of appearance. Specifically, it is a characteristic of those lower or bodily pleasures and pains that are typically thought to elicit non-rational passions. But Socrates also recognises other pleasures, the pleasures of virtue, that don't exert the power of appearance. This has more textual support than one might think. The fact that it is not explicit in 356C8–E4

[15] Why will the akratic regret his choice, after he has sided with the appearance and enjoyed the pleasure? To feel regret he must return to his original assessment, which implies that the appearance's power over his judgement is temporary. It seems plausible that something that vividly appears pleasant can, for just as long as it appears, monopolise the akratic's attention and so distract him from less immediately striking considerations like future ill-health. But after he has consumed the pleasure—when it is a remembered rather than immediate pleasure—less salient considerations are no longer overshadowed and can return, leading him back to his original assessment and causing him to regret his action.

is a reflection of Socrates' dialectical strategy. He requires the many's commitment to the identity of the good and the pleasant, but not to any view of what pleasures there are or how they should be ranked (and since the many lack the art of measurement, it is of course likely that the view they do hold is mistaken). Consequently, Socrates remains neutral on this question during his discussion with the many and only later makes explicit the role played by less appetitive pleasures, like the pleasure of courageous action, that can appear *un*pleasant.

I will look at three passages. First, I look at the central description of the power of appearance, examining in particular what Socrates means by an 'appearance' (356C8–E4). Next, I turn to Socrates' earlier discussion of what the many call 'bad pleasures' and 'good pains' (353C3–354E2). In this passage we are given examples of the 'immediate' and 'later' pleasures and pains that enter into the akratic's deliberations. What we will notice is that they are, in some sense, different in kind. The akratic does not weigh, say, a beer now against two beers in a week's time—a very unrealistic picture of how we deliberate—but a beer now against a more general consideration like future health. Socrates tells us that these later, more general goods contribute instrumentally to a life that is pleasant overall, but he remains neutral about what pleasures such a life consists of, specifically whether it is the same sensory pleasures that mislead the akratic, but enjoyed more wisely, or another kind of pleasure. In a third passage, Socrates' second discussion of the virtue of courage (359A2–360E2), we find a very strong indication that he has in mind another kind of pleasure. Socrates argues that, because it is honourable, the courageous act of going to war must be pleasant. I argue that this pleasure cannot be bodily but must be a non-bodily pleasure, of some sort, that is associated with virtuous action and that does not exert the power of appearance.

1.2.1 *False appearances: 356C8–E4*

Socrates' entire account of the false appearance of immediate pleasure is in the form of an analogy with perceptual appearances: 'things of the same size appear

(*phainetai*) to you larger when seen near at hand and smaller when seen from a distance'. We can see in outline how this analogy should apply to pleasure. The many have just suggested that there is a difference between 'immediate pleasure' (*to parachrêma hêdu*) and pleasure at a 'later time' (*husteron chronon*). Socrates' response seems to be that they are not in fact different, but merely appear so: just as spatially near objects falsely appear larger than equal but distant ones, temporally near pleasures falsely appear better—more pleasurable—than equal but distant pleasures. For a more detailed understanding we need to first consider what it means for something to have a false 'appearance' (*phainomenon*, 356D4, or *phantasma*, D8)?

Just like 'appear' in English, *phainesthai* can have two quite different senses: a judgemental sense in which it refers to what is believed to be the case—for example, a Socratic interlocutor often replies 'it appears so' (*phainetai*) to register his agreement—and a phenomenal sense in which it refers to how something is presented to one, whether or not one believes it. Aristotle gives a very explicit example of the latter sense: the sun 'appears' (*phainetai*) a foot across despite the fact that we believe it to be considerably larger (*De Insom.* 460B16–20).[16] Throughout the description of the power of appearance the latter sense is intended. This is obvious in the case of near objects that appear larger. To say near objects appear larger is just to say that this is how they look, how they are presented to the senses. There are two good reasons to think that pleasure's appearance is similarly phenomenal. First, pleasure's appearance 'makes' (*poiein*; 356D5) our beliefs shift back and forth—this is its *dunamis*, its power or ability—and it does so by contradicting our original assessment of the pleasure. So the appearance is the cause of the error, what leads us to overestimate the value of immediate pleasures, and it could not play this role if it is, or is dependent on, what we *already* believe (a belief that *p* can not explain why we come to believe not-*p* or even, for that matter, why we believe *p*).

Second, the art of measurement is said to make the appearance *akuron*, with-

[16] I discuss this distinction in more detail in Chapter 3, p. 112 ff.

out 'power' or 'authority' (*kuros*). Of course, it could lack authority because it no longer exists, but describing it as *akuron* would be an unusual way to say this (compare saying of a dead king that he 'lacks authority'). It is more likely that the intention is comparable to a use of *akuron* in the *Crito* (50B4). Here *akuron* describes the effect that Socrates' escape would have on the court's verdict: his escaping would not of course annul or change the court's verdict, but it would undermine its authority by making it powerless to control Socrates' fate.[17] Analogously, in the *Protagoras* the art of measurement makes an appearance *akuron* by taking away not the appearance itself, but the appearance's power to control our judgements. And it is easy to see why: appearances control our judgement only if we judge by appearances, but the art of measurement gives us an alternative, accurate way of judging. So taking our cue from Socrates' analogy, we can say that just as near objects continue to look larger than equal but distant ones even if we don't believe they are larger—since we don't judge their size by their appearance—near pleasures continue to appear better even if we now know they are not as they appear. (In the *Philebus* Socrates uses exactly the same analogy: 'does it happen only in sight that seeing objects from afar or close by distorts the truth and causes (*poiei*) false judgement? Or does not the same thing happen also in the case of pleasures and pains?' (41E9–42A3). But here Socrates is quite explicit that false pleasures of this kind, in contrast to anticipatory pleasures, appear as they do independently of our judgement (42A7–C3).)

If this is right, the difference between the akratic and someone who has the art of measurement lies solely in the way in which they form their judgements. The information they have is the same, but what they conclude from this

[17] This is an example Bobonich (2007) draws attention to and I reach a similar conclusion to him. However, it seems that he nonetheless continues to think of the appearances as judgemental, or so it seems from the following commentary: 'what this passage suggests is that even in a person with the measuring art an appearance could be present, it would just not determine how a person acts. But if Plato were to allow this, we seem to get two potentially conflicting sources of judgement and be on the way to a middle-period position.' He qualifies this in a footnote: 'But it is worth noting this possibility does not require the existence of two judgement-makers: all that is required is that the content of a previous judgement still, in some way, remain 'live' after a new one has been formed' (p. 55 and n. 27).

information is different. On the one side is the art of measurement, which includes weighing, counting, and measurement of magnitudes like size, thickness, number, and loudness (356A8–E4). (Socrates excuses himself from an explicit account of the art of measuring pleasure and pain: 'what exactly this art and knowledge is we can inquire into later' (356B5–6)). On the other side is judging the same magnitudes by how they sensibly appear, that is, how they are presented to the senses. Near objects look larger than equal distant ones: if we measure them, we will know they are equally sized; if we judge them by how they look, we will falsely believe they are unequal.

The akratic's mistake, then, is to base his belief on appearances. Judging by appearances is a very easy way to acquire a belief, requiring no more than an act of assent. If one does not engage in any rational reflection, which might give one grounds for doubt, then there is little to prevent one from uncritically assuming that things are simply as they appear.[18] This gives us a sense of why Socrates thinks appearances are a dangerous, powerful thing. In the absence of the appropriate art of measurement, appearances offer an unfortunate mix of being both compelling and false. (All of the relevant uses of *phainetai* and *phainomena* in the *Protagoras* refer to false appearance; a 'true appearance' may not be an oxymoron, but since appearance-language offers a non-factive alternative to verbs like 'perceive', it lends itself to cases where one is at least unsure about whether things are as they appear.) It is, then, easier—it requires less cognitive effort or skill—to come to a false belief about merely apparent pleasure or pain than a

[18] The distinction between reasoning, on the one hand, and judging by appearances, on the other, is explicitly drawn in a strikingly similar passage in book 10 of the *Republic*, albeit with the help of a very different psychology: 'through sight the same magnitude doesn't appear to us to be equal when near and far away … [but] measuring, counting, and weighing give us welcome assistance in these cases, so that we aren't ruled by something's appearing bigger or smaller' (602C7–D8). Socrates argues that the part of the soul that uses measurement in such cases is, unsurprisingly, the part capable of 'reasoning' or 'calculating' (*logismos*), namely the rational part (*to logistikon*; 602E1). But he argues that in addition there's another part of the soul that believes the opposite, that things are as they appear, and it does so precisely because it is a part that is *in*capable of reasoning. At least in the *Republic*, then, judging by appearances is inevitable in the absence of reasoning or reflection. In the next chapter, I argue that we find the same account of judging by appearances in the *Gorgias*.

true belief, with the inevitable result that most people are led astray by immediate pleasures and pains.

The analogy with perceptual appearances leaves one question unanswered. To explain why a near object appears larger we can point to the fact that it takes up greater space in our visual field, but what can we point to to explain why a near pleasure appears more attractive? It seems right that near pleasures do indeed appear more attractive, yet it is surprisingly difficult to explain why this is so.[19] Attempts to explain immediate pleasure's deceptive appearance have tended to be entirely speculative. I think a more textually grounded explanation is possible. In an earlier passage, 353C3–354E2, Socrates not only gives examples of the kind of immediate and later pleasure and pains that we are liable to mis-measure but also, implicitly, explains why immediate ones appear better (or worse) than distant ones.

1.2.2 *The immediate and later: 353C3–354E2*

To show the many that they call pains 'good' only when they believe that overall they promote greater pleasure and relief from pain, Socrates asks them:

> Would you call these things [sc. cautery, surgery, or military training] good for the reason that they provide intense pain and suffering in the present (*en tô parachrêma*) or because at a later time (*eis ton husteron*) they bring about health and good conditions of body and preservation of cities and power over others and wealth? (354B1–5)

Likewise, bad pleasures like harmful food, drink or sex 'are bad not on account of the immediate pleasure (*tên hêdonês tês parachrêma*) they provide, but because of what they bring about later (*ta husteron*), disease and things like that' (353D7–E1). The distinction between pleasures and pains that are 'present' or 'immediate' (*parachrêma*) and ones that are 'later' (*husteron*) is precisely the distinction at play in the 'power of appearance' passage. The pleasures and pains in

[19] Brickhouse and Smith lean heavily on the difficulty this question poses to motivate their view that 'it is strong [non-rational] appetite or passion … that accounts for the object of the appetite or passion having the power of appearance' (2007, p. 2). As I argued, this reading is precluded by, in particular, the account of fear at 358D6–E6, but the alternative answer I give in the following will reinforce the idea that the power of appearance is independent of non-rational desires.

question also seem to be of the right kind: resisting painful cautery despite believing that one ought to endure it for one's health is a plausible example of an akratic action. Moreover, to the many the later pleasures and pains *appear* different to the immediate ones. As we've seen, Socrates' argument aims to address the fact that the many fail to understand why they value the kind of good and bad things described as 'later'. They mistakenly think that they value them on some metric other than pleasure and pain. We are never told why they make this mistake, but a reasonable guess is that they have a general, non-explanatory sense, perhaps based on experience or received wisdom, that things go well when one is healthy or wealthy without having reflected sufficiently to recognise that the only constituent of 'going well' that they will be able to point to is pleasure. The many, then, have a poor understanding of why they value such things as health or wealth, yet have a clear grasp of why they value such things as food, drink, or sex. The answer in both case is the pleasure they provide, but they only recognise this with the latter.

It is easy to see why there is this disparity between the immediate and the later. The immediate pleasures and pains mentioned—drink, sex, cautery, arduous training and the like—are obviously pleasurable or painful, so the many grasp their value easily. In contrast, it requires reflection to appreciate, let alone measure, the pleasure or pain arising from poverty, power, disease, or the preservation of cities. The many fail to engage in this reflection and, consequently, fail to understand why they find them good or bad. This difference is not strictly tied to whether or not they are immediate or later. One can be healthy or powerful in the present and one can enjoy food, drink, and sex in the future. Rather, the temporal difference is explained by the role they play in our deliberations. No one can calculate all the particular pleasures and pains—cakes, beers, wounds, and so forth—that they will experience in the future. In this epistemic sense, specific pleasures and pains are very much tied to the present. When we weigh near pleasures and pains against future ones we typically weigh something definite and particular, like drinking this beer, against something more probabilistic

and general, like one's future health or finances—one rarely plans to drink a beer in a year's time and no one can decide to be healthy or rich right now.

We can draw from this a very simple account of pleasure and pain's 'power of appearance', one that is both plausible and textually motivated. The crucial point is that the akratic does not weigh like against like, but rather one kind of good or bad against another kind: definite, particular pleasures and pains like sex or cautery against more probabilistic, general consequences like health or poverty. Succinctly put, the former appear better (or worse) than the latter because their value is both more prominent and more easily understood. The salient characteristics of a tempting pleasure, which is typically its sensible characteristics, make its pleasure abundantly clear. A glass of ice-cold beer on a hot day, for example, looks refreshing and it clearly brings to mind its pleasant taste. This much is obvious from a *prima facie* assessment. Moreover, the many have a clear grasp of *why* these goods are desirable. They understand that when they pursue pleasure it is simply because it is pleasant, that pleasure is an end for them. In contrast, the latter kind of good and bad things are abstract and difficult to understand. What effect another beer will have on one's long-term health, or even on tomorrow's productivity, demands thought. Since these kind of considerations are not revealed by something's look or taste, to appreciate them one must go beyond appearances—beyond judging something by its sensible characteristics—and engage instead in reasoning and, ideally, the art of measurement. This is difficult for the many not only because reasoning is more difficult than assenting to appearances, but also because they don't sufficiently grasp why they value goods like health. They *do* value them, but since they haven't reflected on why this is so, it amounts to an ungrounded faith in their value. It is not surprising, then, that their commitment to them is subject to confusion and liable to waver.[20]

[20] Compare the connection between understanding and the stability of beliefs in the *Meno*: true beliefs are beneficial 'as long as they remain … but they are not willing to remain long, and they escape from a man's mind, so that they are not worth much until tied down by an account of the reason why'; 97E6–98A4.

Thus, to say that immediate pleasure appears 'larger' than it really is, is to say that its appearance advertises its pleasure (the beer's taste) but not its pain (its effect on one's health). Judged by appearances, it seems to yield greater overall pleasure simply because this affords one an unequal view, revealing immediate, particular benefit or harm more prominently than later, general benefit or harm. The strength of this reading comes from the fact that it doesn't venture far from what Socrates makes explicit. It relies on two claims. First, that we should identify the immediate and later pleasure and pains that the akratic deliberates over with the two groups of immediate and later pleasures and pains that Socrates pointed to earlier, which themselves plausibly capture the kinds of short- and long-term considerations people typically take into account in their deliberations. Second, that to explain the difference between immediate and later pleasures and pains—a difference in salience—we should turn to Socrates' claim that the many have an unequal grasp of the value of immediate and later pleasures and pains.

This still leaves one important question unanswered: what exactly is the difference between the two groups of good or bad things that Socrates introduces? The two groups are (A) the pleasures and pains of things like sex or cautery and (B) good and bad things that are less obviously sensory like health or poverty. It is easy to assume, since their pleasure is obvious, that what makes life pleasant is solely the kind of bodily, sensory pleasures epitomised by pleasant food, drink, and sex. On this reading, if the members of (B) are good or bad it is only because they promote bodily pleasures or pains of kind (A). If we look closely at the central examination of pleasures and pains (353C3–354E2) we will notice that it is in fact, and tellingly, more difficult to pinpoint what Socrates believes makes life pleasant overall.

Note first that even the many have a low opinion of bodily pleasures. Food, drink, and sex are introduced as examples of pleasures that, in their opinion, can be 'ruinous' (*ponêros*; 353C7)—such pleasures often overcome our better judgement and lead to the harmful things of kind (B), such as disease and poverty.

Moreover, as Socrates' argument shows, since disease or poverty are harmful only because they end in greater pain overall, the bodily pleasures of kind (A) leading to them are in fact, all things considered, *painful* things. Likewise, *mutatis mutandis*, for the so-called pains of kind (A)—they are, overall, pleasurable. The one thing we can say for certain is that a life that is pleasant overall requires some goods of kind (B), such as health, power, wealth, and the preservation of cities. These goods, however, are merely instrumental. Socrates does not say that they are pleasant in themselves but that 'they end in pleasure and the relief and avoidance of pain' (354B6–7). What pleasures and pains? Importantly, Socrates never commits himself to a view. We might *assume* that he means a more wisely chosen selection of the sensory pleasures and pains of kind (A)—engaging in moderate, healthy feasting, but also submitting to starvation diets if necessary—but this is not something that Socrates says. Rather, we only hear about sensory pleasure's harmful side. In this part of the dialogue, then, we are simply never told what makes life pleasant overall.

1.2.3 *Two kinds of pleasures and pains: 359A2–360E2*

The suspicion that Socrates holds a more sophisticated view of pleasure is confirmed later in the dialogue, after his discussion with the many is completed. One problem, for Socrates, with a hedonism that includes only sensory pleasures, what is often called 'Sybaritic' hedonism, is that it is difficult to square with his view of virtue. With this kind of hedonism in mind, it is easy to be sympathetic with Zeyl's conclusion that 'there is an irreconcilable incompatibility between the claim that virtue and the care of one's soul is supremely important, and the claim that pleasure is the only ultimate good.' (1980, p. 263). The problem of course is that it implies that the virtues are only valuable to the extent that they help us indulge in pleasures like food, drink, and sex. It is difficult to see how this will not force either an outright rejection of virtue's value or a revision of the traditional view of what the virtues are. For example, in the *Gorgias* Callicles defends a hedonism centred on appetitive pleasures and

it leads him—quite reasonably, given his assumptions—to reject traditional virtue in favour of 'virtues' that are consistent with self-serving tyranny (a position I discuss in the next chapter). Whatever differences there may be between these versions of hedonism, Callicles' point can be applied to the *Protagoras*: if there are only appetitive pleasures and virtue is defined as what allows us to get the greatest amount of pleasure, we have no reason to think that virtue can be other-regarding. Rather, it implies that if treating others badly gets us greater pleasure, we ought to—it is virtuous to—treat others badly. The response that injustice never does result in greater *bodily* pleasure seems naïve at best.[21]

Socrates makes it very clear, in an apparently calculated move, that he is not similarly compromising virtue. What makes it seem calculated is the aptness of the example he chooses. He chooses not only a virtue that has no *prima facie* link to physical pleasure, but an instance of this virtue that involves considerable physical pain, namely the courageous act of going to war. Socrates claims that this is pleasant and he uses a surprising argument to show it. Rather than working forwards from what is widely accepted to be pleasant, Socrates works backwards from what is widely accepted to be good: if the good really is pleasure, then going to war must be pleasant, since it is widely accepted to be honourable and honourable things are good.[22]

[21] Irwin notes that a similar criticism from the *Gorgias*—Socrates' claim that Callicles' hedonism implies that cowardice is, at least, no worse than courage—can also be applied to the *Protagoras*, but he believes that this warrants only the conclusion that Plato changed his mind about hedonism since writing the *Protagoras* (1995, pp. 113-114). It seems to me, in contrast, that the tension between hedonism and virtue is sufficiently obvious and sufficiently serious that Plato could not have missed it any dialogue in which he is thinking about them side-by-side. The very fact that it is not explicitly flagged in the *Protagoras* is, I believe, a reason to suspect that there is a more subtle resolution.

[22] Given this argument's structure, it is worth comparing to Russell's claim that what Socrates defends in the *Protagoras* is 'explanatory hedonism' (2000). This is a form of hedonism that leaves our pre-hedonistic ideas of what is good entirely unchanged, but appeals to pleasure to explain why we believe they are good. The reading I will propose differs in two respects from Russell's. First, Russell argues that this is the position defended—for purely dialectical reasons, he believes—in the discussion with the many. I have argued that in this discussion Socrates remains neutral on the question of what pleasures there are, relying only on the many's commitment to the identity of the pleasant and good. By that point in the dialogue, then, he has not yet shown that hedonism will leave our ideas of, for example, what is honourable unchanged. Second, I do not believe that the *kind* of hedonism Socrates defends makes it a forgone conclusion that our

> S: Is going to war honourable (*kalon*) or disgraceful? P: Honourable.
> S: Then, if it is honourable, we have agreed before, it is also good, for we agreed that all honourable actions were good. P: Very true, and I always believed this.
> S:And rightly; but who would you say are not willing to go to war, war being honourable and good? P: The cowardly.
> S: If a thing is honourable and good, is it also pleasant? P: That was indeed agreed. S: So the cowardly, with full knowledge, are not willing to go toward the more honourable, better, and more pleasant? (359E4–360A5)

Popular opinion would have it that going to war is unpleasant but honourable. If everything honourable is good, then hedonism requires us to revise this opinion: either going to war is not unpleasant or it is not honourable. Sybaritic hedonism would suggest that we revise our view of what's honourable: war, being unpleasant, must be shameful; what is honourable is whatever allows one to indulge in as much food, drink, and sex as possible. (Cf. Epicurus: 'We must esteem (*timêteon*) such things as honour and virtue if they provide pleasure; if they do not, we must dismiss them'; *Dep. Lib. Rel.*, Frag. 22, 4.1–4.)

The fact that Socrates argues for a revision in the opposite direction suggests, I believe, that he has a different kind of pleasure in mind. He relies only on the claim that honourable actions are good and, therefore, pleasant. One response to this claim has been to take the argument to be a kind of *reductio ad absurdum* of hedonism: the pursuit of pleasure so obviously *doesn't* lead to what are normally understood to be honourable actions, especially self-sacrificing actions like going to war, that Socrates must be subtly illustrating that hedonism is incompatible with virtue (See Weiss (1990), McCoy (1998), and Hemmenway (1996)).[23] There are two problems with this. First, Socrates uses this account of courage to draw a conclusion that most commentators believe he intends sincerely, namely the unity of the virtues. Second, the idea that it is a *reductio* derives primarily from these authors' appeal to our own incredulity at the align-

pre-hedonist ideas of what is good will align with what is pleasant. On the interpretation I offer, Socrates relies on the independent claim that honourable actions are pleasant, a claim that is defeasible on its own terms, not as part of the broader hedonist theory he defends.

[23] Adam Beresford, in an unpublished draft, offers a careful treatment of this reading, including an analysis of the passage to which I am largely sympathetic, despite our opposite conclusions (see bibliography for url).

ment of honourable and pleasant actions, rather than an indication of this in the text or structure of the argument. It is true that the claim that going to war is pleasant is surprising and it seems likely that it is intended to be so, but before judging it to be absurd, it is surely preferable to look for a reading on which it reveals a surprising truth, consonant with Socrates' apparent sincerity. I propose that it is introduced to put in question the assumption that the range of human pleasures extends no further than the pleasant physical sensations we get from things like food, drink, or sex.

This has two advantages. First, it is far from an implausible claim. Most of us would agree that we can take pleasure in a wide range of activities and experiences, ranging from lower physical pleasures to higher intellectual pleasures. It is also a thesis that Plato defends elsewhere. In the *Republic*, for example, Socrates distinguishes three kinds of pleasures—appetitive pleasures, pleasures of honour and victory, and intellectual pleasures—and argues that it is the virtuous man who enjoys the greatest overall pleasure, while those devoted only to appetitive pleasures live a life that is vicious and painful overall (580D–588A). Secondly, and most importantly, it explains how Socrates can argue for hedonism and still maintain his commitment to the value of virtue, thereby allaying one of the principal reasons for doubting that Socrates could seriously be commitment to the hedonism he very much seems to defend.

If this is right, then it is also the case that honourable actions must afford a pleasure that far outweighs the pleasures of food, drink, or sex. War not only involves hardship and injury, it also involves a very real possibility of death. In line with the traditional conception of courage and honour, what Socrates calls pleasant is not fair-weather soldiering, joining only those battles in which ones survival is ensured, but rather simply 'going to war', which is as likely to involve one's own death as it is the enemy's. His argument is that going to war is pleasant—not that it is courageous to go to pleasant wars.[24] This entails that

[24] Contra Weiss's reading (2006, pp. 63–67). Note that this also excludes the possibility that Socrates is treating going to war like cautery: painful in the moment, but in the long run leading to greater overall bodily pleasure. Even if one could adduce pleasant rewards that outweigh the

courageously going to war, even if one will die, is more pleasant than cowardly avoiding it, even if one would have lived a long life of pleasant food, drink, and sex. This does not necessarily mean that a single courageous act is more pleasant than a whole lifetime of lesser pleasures. It is more plausibly interpreted as the claim that surviving is not worthwhile if one is prepared to live a life of shame and vice, no matter how many lesser pleasures one enjoys. Again, this is in line with traditional views of virtue, while revising the traditional view of the most pleasant life. It is also in line with what Socrates says about virtue elsewhere. For example, in the *Apology* he says that a good man should not 'take into account the risk of life or death; he should only look to this in his actions: whether what he does is just or unjust' (28B5–9).[25]

Courage in war brings out the contrast between the virtuous life and a life of bodily pleasure especially starkly, but his argument can be applied equally to the other virtues: wisdom, justice, piety, and temperance. This is implied by the unity of the virtues and the fact that all the virtues are—on Protagoras's reckoning, although we can safely assume that Socrates agrees—'as honourable as anything can be' (349E3–8) and, therefore, by the same reasoning, as good and pleasant as anything can be. The conclusion that Socrates' treatment of going to

pains of war, these pleasures could only make war pleasant and honourable for those who return home. Those willing to face death would be facing only pain and therefore—if overall pain is its measure—only *dis*honour.

[25] We find the same claim in the *Crito*, 48D3–5, and *Gorgias*, 511B1–512E5. Gosling and Taylor offer a number of ways to reconcile courage in the face of death with a more Sybaritic hedonism (1982, pp. 62-65). They appeal first to the pleasures of an afterlife. I find this unconvincing: if the pleasure of going to war depended on this—a claim that is controversial within Greek religion—Socrates would hardly have left it implicit. Moreover, as they recognise, in the *Apology* Socrates couples the same claim with an agnosticism about the afterlife (40C4–9). I'm similarly unconvinced by their suggestion that facing death might be an unfortunate but necessary consequence of having the courageous disposition needed to avoid a life of fearful distress. Courage is knowledge of what is and is not to be feared, that is, of what is and is not bad. The only relevant disposition, if we can call it that, of a courageous person is to avoid, and fear, what is in fact bad. This disposition could never make it necessary to pursue something harmful. Therefore, a courageous act must be a beneficial act, even if it might cause one's death. Its benefit is best seen, I believe, as the lesser of two evils: it is better to die virtuous than live viciously. If one is prepared to do something vicious, like cowardly avoiding war, for the sake of what one believes to be a good life, than one must have a mistaken conception of the good life, one that neglects virtue and, therefore, leaves one with the pain of vice.

war suggests, then, is that the virtuous life is pleasant and pleasant in a way different from, and superior to, a life devoted to bodily pleasures. This is a strong thesis, but it is a far from surprising one to find in a dialogue in which Socrates is committed to both virtue and hedonism.

We can now complete Socrates' earlier account of pleasure and pain. As we saw, he introduces an instrumental structure without a clear end: certain immediate 'good' bodily pains lead in the long run to later goods like health or wealth that, in turn, 'end in pleasure'. Socrates' account of courage suggests that this final end, what makes life pleasant overall, is virtue. Precisely what kind of pleasure virtue provides is an important question, but one that need not concern us here.[26] The simple fact that it is pleasant is sufficient for the point I wish to make: certain pleasures, like the pleasure of courage, are far from obvious, since it is argument rather than appearances that reveals that they are pleasant. The power of appearances, then, is not a power held by all pleasures. It is a power held specifically by bodily pleasures, pleasures that often lead to greater overall pain. We see this very clearly in the coward's attitude to going to war.

Socrates' account of cowardice is implicit but easily reconstructed. It should be seen, I believe, as an application of his account of the experience the many call 'being overcome by pleasure'. The many believe not only that people can be overcome by pleasure or pain, but also by the pleasure- and pain-involving passions of lust, anger, and 'often fear' (352B5–8). In addition, most people think that fear makes the difference between the courageous and the cowardly: the former go towards and the latter avoid fearful things (359C5–7). Presumably, on this popular conception, it is not that cowards don't believe actions like going to war are honourable, and therefore good, but that they are simply too afraid to do them—they are overcome by fear. Socrates' response is the same as it was for the idea that people are overcome by pleasure: since honourable actions are

[26] Rudebusch (1999), drawing on a number of Socratic dialogues, argues that Socrates sees virtue as a 'modal' pleasure, an activity that is done or happens in a certain way—comparable to Aristotle's view of pleasure as an 'unimpeded activity of the natural state'—rather than a 'sensate' pleasure, a feeling or sensation that accompanies an activity.

themselves pleasant, it is a mistake to think that the honourable and the pleasant can pull us in opposite directions. Once we recognise this, the popular conception of cowardice can be restated as the absurd idea that because of fear, which is an expectation of pain, 'the cowardly, with full knowledge, are not willing to go toward the more honourable, better, and more pleasant' (360A4–5). Socrates own account, again, is to deny that they do so 'with full knowledge'. Rather, cowards are fearful of going to war because of ignorance. He does not spell out what this ignorance consists in, but it is surely ignorance of what he has argued, that courageous action is pleasant. The cowardly, then, see the obvious physical pains and hardships of going to war—these are the features of going to war that exert the power of appearance—but they fail to see that this is outweighed by the pleasure of honourable, and pain of shameful, action. What leads the coward astray is the fact that the physical pains he anticipates are highly salient but the most relevant pleasures and pains are invisible to him. He judges going to war by its appearance, but the pleasures and pains of virtue and vice can only be recognised through reasoning.

Thus, while Socrates' intellectualism entails that all desires are rational and his hedonism entails that the only good is pleasure, he can nonetheless distinguish between two kinds of desires, each characterised by the kind of object they have. The experience the many call 'being overcome by pleasure' occurs when a person is liable to be led by appetitive desires, that is, by desires for appetitive pleasures. In other words, they are left open to the danger of deriving their beliefs about what is most pleasant from what deceptively appears pleasant and, consequently, to desire and pursue immediate, bodily pleasures that are often harmful. In contrast, someone with the art of measurement will be immune to false appearances and will therefore consistently desire and pursue a different kind of pleasure that is more than a mere appearance, the pleasure of virtue.

The many, then, were half right and half wrong in their assumption that there are two kinds of object of desire. They were right that value is not exhausted by the pleasures of food, drink, or sex and that they do not value 'later'

goods like health or wealth simply because they lead to these pleasures. They were wrong, however, to conclude that this implies that there are two incommensurable, potentially conflicting kinds of object of desire, pleasant things and good things. All good and pleasant things, all objects of value, are measured on a single scale, but they have failed to sufficiently understand their reason for valuing goods other than bodily pleasures. They feel the pull, for example, of honourable action without having fully worked out where its value lies, namely that it is something that is pleasant and good in itself. Thus, when forced to make difficult decisions like whether or not it is best to go to war, they are liable, under the influence of its painful appearance, to choose the cowardly option while nonetheless, once their attention is diverted from its pain, feeling the shame of having acted dishonourably.

The *Protagoras* is the first dialogue in which appearances play a significant role but, remarkably, it is also more or less the same role that they play in the many subsequent Platonic dialogues in which they appear. There are in particular three closely related constants, all of which are prominent in the *Protagoras*. First, Plato is especially interested in evaluative appearances and these appearances, he believes, are almost invariably false: we live in a world in which what appears to be good or bad is typically not in fact so. Second, and consequently, appearances are the main cause of practical error. If what appeared good was indeed good, doing what is best would be a straightforward, cognitively undemanding task. But as this is not the case, appearances lead people to have false beliefs about what is good and bad and practical success, and ultimately virtue and happiness, is very difficult to achieve, requiring one to investigate and acquire knowledge of what's good and bad. Third, our lower desires are controlled by appearances. This claim does differ in the early and middle dialogues, but not as much as one might expect. Common to both periods is the claim that lower passions—the passions that relate to bodily pleasures and pains—are a result of an unreflective acceptance of false evaluative appearances: by either the whole person, in the early dialogues, or a part of a person's soul, in the *Republic*. In the

next chapter I examine how these three claims are developed in the *Gorgias*.

CHAPTER 2

Appearances in the *Gorgias*

In stark contrast to the *Protagoras*, Socrates strenuously rejects hedonism in the *Gorgias*. Far from arguing that a virtuous life is the most pleasant, he argues that virtue is better than vice even if it is accompanied by great pain. Nonetheless, when it comes to lower, bodily pleasures and pains—the only kind of pleasures and pains it explicitly examines—the *Gorgias* is consistent with the *Protagoras*. Like the *Protagoras*, the *Gorgias* presents bodily pleasures and pains as deceptive and error-inducing: they appear better than they really are—not more pleasant or painful this time, but rather falsely good or bad—and thus mislead the majority of people who tend to judge things by how they appear. The *Gorgias*, however, examines this effect both in more detail and considerably more broadly. While the *Protagoras* considers how appearances can cause us to act contrary to our better judgement, the *Gorgias* looks at the consequences of having a 'better' judgement that agrees with appearances *tout court*.

The first two of the three sections of this chapter develop an interpretation of the *Gorgias* as a dialogue that is guided by the following question: what are people likely to believe if they lack knowledge of what is good or bad? The *Gorgias*'s examination of oratory can be seen as an examination of what ordinary people—those who lack the relevant knowledge—find persuasive. The short answer is that they are persuaded by what merely appears to be good or bad, namely pleasure and pain. In §2.1 I look at Socrates' definition of oratory and offer an interpretation of its relevance to another central theme in the *Gorgias*, the ques-

tion of what kind of life we should lead. With increasing detail and decreasing attachment to conventional ideas of virtue, Gorgias, Polus, and Callicles each offer their view of the best life. Its apogee, with Callicles, is a might-is-right hedonism: the pleasant and the good are identical and, therefore, the best life is the one that allows one to secure the greatest share of life's luxuries without fear of reprisal, that is, the life of a tyrant. I argue that oratory and the pursuit of pleasure are two sides of the same practice: one side providing pleasure for others and the other providing it for oneself. They also involve the same account of persuasion, that is, the same account explains how both the orator's—the democratic analogue of the tyrant—and his audience's beliefs come about: both uncritically accept appearances, leading them to believe that what appears best, is best.

In §2.2 I take a closer look at three aspects of Socrates' account. First, his description of how the pleasure-related practices are a kind of non-rational or unreasoning (*alogos*) knack that merely guesses at what's best. Second, his remarkable claim that, *prima facie*, what merely appears good and bad, pleasure and pain, is *more* convincing than what is in fact good and bad. Finally, through an analysis of Polus and Callicles' conceptions of the good, I look at the claim that following appearances to their conclusion leads one not just to error but to the worst mistake one can make, namely treating as the very best life what is in fact 'by nature the greatest and first of all bad things' (479D4–6), a life of unfettered injustice.

In §2.3 I turn to the psychological theory we find in the *Gorgias*, a topic I remain neutral on in the first two sections. Finding a coherent reading of the *Gorgias*'s psychology has proven difficult and divisive. While Socrates argues—it seems—for the same intellectualist psychology we found in the *Protagoras* there is also evidence—it seems—that he adheres to a different and incompatible psychology that is closer to what we find in the *Republic*. While it does not depend upon it, I argue that the reading I develop in §§2.1 and 2.2 strongly favours a consistently intellectualist reading of *Gorgias*. My central claim is that in the *Gorgias* Plato develops the intellectualist account of appetitive desire that was suggested in the *Protagoras*: since what appears best is lower, bodily pleas-

ure, unreflectively accepting appearances leads one to blindly pursue appetitive pleasures like food, drink, and sex. Consonantly, Socrates' account of disciplining or restraining appetites does not, on examination, describe a process of holding back non-rational desires through an exertion of will-power, but of acquiring and maintaining the true beliefs that guide our rational appetites. I conclude that, counter-intuitive as it may seem, this intellectualist reading of the *Gorgias* allows a more meaningful comparison with the psychology of the *Republic*, especially—again, surprisingly—with respect to appetites.

2.1 ORATORY AND THE PURSUIT OF PLEASURE

The *Gorgias* has two principal topics. First, it considers the nature and value of two political practices: oratory and a certain craft of justice. Second, it considers two conceptions of the good life—of 'who is happy and who is not' (472c9–d1)—and two corresponding practices: a practice for getting pleasure, especially bodily pleasures, and a practice for maintaining a virtuous soul. These are undoubtedly related topics and the dialogue's discussion of them is intertwined, but precisely what relationship they have requires investigation. My aim in this section is to shed light on the relationship specifically between oratory and the pursuit of pleasure. The connection between them, I will argue, is found in the persuasiveness of pleasure. Both oratory and the pursuit of a pleasant life are practices that, first, take at face value pleasure's *prima facie* appearance of being good and, second, employ a 'knack' that allows one to both provide (oratory) and procure (the pursuit for pleasure) this apparent good.

2.1.1 *Defining oratory*

The first twenty or so Stephanus pages of the *Gorgias* are devoted to defining oratory. The first attempt looks at Gorgias's account of oratory as a 'producer of persuasion' (453a2). Socrates calls for more precision. To be able to distinguish oratory from other crafts that equally produce persuasion, we must know two

further things: first, what *type* of persuasion oratory is and, second, what *subject* it is persuasive about. As an example answer, Socrates defines arithmetic as 'persuasion through teaching about the extent of the even and odd' (454E7–455A1). To meet the first requirement, Socrates distinguishes between two types of persuasion: 'one producing conviction without knowledge and the other producing knowledge' (454E2–4). Crafts like arithmetic, which persuade through teaching, produce knowledge, but oratory's persuasion is agreed to be a type that produces only conviction (at this stage there is a notable silence about what method of persuasion, corresponding to teaching in arithmetic, is used to produce conviction). Such conviction without knowledge can be produced about subjects about which the orator himself has no knowledge—a considerable benefit in Gorgias's eyes—but can equally only be instilled in those who also lack the relevant knowledge (459B6–C5).

The discussion is less successful on the second requirement. When pressed for an answer, Gorgias says that oratory is 'about the things that are just and unjust' (454B7). This makes some sense: Gorgias has emphasised that oratory is principally practised in the political realm—in law courts, the assembly, or 'any other political gathering that might take place' (452E3–4)—where matters of justice and injustice are at least a major concern. However, Gorgias also wishes to celebrate oratory's extraordinary breadth, claiming that the orator has 'the ability to speak against everyone on every subject, so as in gatherings to be more persuasive, in short, about anything he likes' (457A5–B1). He has, then, claimed both that oratory's subject is the just and unjust and that it not confined to any special subject.[1] This contradiction marks the end of any progress in the first attempt to define oratory, but we soon see that it is a telling failure. The contradiction arises because oratory is not in fact a craft: it is a practice that merely *seems* to be the craft of the just and unjust, but in fact lacks the ability to give an explanatory account of *any* subject (a necessary condition for being a craft: 465A2–6) and,

[1] The same tension between oratory's popular association with the subject of the just and unjust and its ability to persuade in any subject is brought out in the *Phaedrus*. See 260E ff.

as such, has no special subject it can claim as its own—it can persuade on any subject because it is the craft of none.

The claim that oratory is not in fact a craft forms the basis of the second attempt to define oratory, which sees Socrates, in an uncharacteristically didactic turn, presenting his own sophisticated theory of oratory and related practices. Socrates distinguishes two families of practices: those that really benefit the body and soul and those that merely seem to benefit them. The first are genuine crafts that provide care with a view to what is best for the body or soul and that proceed from knowledge of their subject. These are gymnastics and medicine for the body, and legislation and justice (together the craft of politics, 464B4) for the soul. The other family of practices Socrates calls 'flattery' or 'pandering' (*kolakeia*) and they are not crafts but mere 'knacks' (*empeiriai*)—non-intellectual abilities acquired through experience (see p. 52–54 below)—for 'producing a certain gratification and pleasure' (462C7). This family of practices is comprised of cosmetics and pastry baking for the body and sophistry and oratory for the soul.

In contrast to the genuine crafts, flattery merely appears to provide benefit. Each branch of flattery is a sort of pseudo-craft: while they are not crafts, since they don't work from knowledge but from a sort of 'guessing', they nonetheless *seem* to be crafts (463B3). They manage this by each imitating one of the genuine crafts for the care of the body and soul, masking themselves with the character of the craft it is pretending to be (464C7–D1): cosmetics pretends to be gymnastics; pastry baking pretends to be medicine; sophistry pretends to be legislation; and oratory pretends to be justice. The claim is not quite that a craft such as pastry baking aims to be *mistaken* for medicine—even if the political crafts may indeed be thought to be the preserve of oratory and sophistry—but that pastry baking competes with medicine for the title of being the craft that does most good with respect to the body's diet:

> Pastry baking has put on the mask of medicine and pretends to know the foods that are best for the body, so that if a pastry baker and a doctor had to compete in front of children, or in front of men just as foolish as children, to determine which of the two, the doctor or the pastry baker, had expert knowledge of good food and bad, the doctor would die of starvation. (464D3–E1)

Like pastry baking's imitation of medicine, oratory pretends to be the craft of justice. It presents itself as having a craft-like knowledge of what is good and bad for the soul and it practices its knack in the craft of justice's rightful terrain, the law courts, assembly, and other political gatherings. It thereby manages to convince its audience, who lack knowledge, that it is the true craft of the just and unjust and so appears to adults to benefit their souls just as the pastry baker appears to children to benefit their bodies.

How does oratory, and flattery in general, succeed in perpetrating this deception? Oratory was earlier agreed to be a practice for 'producing conviction without knowledge' and now it is described as a practice for 'producing a certain gratification and pleasure'. There is, then, a connection between persuasion and pleasure. This points us towards an account of *how* oratory persuades, that is, to whatever it is that corresponds in oratory to the method of teaching in knowledge-producing forms of persuasion like arithmetic.

The answer Socrates offers is that oratory persuades by exploiting pleasure's natural persuasiveness. Socrates introduces a distinction between the real and apparent good to prepare Gorgias for his distinction between crafts and flattery. He invites Gorgias to agree that there exists both a real 'good condition' (*euexia*) for the body and soul and 'an apparent state of good condition, one that isn't real' (464a3–4). These apparent good conditions will fool those who lack the relevant knowledge—anyone who is not an expert in the relevant good of the body or soul—into thinking that it is a genuine good condition: someone can, for example, 'appear to be physically well, and unless one is a doctor or a gymnastic trainer, one wouldn't easily notice that they're not well' (464a4–6). This is established prior to his account of flattery, which indicates that they are conditions that, by themselves, have this deceptive appearance. Flattery succeeds by exploiting these natural deceptions, using the fact that the inexpert are susceptible to them to deceive people into believing it is genuinely benefiting the body or soul. As becomes clear as the passage continues, what flattery uses to persuade, and what naturally has this appearance of being good, is pleasure: the pastry

baker seems superior to the doctor by giving people pleasant food in place of the doctor's beneficial, though often unpleasant, diet and cosmetics gives people an appealing façade of beauty with eye-catching but empty features, 'deceiving by shaping and colouring, smoothing and draping' (465B4–5). In a more general characterisation, Socrates says that flattery 'with the lure of what's most pleasant at the moment, sniffs out folly and deceives it so that it gives the impression of being most worthy' (464D2–3). Oratory persuades, then, not by teaching but by 'gratifying and pleasing' its audience and thereby exploiting the fact that to the inexpert what gratifies and pleases seems to be beneficial.

This is a brief sketch of the connection between pleasure and persuasion, which I will elaborate on below, but first I wish to look at oratory's relationship with a certain pleasure-centred conception of what's good. As we have seen, oratory, and flattery in general, is a practice based on, or employing, false but convincing candidates for what's good and bad: pleasure and pain. This use of pleasure or pain mirrors a certain conception of the good that the orator himself possesses. This is the conception of the good espoused, with increasing clarity and consistency, by Gorgias, Polus, and Callicles.

We hear first about how oratory has an instrumental benefit for the orator. All three of Socrates' interlocutors extol the orator's ability to get whatever he desires through his skills of persuasion: Gorgias sees oratory to be the 'greatest good' since it is 'the source of rule over others in one's own city' (452D5–8; cf. 452E4–8 and 456B ff.); Polus sees the orator as the democratic analogue of the tyrant, being the most powerful person in a democratic city and capable of doing 'whatever seems best to him', be it slaughtering, exiling, or confiscating (466B4–C2); and Callicles praises the tyrant and orator as those who can rule over and get more than others and who, moreover, are immune from punishment for their actions.[2] Oratory can, then, be a useful tool in the life Socrates' interlocutors take to be good, but this by itself is only an incidental relationship. After all, for

[2] In fact Callicles' praise is only explicitly for tyrants but that the orator is another, perhaps less complete, example of his 'superior' man of action is implied (e.g. 485E6–486A3) and is assumed by Socrates later in the discussion (e.g. 500C4–7)

Polus and Callicles the tyrant is the archetype of the happy man and the orator is treated as happy to the extent that he is like a tyrant—presumably tyrants have no need for oratory, using force instead of persuasion.[3]

While oratory may be dispensable as a tool, there is a stronger relationship between oratory and the conception of the good developed by Socrates' interlocutors, namely that both start from the assumption that pleasure is good and pain is bad.[4] The most forthright account—the account whose sincerity is least tainted by shame, 486E5–D7—is given by Callicles:

> The man who'll live correctly ought to allow his appetites to grow as large as possible and not restrain them, and through courage and intelligence ought to be competent to minister to them at their greatest and to satisfy every appetite with what it craves.… Wantonness, indiscipline, and freedom, if available in good supply, are virtue and happiness. (491E8–492C6)

I discuss this view of the good further in §2.3, but for the present it can be summarised as follows. Callicles believes one should satisfy all appetite because he believes, first, that pleasure and the good are identical (495A5–6) and, second, that the satisfaction of an appetite is, or is at least very closely associated with, pleasure (Socrates gives the most explicit statement of the latter, with Callicles'

[3] It is worth noting the criticism of democracy implicit in the dialogue's comparisons between the orator and tyrant. The formal power of the Athenian democracy was held by collective bodies, such as the assembly or courts, that consisted of a large number of citizens with equal authority. In this equal capacity, these members could only enjoy 'power' in the sense that they are a drop in a powerful river. Personal, non-collective power lay with those who could change the river's course by swaying popular opinion. While not official roles, orators were those who chose to devote themselves to regularly giving the speeches that influenced popular opinion, making them, if they were successful, the most powerful individuals in the *polis*. With respect to power, then, an orator can be seen as the democratic analogue of the tyrant and it is for precisely this near-tyrannical power—getting whatever one desires—that Gorgias, Polus, and Callicles value oratory. Plato is, then, criticising democracy on its own terms: it was a political system designed to avoid usurpation by oligarchy or tyranny, attempting to restrict decisions to collective rather than personal will; by portraying those who are effectively its leaders as aspirational tyrants, Plato implies that it gives rise to what it aims to avoid.

[4] While it would be going too far to say that Gorgias, Polus, and Callicles have the *same* account of the good, since they differ at least on justice, their accounts are in most respects consistent and the differences between them are largely not in their content but in the detail and clarity with which they express them. It is tempting, then, to see their accounts as greater and lesser developed versions of the same view and to attribute their differences on justice to the fact that Gorgias and Polus—who, in Callicles' opinion at least, defend conventional justice only out of shame—haven't developed their views sufficiently to see its full implications. I discuss in detail the relation between Polus and Callicles' conceptions of the good below.

assent, in his account of hunger and thirst: 496C6–E2). In short, then, the reason he believes the happiest life consists in maximally indulging one's appetites is that he believes this is how one can maximally experience pleasure and he believes, in turn, that experiencing pleasure is the good.

Two brief clarifications about Callicles' hedonism. First, while it is possible that Callicles recognises more refined pleasures or pleasures of the mind,[5] his position clearly centres on the kind of physical pleasures that typically satisfy appetites, epitomised by the pleasures of food, drink, and sex (see especially 494C1 ff.). Moreover, whatever pleasures he recognises, he believes that physical pleasure is *sufficient* for a happy life, as is clear from the fact that he admits, despite being displeased by the example, that a life devoted to the pleasure of endlessly scratching a perpetual itch counts for him as a happy one (494C6–D7). Second, Callicles says that happiness is to have many and large appetites *and* possess the virtues of courage and intelligence needed to 'be competent to minister to them at their greatest'. Callicles' ideal, then, is to constantly maintain many large appetites specifically in a way in which one can, and only if one can, constantly satisfy them. Note that this is not possible if one pursues immediate pleasures at the *expense* of one's future ability to either have or satisfy large appetites (e.g. overindulging to the point of illness). Callicles' view is a maximising one, requiring one to organise one's life so that one will both have appetites that are 'as large as possible' and the capacity to satisfy them, so that one has 'as much [pleasure] as possible flow in' (494B2).[6]

[5] See Rudebusch (1999, pp. 36–37) for evidence of these pleasures.

[6] Here I disagree with a number of readers who take Callicles' hedonism to involve pursuing short-term pleasure at the expense of greater long-term pleasure: 'Instead of working out the total consequences [Callicles] is led astray by the attractions of each pleasure as it occurs, and so gives himself up to a self-frustrating way of life' Gosling & Taylor (1982, pp. 70–71); 'Callicles is committed to the view that drinking cyanide is pleasant and good, provided one has a craving for the taste of bitter almonds' Klosko (1984, p. 132). See also Berman (1991, pp. 125–126). What most suggests this is Callicles' rejection of temperance and self-control, which might seem to imply that he is against any restraining of appetites, even for the sake of greater long-term appetite-satisfaction. But we should note that Callicles equates temperance with the insensible life of a corpse or stone, one in which we experiences no pleasure or pain at all, either now or in the future (492E5–6; 494A6–B1). Socrates, in turn, sees temperance as, ideally, a state in which we have all our appetites already satisfied, a state that affords no opportunity for pleasure, since

What we find, then, is that Callicles and oratory see eye to eye. Callicles' conception of a happy life begins with the assumption that the good is pleasure, the very same assumption that underlies flattery's ability to persuade. Oratory, then, is not simply an instrument in the orator's pursuit of the good, but a practise that uses arguments of a sort that he himself finds persuasive and that targets an audience that agrees, at least to the extent that they are persuaded, with the basic tenet of his conception of what's good. In the remainder of this section I will argue that there is a yet stronger relationship: oratory and the pursuit of pleasure are essentially the same practise, but applied to different subjects. One 'gratifies and pleases' the practitioner's audience; the other 'gratifies and pleases' the practitioner himself.

2.1.2 *Self-flattery*

In the closing lines of the *Gorgias*, in a summary of what can be concluded from the dialogue's discussion, we are told 'that every form of flattery, both *the form concerned with oneself* and that concerned with others, whether they're few or many, is to be avoided' (527C1–3; my emphasis). We get no indication that there is a form of flattery 'concerned with oneself' in Socrates' first statement of his account of flattery (462C–463D) but it appears, albeit obliquely, when Socrates returns to the topic from 500A ff. This passage occurs immediately after Socrates' most significant dialectical victory over Callicles: forcing him to admit, in refutation of his view of the good, that pleasure and the good are not in fact identical. Socrates uses this admission to claim that getting the good and getting pleasure are not, as Callicles would have it, one and the same practice, but two separate

pleasure can be had in the process of satisfying desires (on Socrates' view of temperance see §§2.3.3 and 2.3.4 below). There is no indication, then, that either Callicles or Socrates sees it as temperate to restrain one's appetites now for the sake of satisfying larger appetites later. Moreover, the condition that one must be competent, through courage and intelligence, to satisfy one's appetites at their greatest suggests that if one has large appetites that one *can't* satisfy, it is better to restrain them than continue having the burden of frustrated appetites. Indeed, Callicles suggests that the many who lack the required courage and intelligence are better served by conventional justice and temperance precisely because these virtues fit the more limited reach of their appetites (492A3–B1).

practices:

> Given that we're agreed, you and I, that there is such a thing as good and such a thing as pleasant and that the pleasant is different from the good, there is a practice for each of them and a procedure for obtaining it, the quest for the pleasant on the one hand and that for the good on the other. (500D6–10)

The context makes it clear that these practices aim to get pleasure or the good for *oneself* and not, or at least not only, for others. Socrates has just set out their discussion's central agenda: there are 'two distinct lives', one that Callicles recommends living and one that Socrates recommends living, and they must 'examine how they differ from each other and which of them is the one we should live' (500D3–4). The discussion, then, is centred on the clearly prudential question of which of two lives one should live and the opposing practices are introduced to clarify this distinction (thus, Socrates speaks of each as a 'procedure for *obtaining*' (*paraskeuên tês ktêseôs*) and as a 'quest' or 'hunt' (*thêran*) for either the good or the pleasant). However, despite this, it occurs alongside a restatement of the distinction between the practices that comprise flattery and the practices that truly benefit the body and soul. It is hard to see what relevance these have—Socrates' transitioning comments are unhelpful[7]—unless Socrates believes the following: one and the same practice, using the same knack, aims at pleasure whether it is providing it for oneself or providing it for others. This is consistent with the descriptions of these practices that we find in the passage, which capture what is common both to the two lives that Socrates and Callicles disagree about and the two sets of other-regarding practices introduced in the earlier discussion with Polus:

> There are some practices that concern themselves with nothing further than pleasure and procure only this, practices that are ignorant of what's better or

[7] The order of the discussion is as follows: (1) it is claimed that we should do all things for the sake of the good (as agreed with Polus earlier) and that discerning which pleasures are a good and which bad is a matter of craft (499E6–500A5); (2) Socrates recalls—introduced with no more than 'let's recall' (*anamêsthômen*)—the distinction between the flattery that aims at pleasure and genuine crafts that aim at the good (500A6–B5); (3) Socrates interrupts himself to plead for Callicles' sincerity, emphasising that their discussion (which discussion?) is about the important matter of which of two lives one should lead and introducing their corresponding practices (500B5–E2); (4) Socrates returns, now looking for Callicles' agreement, to the distinction between pleasure-aiming knacks and good-aiming crafts (500E3 ff.).

> worse, while there are other practices that do know what's good and what's bad. (500B1–3)
>
> [Some practices] are of the order of crafts and possess forethought about what's best for the soul, while others slight this and have investigated only … the soul's way of getting its pleasures. (501B3–6)

From one perspective this is not so surprising. If we consider flattery just insofar as it is a practice for providing pleasure, then naturally it is a counterpart to the practice for providing the very same thing for oneself. It is also highly plausible that mastering the practice of providing pleasure for others will at the same time enable one, by virtue of the same ability, to procure one's own pleasures. Certainly this is the case if they do indeed parallel the practices for providing and procuring the good, since surely we are to assume that only one ability is behind both of these practices, namely knowing what is good and bad for the body and soul.

From another perspective, however, it is indeed a substantial and surprising claim. Flattery is not simply a knack for providing pleasures. It is also a practise that fools its audience, leading them to false beliefs about what's good. Thus its branches include not only the straightforwardly pleasure-providing practices of pastry baking and cosmetics, but also the practices of oratory and sophistry. What is surprising, then, is that this suggests that the pursuit of one's own pleasures is a practice that is in some way comparable to oratory and sophistry—but what sense can be made of the idea of self-oratory or self-sophistry? The answer, I believe, is that the same account of persuasion—of how one forms certain false evaluative beliefs—underlies both flattery of others and flattery 'concerned with oneself'.

As we saw, flattery is described as a deliberately deceptive practice: it 'pretends' (*prospoieitai*, 464C7) to be one of the real crafts and 'hoodwinks' (*exapata*, 464D2) people with the lure of pleasure so that it seems to be best; pastry baking pretends to know what is best for the body (464D4–5); and cosmetics is 'illusion-producing' (*apatêlê*, 465B3), creating a false façade of beauty. We should note, however, that it is flattery itself—Socrates' personification of the practice—that is said to be the deceiver and not the orator or chef. Of course an orator can delib-

erately mislead his audience and, not being known for justice, probably will do so (although he is, like the pastry chef, primarily understood as someone who successfully provides pleasure for others: e.g. 462C8–9, 503C1–6, 517B2–C1). However, since they lack knowledge of the nature of their own practice, they are not party to the fundamental deception by which it works. Flattery's deception is that in providing the merely apparent good of pleasure it is providing what is best for body or soul and that, thereby, it is a craft like medicine or justice, working from knowledge of what's best. Far from thinking of them as deceptions, these are claims that form the fundamental beliefs of Socrates' interlocutors: Gorgias's attempt to define oratory ultimately fails because he treats it as a genuine craft; Polus confuses the apparent and real good, failing to see that doing what seems best is not the same as doing what is best; and Callicles defends flattery's central pretence that pleasure and the good are the same. The orator, then, is as much a victim of his practice's deception as his audience. He is in the same position with respect to the soul as the flatterer of the body who 'because of his character appears both *to himself* and to others a true minister to the body' (517E2–3; my emphasis).

One might wonder how an orator can be competent and yet be unaware of the deceptions that allow his practice to function. The answer lies in the role pleasure plays in flattery. The flatterer persuades *with* pleasure (464D2) and not by persuading his audience *that* pleasure is good. The pastry chef, for example, simply points to the fact that his practice gives pleasure while the doctor's causes pain to convince his audience that he benefits the body and the doctor harms it (521E5–522A3). Flattery, then, is a practice that persuades by taking advantage of the pre-existing fact that pleasure itself plausibly appears good: it produces something that tastes or looks pleasant, as in cookery or cosmetics, or speeches that emphasise the pleasure and pain in legal or political matters, as in oratory, and it thereby convinces an audience *already* inclined to take pleasure and pain as the measure of what's good or bad. Thus, the persuasiveness of pleasure comes prior to flattery and flattery persuades by exploiting the apparent goodness of pleasure.

Consider again the two apparently different accounts of oratory that we find in the *Gorgias*: in Socrates' initial discussion with Gorgias, oratory was defined as a 'producer of persuasion' (453A2), specifically that variety of persuasion that produces conviction without knowledge, while in his later account it was said to be a knack for 'producing a certain gratification and pleasure' (462C7). We can now see very clearly how these two accounts fit together: oratory is a producer of persuasion because, and just in so far as, it is a producer of gratification and pleasure.[8]

The fact that oratory is a producer of persuasion just insofar as it is a producer of pleasure helps explain how the flattery 'concerned with oneself' can be construed as a sort of 'self-oratory'. The beliefs of the tyrant or ambitious orator who devotes his life to pleasure are formed in precisely the same way, since they are persuaded by the same things, as an orator's audience: they fail to reflect on the good and so are persuaded by what appears good at first sight. They are, however, unlike the many insofar as they possess the knack of flattery, that is, they are especially good at 'guessing' at what will be most pleasant and so they can provide and procure pleasure more effectively than others. In its oratorical guise, this knack amounts to the ability to find, for example, the most compelling *prima facie* reasons for taking a course of action by drawing out, or taking an 'expert' guess at, the pleasures or pains it's likely to entail. By giving these reasons to others they will be able to lead others to pleasures; by giving these reasons to themselves they will be able to lead themselves to pleasures. As Socrates says of Themistocles and other orators of the past, they possess an 'excellence' that amounts to 'the filling up of appetites, *both one's own and those of others*' (503C4–6; my emphasis).

[8] Moss defends a similar view: 'flattery of all kinds is persuasive *because* it gratifies people's appetites, causing pleasure; correction of all kinds [e.g. Socratic refutation] is unpersuasive *because* it frustrates people's appetites, and causes pain' (2007a, p. 236). Where Moss emphasises the procedural role pleasure plays in persuasion—e.g. the orator's words persuade in part because he is, while speaking, giving his audience pleasure by, say, praising them—I would rather put the emphasis on the pleasures the orator points to in the content of his argument: e.g. he persuades his audience that *x* should be done by showing them that *x* is pleasant, thereby persuading them that it is good. On my view of why Socrates fails to persuade, see §2.2.2 and pp. 61–62.

What I hope I have shown so far is that two of the major themes of the *Gorgias*—the knack of oratory, on the one hand, and the life of pleasure epitomised by the tyrant, on the other—are introduced to answer the same question, since they represent essentially the same pleasure-guided practice. In both, the real persuaders are not orators, but pleasure and pain: these themselves are producers of conviction, a false conviction that what's pleasant is good and what's painful is bad. By investigating oratory and how it succeeds in persuading the masses with its appeal to pleasure and pain, the *Gorgias* is at the same time shedding light on what leads someone to envy the unjust but appetite-satisfying life of a tyrant. Now that we see that it is a topic explored in both of these sides of the *Gorgias*'s discussion, we can take a closer look at its principal assumption, the claim that pleasure and pain are what determine what most people, if they fail to reason appropriately, are liable to believe.

2.2 TAKING THINGS AS THEY APPEAR

Pleasure presents a *prima facie* appearance of being good. This claim leads to three associated claims in the *Gorgias*. First, pleasure is said to persuade specifically the ignorant, that is, those who lack the expertise needed to recognise what is genuinely good for the body or soul and, accordingly, to spot false appearances of these goods. For example, a merely apparent good state of the body, as we saw, makes one 'seem to be well in body, so that it would not be easy for one to see that they are not well unless one is a doctor or a gymnastic trainer' (464a4–6). Second, there is the remarkable claim that, to the inexpert, pleasure's apparent goodness is not just convincing, but is in fact a *more* convincing candidate for the good than what really is good. Finally, it seems that in the *Gorgias* Plato treats evaluative appearances as almost an exact inversion of the truth: what is bad appears good and what is good appears bad. Behind all these claims, I believe, is the fact that pleasure's false appearance is in a strict sense *prima facie*: pleasure is what appears good at first sight, before one has, in any active way, thought about

what's good or bad.

Consider first a passage in which Socrates describes what the crafts seem to be from flattery's perspective:

> If the soul didn't govern the body but the body governed itself, and if pastry baking and medicine weren't kept under observation and distinguished by the soul, but the body itself makes the judgements about them, making its estimates by the gratification it receives, the world according to Anaxagoras would prevail, Polus my friend—you're familiar with these views—all things would be mixed together in the same place and there would be no distinction between matters of medicine and health and matters of pastry baking. (465C7–D6)

Although this claim is presented counterfactually, it captures an important insight. The image we are given is of the body making its own judgements, unguided by a soul capable of making the discriminations needed to see the difference between pastry baking and medicine. What this is intended to capture, it seems, is the fact that since flattery and those under its sway fail to examine what is best for the body, the judgements they form will be no more than what can be unthinkingly concluded from one's bodily experiences, thus letting their body be the measure. (We find a very similar claim in the *Phaedo*: 'every pleasure and every pain provides, as it were, another nail to rivet the soul to the body and to weld them together. It makes the soul corporeal, *so that it believes that truth is what the body says it is*'; 83D4–7, my emphasis.) The body makes 'its estimates by the gratification it receives', which presumably means it judges what's pleasant to be good and what's painful to be bad. Why? The answer may be obvious but it is worth stressing: pleasure *seems* good and pain *seems* bad and they appear this way in a very immediate way that requires little or no thought. Consider the immediate, unthinking aversive behaviour of someone who places their hand in a flame, as if their hand itself has decided to move away. Thus, if one fails to reflect on them, the apparent value of one's bodily sensations will be automatically accepted—if something tastes good, then it is good and if it feels bad, it is bad. In general, the conclusion that pleasure is the good will be the default position for anyone who fails to examine, and so unreflectively accepts, what appears to be the case about the good of the body or soul. From such a perspective medicine,

the craft of the body's good, will be indistinguishable from pastry baking—that is, the craft of the body's good will *be* pastry baking—because the real good will collapse into the apparent good: what *is* good for the body will be indistinguishable from what *appears* good for the body, pleasure.

In the following, I hope to show just how central this error—the error of unreflectively accepting appearances—is to Socrates' account of both flattery and the view of the happy life his interlocutors defend. I begin by looking at what it means to judge by appearances (§2.2.1). I turn then to Socrates' analogy of the pastry chef and the doctor, where we find the claim that, to the inexpert, evaluative appearances are more persuasive than the truth (§2.2.2). Finally, I look at the ethical consequences of uncritically accepting appearances: the worst and most unjust life is mistaken for the best and most just life (§2.2.3).

2.2.1 *Judging by appearances*

The error of uncritically following appearances is exemplified in Polus's defence of his belief that a happy life is one in which we manage to get what *seems* best to us. Socrates challenges Polus's belief that orators have great power because 'like tyrants, they kill whomever they want and confiscate the property and banish from their cities whomever seems fit' (466B11–C2). He maintains that Polus has, unwittingly, made two claims, only one of which he accepts: orators and tyrants do not in fact do what they want (*ha boulontai*) even though they do indeed do what *seems* best to them (*ha dokei autois beltista einai*). Socrates' rather difficult argument for this distinction is, in rough summary, that what everyone ultimately wants is what is *in fact* good, so if what seems best to us is not in fact good, then it is not what we want (466A–468B).[9]

[9] This argument has raised a great deal of controversy, with many readers finding its conclusion difficult to swallow: how can one do what seems best to one and yet, without even realising it, not what one wants? There are three particularly influential responses. The first, introduced by Gulley (1965), is to weaken the claim—at the cost of textual fidelity—to a more palatable counterfactual: doing what seems best to us is not what we *would* have wanted, *had* we known it was bad. The second, found in McTighe (1984), claims that the argument trades on a fallacy: Socrates shows that 'we do not desire bad things' as a *de dicto* ascription of desire (roughly, what seems to us to be bad) but illicitly concludes that 'we do not desire bad things' as a *de re* ascription

Polus is extremely resistant to this distinction; he finds it 'outrageous' and 'monstrous' (467B11). He believes that being able to do whatever seems best is the happiest life and, so, most completely what one wants—clearly such a life *must* be good if doing what seems best is identical to doing what is best. Although Polus eventually concedes that there is a conceptual difference between what seems best and what one wants, he continues to deny that there is such a difference in practice. He concedes only the conditional claim that *if* orators and tyrants do what seems best to them, but is in fact bad, then they fail to do what they want, but he maintains that the antecedent is false. At the close of the argument Polus is unshaken in his belief that the best life is 'being in a position to do what seems best' (468E6–7); he believes, as Socrates expresses it, that 'doing what seems best coincides with acting beneficially' (470A10–11). In short, then, Polus is not prepared to believe that the evaluative world is complex and deceptive. He believes that what is good is just what it appears to be; the appearance of the good is no more deceptive than, say, the appearance of tables and chairs.

Polus's refusal to go beyond what seems good fits Socrates' general charac-

(roughly, what is in fact bad). I believe the right reading is a variant of a reading offered by Penner. Penner (1991) argues that there is a perfectly plausible desire that is independent of what seems good to us, the desire for the good 'whatever that is and even if it differs from what I think it is'. Unlike Penner, however, I believe that Socrates does not deny that we also desire what seems good to us. Consider the analogy of a detective who wants to catch the person who killed A. The killer is, unbeknownst to him, person B. Unfortunately, the detective acquires evidence that leads him to believe, falsely, that person C killed A. While it now seems true that the detective wants to catch person C, there is another sense in which C is not the person he wants to catch. The detective, we assume, simply wants justice to be served, so one desire that he has is the desire to catch *whoever it is* that killed person A. We can be certain at least that this is the desire that motivated him before he had a suspect; he might have said, for example, 'I don't yet know who killed A, but whoever it is, I want to catch them'. The crucial claim is that this desire (let's call it d^1) does not disappear when—is not superseded by—the desire to catch person C (let's call this d^2). Rather, d^2 is explained by and dependent upon d^1: the detective desires to catch C *because* he desires to catch whoever killed A and believes that C is this person. Catching C would satisfy d^1, but it fails to satisfy the detective's more fundamental and important desire, d^2 (more fundamental because d^2 explains d^1 but not *vice versa*; more important because, of the two desires, d^2 is the one the detective himself puts priority on satisfying). It is in this sense, then, that catching C is not what the detective wants. By analogy, we all desire the good, whatever it is, and for this reason desire what seems good, like pleasure. But pleasure is not what we want insofar as it does not satisfy the most fundamental and important desire we have, the desire to get what really is good for us.

terisation of the flatterer. Socrates emphasises that the flatterer is someone who proceeds unreflectively, 'without giving any thought to what's best' (464D1) and who has 'considered only … the soul's way of getting its pleasure, without examining (*skopoumenai*) which of the pleasures is better or worse' (501B5–C1; see also 465A1–2, 501A3–6, and 501C4–5). We should understand the idea that the flatterer gives 'no thought' (*ouden phrontizei*) to the good with an active sense of 'thought' in mind (in line with the *skopoumenai* at 501B8): he never questions what the good is, but instead passively accepts that it is just what it appears to be. Socrates complements this negative claim with a positive account of the type of thinking the flatterer does engage in. Firstly, the flatterer does not only fail to investigate what's good or bad, he is someone who in general proceeds without any examination, even of his pleasures. Flattery is described as something that is practised by 'a mind given to making guesses (*stochastikês*)' (463A7), that 'guesses (*stochazetai*) at what's pleasant with no consideration for what's best' (465A2). As a practice that merely guesses rather than looks for an account of its subject, we can understand why flattery can only be called an *empeiria*, a knack or ability based on experience rather than knowledge. We find a quite careful account of what this knack of 'guessing' consists in when Socrates restates his account of flattery for Callicles (500A7–501C6):

> Flattery proceeds towards its object in an altogether artless (*atechnôs*) way, without having at all examined either the nature of pleasure or its cause. It does so completely non-rationally (*alogôs*), with virtually no discrimination (*diarithmêsamenê*). Through habit and knack (*tribê kai empeiria*) it merely preserves the memory of what usually happens and that is how it procures pleasures. (501A4–B1)[10]

Here Socrates is unequivocal about flattery's failure to use reason: it doesn't consider the nature or cause of its subject, it proceeds non-rationally (*alogôs*), and it fails to discriminate (perhaps between good and bad pleasures: see 501B7–8). Accordingly, the procedure for getting pleasure that Socrates describes requires no, or at least minimal, use of reasoning. Expanded a little, it would seem it is the

[10] Compare Polus's theory that crafts are 'experimentally devised by experience, for experience guides our lives along the path of craft' (448C5).

knack of remembering what things have typically given pleasure in the past and then repeating what one remembers to get further pleasures. Such a technique relies only on experience and memory, thus earning it the title of an *empeiria*, a knack or something learned by experience. It is *alogos* insofar as it guides action without reflecting on these experiences or discriminating between them with reference to an account of their nature or cause.

Compare a passage from what is generally taken to be a much later dialogue, the *Philebus*. Despite the otherwise large philosophical distance between these dialogues, we find a description of a set of practices that has remarkable theoretical and linguistic similarities to the *Gorgias*'s description of flattery:[11]

> S: If someone were to take away all counting, measuring, and weighing from the arts, the rest might be said to be worthless. P: Worthless, indeed! S: All we would have left would be conjecture (*eikazein*) and the training of our perception through knack and habit (*empeiria kai tribê*). We would have to rely on the ability to make guesses (*stochastikês*) that many people call craft, once it has acquired strength through practice and hard work. (*Philebus* 55E1–56A1)

Again we have a set of practices that lack any rational procedure—'counting, measuring, and weighing'—and rely instead on mere 'knack and habit' to make 'guesses'. Socrates gives the example of musical proficiency, which is acquired not by reasoning or understanding but by pure experience: testing, observing, and remembering what gives the right and the wrong kind of sound (56A3–7). This remarkable similarity to the *Gorgias* tells us two things. First, and most importantly for the present context, it allows us to more confidently take Socrates' description of flattery to be as careful and defining as it seems. He is describing an important kind of cognition, which I will call judging by appearances, that has in particular two important features: it is *alogos* to the extent that it proceeds without reasoning and it is empirical—a kind of experientially-based guessing—to the extent that it is confined to uncritically accepting, and remembering and predicting, how things appear to one. Second, it indicates that this low form of cognition is consistently important for Plato, despite many other changes in his

[11] The comparison is also invited by what appears to be a back-reference to the *Gorgias* at 58A7–B6.

psychology. In the *Protagoras* we saw it in the form of being prey to the 'power of appearance' when one lacks 'the art of measurement'; in the *Republic*, as I argue in chapter 5, it reappears in the descriptions of *eikasia* in books 6 and 7 and Socrates' description of non-rational cognition in book 10.

It is worth also briefly noting a similar description Plato gives in the *Republic* of the limited thinking of another group of people confined to appearances, the prisoners in the cave who perceive only shadows. Among them, the one who is praised and holds power—very likely the orator or tyrant—is said to be:

> [T]he one who was sharpest at identifying the shadows as they passed by and who best remembered which usually came earlier, which later, and which simultaneously, and who could thus best divine the future. (*Rep*. 516C8–D2)[12]

Like the flatterer, the prisoners do no more than memorise what 'usually happens' and thereby make their estimates of what will happen in future. The prisoners are confined to this limited cognition because their restraints confine them to perceiving shadows, while explaining or knowing what the shadows represent requires contact with their originals. The flatterer, in comparison, deals only in what appears good, while making no attempt to examine what the real good is. Both amount to an account of what thinking is like if one's information is confined to appearances: taking one's experiences at face value and then using only memories of these experiences, perhaps combined into a general picture of what usually happens, to guide one's future actions. This is a process which presents no opportunity to discover what's behind how things appear and, so, will lead one to the conclusion Polus reaches, that the real good is identical to the apparent good.

The flatterer is of course not just anyone who judges by appearances, but someone who, like the one holding power in the cave, is an 'expert' at it, practising it professionally. Using his skill at 'guessing' at what's pleasing, he is able to gratify others and thereby persuade them that he's providing, and is the craftsman of, a good. In the next section I look at why the orator is so successful despite the fact that he presents people with what is a merely apparent good.

[12] I look at this passage in detail in chapter 4.

2.2.2 *The apparent versus the real good*

So far we have seen that the *Gorgias* presents two conceptions of the good, one being the view one gets if one reflects on what's good, like the doctor or philosopher, and the other being the view one gets if one fails to reflect on the good, like the pastry chef or orator. On reflection, it is surprising that the latter leads to a *specific* conception of the good. This is a far from trivial claim. One might think that failing to reflect on what's good is consistent with all sorts of views of what a good life is, as varied as where people's whims and fancies take them. On the other hand, one might think, as Socrates' interlocutors do, that the good is simply what it appears to be, having no more a misleading appearance than tables or chairs do. What we find in the *Gorgias*, however, is a more pessimistic view. The good is far from apparent and something specific that is not good but actually harmful appears good instead. Thus we live in an inherently and consistently misleading world, where we can only avoid what's bad by questioning what seems plainly to be the case. Worse still, for those who lack the relevant expertise, pleasure not only appears good but appears so *more* convincingly than what is good. Consequently, to the majority of people the flatterer is more persuasive than the genuine craftsmen he imitates.

This is a claim that appears a number of times in the *Gorgias*. Gorgias himself is the first to suggest it, making it the central point in his praise for oratory. He claims that 'there is no subject on which an orator would not speak more persuasively than any other craftsman, before a crowd' (456c4–6), where 'before a crowd' marks the one limitation on his power, being later defined as before those who lack knowledge of the subject the orator speaks on (459a4). So persuasive is oratory, Gorgias maintains, that even if it came to deciding who was to be appointed doctor, an orator could make a more convincing case for being fit for the job than a real doctor (456b6–c2).

Socrates agrees, although he does not value this ability as highly as Gorgias. He adapts Gorgias's example of the orator competing with a doctor, replacing the orator with a pastry chef:

> Pastry baking … pretends to know the foods that are best for the body, so that if a pastry chef and a doctor had to compete in front of children, or in front of men just as foolish as children, to determine which of the two, the doctor or the pastry chef, had expert knowledge of good food and bad, the doctor would die of starvation. (464D4–E2)

Dodds takes the emphasis of this passage to be the fact that it is a competition judged by 'men just as foolish as children' and, thus, sees it as a 'parody' of Gorgias's earlier claim (1959, p. 228). I doubt, however, that Gorgias—and even more so Polus or Callicles—would feel the sting of this remark, which is more an insult to the majority of Athenians under the sway of the orators than the orators themselves. Gorgias's contention was only that oratory was more persuasive 'before a crowd' (he even makes sure to add this when Socrates leaves it out, 459A3) and he readily admits this means before those who do not know. Before a crowd, we should remember, means before the assembly or law-courts, before the Athenian citizens, and it is precisely within this political arena that Gorgias boasts oratory's unchallenged dominance. What will sting, however, is the idea that the ability for which Gorgias believes orators should be revered is shared by pastry chefs—the orator's ability is as unsophisticated as enticing children with tasty morsels. That is, he is more persuasive before a crowd for no more impressive a reason than the fact that he, like the pastry chef, waves enticing pleasures before a pleasure-hungry crowd.

Socrates is of course not suggesting that adults will be fooled in medical matters by a pastry chef. Pastry baking is not the most harmful branch of flattery and the significance of Socrates' comments comes from the fact that it is analogous to flattery of the soul. As Irwin says, Socrates is suggesting that 'most people's moral and political judgement is as ignorant as children's judgement about healthy foods' (1979, p. 134). Socrates' aim, of course, is not simply to show the ethical ignorance of the majority but to point out what, given that they lack knowledge of what's good, they are liable to be convinced by and how such conviction comes about. Like children charmed by a pastry chef, when adults are fooled by flattery, it is the attraction of pleasure that fools them.

Socrates returns to the analogy of the doctor and pastry chef later in the dia-

logue, giving us a more detailed picture of the case he envisages the chef will make. This time, in a clearly premonitory passage, Socrates says that because he aims at what is best and not at what is most pleasant (521D9–E1) he would fare badly if *he* were to be brought to court:

> For I'll be judged the way a doctor would be judged by a jury of children if a pastry chef were to bring accusations against him. Think about what a man like that, taken captive among these people, could say in his defence, if somebody were to accuse him and say, 'Children, this man has worked many evils on you, yes, on you. He destroys the youngest among you by cutting and burning them, and by slimming them down and choking them he confuses (*aporein*) them. He gives them the most bitter potions to drink and forces hunger and thirst on them. He doesn't feast you on a great variety of sweets the way I do!' What do you think a doctor, caught in such an evil predicament, could say? Or if he should tell them the truth and say 'Children, I was doing all those things in the interest of health,' how big an uproar do you think such 'judges' would make? Wouldn't it be a loud one? (521E3–522A7)

Although he may mislead, the pastry chef does not in fact lie to the children about what he or the doctor does. He does not, for example, claim that the doctor causes pain gratuitously or for a malicious reason or that he himself gives them any greater benefits than pastries. Nor does the doctor deny the charges laid by the chef (522A5–6) and it is not even clear that the children doubt the doctor's defence, that he provides health (although they certainly fail to appreciate the connection between health and what's genuinely good for the body, as Polus fails to see the connection between justice and what's good for the soul). Rather than about empirical facts, the disagreement is about the evaluation of what the doctor and chef do. The competition is between which condition, pleasure or health, is the body's real good. The case that convinces the children is simple: the doctor causes them pains and the pastry chef gives them pleasures. The chef merely has to point to the fact—he offers no argument—that the doctor causes pains and he gives them pleasures; once he has pointed these facts out, the apparent value of pain and pleasure will do the rest of the work and he can leave the children to draw their own conclusion.

Why is the pastry chef's case more persuasive than the doctor's? Part of the reason is that the good the pastry chef is offering has a *prima facie* appearance

that can be grasped by children in a way that the good the doctor offers, health, can not. Tasty pastries catch a child's attention and the putative good they offer, pleasure, is something a child can understand. The doctor, in contrast, can't simply point to the value of health, he needs to explain it to the children and they will, quite likely, find it challenging to understand him. More seriously, the doctor's practices do not merely fail to seem good, they actually appear *bad* to the children, on account of the pain they inflict.[13] One might think, especially with dialogues such as the *Protagoras* or *Republic* in mind, that despite the pain of its means, health will in fact result in greater pleasure overall. However, unlike discussions of pleasure in other dialogues there is no attempt in the *Gorgias* to deny that Socrates' idea of a happy life is not the most pleasant life. Indeed, it is not even said to be *a* pleasant life and is consistent with great strife and pain, most prominently the possibility of suffering injustice such as unfair trial, physical punishments, and even death.[14] Socrates says he would be 'fool' to deny this possibility, at least 'in this city' (521C7–8).[15] In any case, even if the dialogue does leave space for a more complex view of pleasure, its concern is the fact that pleasure and pain are not a good *guide* to what's good and bad. Even if the genuinely happy life is after all pleasant, it is certainly not pleasant in the direct, conspicuous way necessary to appear good *prima facie*. Thus, just as the children will only see the pain the doctor causes, Polus and Callicles only see the strife of the just life (for example: they 'focus on [just punishment's] painfulness, but are blind to its benefit and are ignorant of how much more miserable it is to live with an unhealthy soul than with an unhealthy body'; 479B6–8).

[13] See Moss (2007a) who defends a similar interpretation.

[14] We get a straightforward example of something being better yet more painful in Socrates' earlier defence of the claim that doing is worse than suffering injustice (474B1–475E6), where it is concluded that even though suffering injustice is more painful it is nonetheless better than committing injustice. See p.63 below. Again, in the discussion of which of the two forms of 'protection' (*boetheia*) is best (509C–522D)—protection from suffering injustice or protection from committing injustice—it is taken for granted that the just man lacks the former protection, leaving himself open to unjust and certainly painful treatment.

[15] At the outset of the discussion Socrates mentions a power that offers both protections, protection against both committing and suffering injustice, but we never learn what this is (509C6 ff.). We might suspect that it requires not only being just, but being just in a just city.

The merely apparent good, then, is more persuasive to most people because it rules at a *prima facie* level: if it is not critically assessed, it appears good while what genuinely *is* good remains largely hidden or, worse still, seems like something to be avoided. This is the basic error that the orator and his audience make. It is also the seed of the more serious error of failing to see the value of justice and, at its extreme, finding value in injustice. This is the error that is illustrated, with increasing severity, in the lives that Polus and Callicles defend.

2.2.3 *The appearance of justice and injustice: Polus and Callicles*

The doctor and the pastry chef reveal a remarkable feature of appearances. Appearances seem to present the inverse of the truth: something bad (unhealthy pastries) appears to be the good and something good (medical treatment) appears to be bad. Is it strictly the case that evaluative appearances are an inversion of the truth? To answer we need to take a closer look at Socrates' and his interlocutors' different evaluations of the two candidates for the good life. What I wish to emphasise is that their disagreement is not simply about what is good and bad, but about what is best and worst, the happiest and most miserable life. Following what merely appears good does not only lead one into error about the good, but leads us towards what Socrates sees as the worst possible life.

So far we have seen that pleasure leads those without knowledge to a false conception of what the best life is. What has not yet been discussed is the fact that the unfettered pursuit of pleasure is, as Callicles' vision of the good life illustrates, an intemperate and unjust life, the life exemplified by the most vicious of tyrants. Socrates sees such uncontrolled injustice as not only bad but the 'greatest' (*megiston*) or 'utmost' (*eschaton*) of evils:

> Doing what's unjust is second among the great evils (*tôn kakôn megethei*); not paying what's due (*didonai dikên*) when one has done what's unjust is by nature the greatest and first of all bad things. (479D4–6)[16]

Callicles has, then, mistaken the very worst life for the very best. With the harm-

[16] This claim is repeated throughout Socrates' discussions with Polus and Callicles: 469B8–10; 476A4–6; 479C1–3; 480D5–6; 482B4; 509B1–3; and 522E3–4.

fulness of injustice being a common thesis in Plato it is easy to fail to appreciate just how prodigious Callicles' error is: in searching for what is best in life, Callicles has not only missed it—not surprising in itself—but has ended up actively pursuing the most extreme evil, the greatest source of misery, a human can suffer. Conversely, he believes the best life is harmful: justice requires the frustration of the pursuit of what he thinks best, pleasure, and leaves one unprotected from the worst evil, suffering injustice. This inversion of values—the directly opposing ethical stances Socrates and his interlocutors hold—is a prominent theme in the *Gorgias*. The majority of Socrates' discussion with Polus is devoted to convincing him that his values are ordered the wrong way around, specifically that suffering is better than committing injustice and being justly punished is better than avoiding punishment, rather than *vice versa* as he believes. In his discussion with Callicles, as we saw, Socrates argues that Callicles' best life is in fact the worst and Callicles, for his part, defends the opposite view.

Callicles is fully aware that this is what Socrates is doing. He is aware that Socrates is not saying only that his view of the good life is false but making the stronger claim that he has got the good and the bad back to front:

> If you are in earnest and these things you're saying are really true, won't this human life of ours be turned upside down and won't everything we do evidently be the opposite of what we should do? (*allo ti ê hêmôn ho bios anatetrammenos an eiê tôn anthrôpon kai panta ta enantia prattomen, hôs eoiken, ê ha dei;*) (481C1–4)

What Callicles is referring to here is Socrates' response to Polus's claim that tyrants, even if unjust, are happy because they posses the 'power' to do whatever seems best to them, a position which, as discussed above, follows from Polus's insistence that doing what seems best is identical to doing what is best. Socrates' response is the claim that the extent to which one is *just*, rather than the extent to which one gets what seems best, determines one's happiness. On this new scale, Polus's model of happiness, the unjust tyrant, turns out to be a model of the worst misery (470E4–471A3). As the discussion continues, Socrates adds finer grain to this re-ordering by isolating two claims that Polus denies: (1) suffering injustice is better than committing injustice and (2) if one does commit injustice

it is better to suffer just 'punishment' or 'discipline' (*kolazein*) than continue unhindered. Thus, Socrates has ordered four states from best to worst: being just and not suffering injustice; being just and suffering injustice; being unjust and 'suffering justice'; and being unjust and not 'suffering justice'.

In contrast to Callicles, Polus has no systematic criticism of justice and his problem with (1) and (2) is not that he believes there is something good about injustice or bad about justice *per se*. Instead, his argument is simply an attempt to make vivid for Socrates the harm he sees in his candidates for the worst states, suffering injustice and being justly punished.[17] His idea of harm is intense pain: his attempt to convey to Socrates the most heinous, misery-causing harm has him imagining one suffering, and seeing one's family suffer, the likes of being put on the rack, castrated, tarred and impaled (473B12–D2). His argument against (1) and (2), then, is exceptionally simple and a perfect example of the oratorical 'knack': just like the pastry chef criticising medicine, he attempts to bring to the forefront of Socrates' mind the pain involved in suffering injustice or punishment, in the belief that its apparent badness will ring as true for Socrates as it does for him.

Polus himself finds this appeal to pain eminently plausible. One of the most remarkable things about Polus's reaction to claims (1) and (2) is the sheer degree of his incredulity, which is certainly among the most pronounced of any of the interlocutors in Plato's dialogues: he thinks that even a child could refute Socrates (471C4–5); that Socrates should find the happiness of Archelaus, the unjust tyrant, 'immediate' or 'obvious' (*autothen*; 470E1–2); that Socrates couldn't be serious that he disagrees: 'you are just unwilling to admit it, for it really seems to you

[17] Polus seems to treat these as equally bad. At one point Socrates attributes to Polus the view that being unjust and justly punished is the greatest evil: they are investigating 'whether a wrongdoer's paying what is due is the greatest evil (*megiston tôn kakôn*), as you were supposing, or whether his not paying it is even worse, as I was.' (476A4–6). However, earlier Polus treats suffering injustice as the greatest evil: 'Socrates: Because doing what's unjust is actually the greatest evil (*megiston tôn kakôn*). Polus: Really? Is *that* the greatest? Isn't suffering what's unjust even worse?' (469B8–10). This inconsistency is resolved if suffering injustice and being justly punished are equally worst for Polus insofar as they share, indistinguishably, what he sees as the worst form of harm.

to be as I say it is' (471E1); he simply laughs at the claim that a tyrant, being unjust, is miserable, and declares that Socrates is saying something 'the likes of which no human being would maintain' (473E4–5); and even his final, conceding statement grants only formal agreement and makes it clear that his intuitions are entirely unchanged: 'I think these statements are absurd (*atopa*), Socrates, though no doubt you think they agree with those expressed earlier' (480E1–2). Why is Polus's incredulity so pronounced? An explanation can be found in Polus's defining error, his failure to see the difference between something that very much seems to be the case and a conspicuous truth. Just as his idea of the good arises from his belief that what seems best is best, what he takes to be the most miserable state is what *seems* to be the worst. We might have some sympathy for Polus here. Extreme pain *does* seem very bad and it is not at all hard to see why someone might believe that their worst fate would be to suffer the horrendous tortures Polus lists. On the other side, the idea that being unjust is worse than this—that even living an unjust life of luxurious enjoyment is worse than being viciously tortured, but just—is not, to put it mildly, immediately plausible. It is a highly inconspicuous truth: it is not simply different from but the opposite of what is strongly suggested by an initial appraisal, where pleasure and pain are the measure. For one to begin to accept this as the truth, one must grasp lengthy and difficult arguments, no less than the arguments to which Plato devotes the *Gorgias*. Thus, Polus finds the falsity of Socrates' position to be 'obvious' and childishly easy to demonstrate because he is looking at evidence that is immediately plausible and intellectually unchallenging, evidence that would convince most people as easily as the pastry chef convinces his jury of children. On the other hand, what Socrates is claiming is something initially *im*plausible, very intellectually challenging, and the opposite of what seems 'obviously' to be true.[18]

While following how things seem leads Polus to hold the contrary of (1) and

[18] This explanation for why Polus remains unpersuaded, which can be equally applied to Callicles, strikes me as sufficient. We do not need in addition the (entirely speculative) suggestion that his and Callicles' resistance to rational persuasion is due to countervailing non-rational motivations, a claim found in Klosko (1983), Scott (1999), Woolf (2000), and Moss (2007a).

(2), his position falls short of being a complete inversion of Socratic values. It falls short because he does not claim that there is anything essentially harmful about justice or essentially beneficial about injustice. While he believes that the just course of action is often inconsistent with doing what seems best and so a hindrance to happiness (e.g. 471A4–D2), he does not think that whether one is just or not has an essential bearing on one's happiness: one is happy as long as one can do what seems best, which could be either just or unjust (468E6–469A1). It is this chink in Polus's position that eventually allows Socrates to force him to agree (although it is a purely dialectical victory, since he still finds what he is agreeing to 'absurd'). In brief, Socrates gets Polus to agree to the traditional assessments of justice and injustice as admirable and shameful respectively and, therefore, to admit that doing injustice is at least less admirable than suffering injustice. Socrates then draws his attention to the two things one can mean in calling something 'admirable': something is called 'admirable' only when one thinks it is either beneficial or pleasurable. Polus agrees that committing injustice is not more *painful* than suffering injustice, so cannot be less admirable in this respect, and this allows Socrates to conclude that it must be so in virtue of being more *harmful* than suffering injustice.

Callicles takes up the gauntlet where Polus left off, presenting a more complete view of the good by filling the gap in Polus's position, namely his lack of a consistent account of the relation between the good or bad life and a just or unjust life. Callicles believes Polus has, out of shame, admitted something he doesn't really believe and he accuses Socrates of having achieved this through the 'trick' of drawing conclusions about the nature of justice from premises that were only intended to be about conventional views of justice. Callicles sets out to fix this error and aright the life Socrates has 'turned upside down'. He argues that the conventional views of justice and injustice are radically mistaken. He defends two related claims: first, he rejects Socrates' use of the terms 'just' and 'unjust', 'admirable' and 'shameful', arguing that what is really—what is by 'nature' (*phusis*)—just and admirable is what Socrates calls unjust and shameful and *vice*

versa; second, he rejects Socrates' evaluation of what he calls justice and injustice, arguing that, in fact, the former is harmful and the latter beneficial.

Callicles believes Socrates' conception of the just and admirable is nothing but law or convention (*nomos*), a convention devised by the inferior majority to prevent a superior minority from exploiting their natural advantage. He contrasts this popular idea of justice with what he believes to be just by nature, a state in which the superior rule over the inferior and 'get more' (*pleon echein*) than them. Callicles characterises the 'superior', after some cross-examination, as those who possess two virtues: they are intelligent in the affairs of the city—the central skill of a powerful orator or tyrant—and possess the courage required to accomplish what they want 'without slacking off because of softness of spirit' (491B4). What the superior man will use his intelligence and courage to 'get more' of is the satisfaction of his appetites, which, as we saw, he identifies with pleasure. Callicles claims that this natural justice is in fact the same as what Socrates and the many, misled by convention, call *in*justice: 'they say that getting more than one's share is 'shameful' and 'unjust' and that this is injustice, to seek to get more than others' (483C4–5, cf. 492B9–C3).[19]

Thus, even though Callicles agrees with the premises of Socrates' argument against Polus, at least in word—that is, he believes both that injustice is shameful and that it is so because it is bad, for, as he says, 'by nature all that is worse is more shameful' (483A7–8)—the fact that he calls 'justice' what Socrates called 'injustice' means he is led to the opposite conclusion: as they are conventionally understood, injustice is beneficial and justice is harmful. The kind of benefit and harm he has in mind is revealed by his distributive account of justice and injustice: conventional justice aims at an equal share while conventional injustice gives a larger share to the stronger. Thus, while it is in the weaker majority's interest to seek justice, since an equal arrangement leaves them with more goods than they

[19] This differs from the account of natural justice in not mentioning the rule of the superior. Nonetheless as the discussion continues it would seem that Callicles considers ruling to be part of natural justice only insofar as it happens to be the best means of getting more than others, so what is definitive of natural justice, as of conventional injustice, is simply getting more.

would otherwise have (483c5–6), for the stronger few to act justly is against their interest, since they are sufficiently courageous and intelligent to get more than an equal share. Acting justly for them would simply be a way of getting less of what's good.

Provocative as Callicles' view of justice is, it is a reasonable view to have *if* his view of the good is correct. He believes that pleasure and the good are identical and that the greatest pleasure is achieved by satisfying many and large appetites without restraint. Since Callicles' account of pleasure barely goes beyond the bodily pleasures of satisfying appetites for food, drink, and sex—an account that Socrates never contests—the value of justice and injustice can only be measured by its instrumental role in getting such pleasures or pains (crucially, it leaves no space for the possibility that being just is intrinsically pleasurable). Consequently, if he is right about what's good, he is also right to believe justice is harmful and injustice is beneficial, at least on his distributive interpretation of them: justice requires people to frustrate their appetites in the name of equality, getting less pleasure and, therefore, being worse off; injustice allows one to indulge any appetite without curbing them for the sake of others, getting as much pleasure as one's abilities allow, and therefore living happily.

Callicles' position is a complete inversion of Socrates'. What is for Callicles the best life, a life of unrestrained injustice, is for Socrates the very worst life—'by nature the greatest of all and first among bad things' (479d6). Again it is worth emphasising how surprising this result is: Callicles' has acquired the most harmful beliefs one could have, beliefs that lead one to the worst misery, just by following to its natural conclusion his belief that the good and bad are simply what they seem to be. So despite how unconventional his position is, the error he starts with is the same as the common error made by the majority who are taken in by oratory, a majority who lack knowledge of, and fail to reflect on, what's good or bad and therefore uncritically follow pleasure and pain's *prima facie* appearance. Of course Callicles' immoralism is not as widespread as the mistake that leads to it—he himself criticises the many for their adherence to conventional morality—

but this is hardly surprising. Callicles is unusually immune to shame, an emotion that leads most people, including Polus, to feel the pull of conventional morality. Moreover, he is undoubtedly unusually clever and his views, although mistaken, are sophisticated enough to be beyond the reach of most people. Nonetheless, even if most people do not go as far as Callicles, his view shows us the direction in which they are led to the extent that they fail to examine what's good or bad, and instead take them to be simply how they appear. The world has a natural evaluative grain that leads in the direction of injustice and unhappiness; to achieve justice and happiness one needs to go against this grain, questioning what very much seems to be the case.

2.2.4 *Conclusion*

While it is unwise to declare any one thing to be the theme of a complex dialogue like the *Gorgias*, the last two sections have shown that there is one clear thread that runs through the dialogue's various topics. This is the attempt to spell out what people are liable to believe if they lack knowledge of what is good or bad or, more precisely, lack the craft that grants that knowledge. The *Gorgias* looks at both the cause and the result (and, in less detail, the cure). The cause is the highly persuasive but merely apparent values of pleasure and pain—and the far from obvious values of what really is good and bad—coupled with most people's assumption that evaluative facts are relatively obvious, the kind of facts that are revealed by a *prima facie* assessment. The result, or at least the direction in which it leads, is the highest of all possible errors, believing that the unhappiest and most unjust life—'the utmost of all evils' (482b4)—is the best life. This is an ambitious theory, a theory that offers an almost comprehensive account of why humans err. It is also a theory that we found the seeds of in the *Protagoras* and that, as I will argue in the following three chapters, is given its most detailed defence in the *Republic*.

In the next section I turn to a new question. A major difficulty of the *Gorgias* is that it seems to present elements of two inconsistent moral psychologies, the

intellectualist psychology defended in *Protagoras* and the very different psychology we find in middle-dialogues like the *Republic*. The reading I've developed so far is consistent with both of the two most common attempts to resolve this apparent inconsistency, but in the remainder of this chapter I will suggest that it lends itself especially well to the view that the *Gorgias*'s psychology is consistently intellectualist.

2.3 THE PSYCHOLOGICAL THEORY OF THE *GORGIAS*

Whereas the *Protagoras* uses appearances to explain the specific error of *akrasia*, the *Gorgias* uses them to give a more comprehensive account of why and in what way humans are liable to err. Rather than the failure to stand by one's better judgement, it explores the beliefs that comprise the majority's 'better' judgement, the beliefs about the good and just that guide their day-to-day lives. As we have seen, Plato is highly pessimistic: what appears best is very different from what is best and most people, since they fail to reflect on what's good and bad, end up basing their beliefs on the false appearances presented by what's pleasant and painful—appearances that, if taken as a comprehensive measure of what's good and bad, make the most miserable and vicious life seem to be the most happy and virtuous. What I wish to focus on here is the fact that this account can form the basis of a challenge to traditional explanations that attribute error to misleading desires rather than misleading appearances.

From what we have seen so far, the *Gorgias* presents pleasures and pains as especially error-inducing not, as is usually thought to be the case, because it is in our nature to desire them but because they appear to us in a deceptive way. We desire pleasures because they seem good and avoid pains because they seem bad. The order of explanation goes from appearance to belief to desire: what's pleasant appears good, persuading us to believe it is good, and thereby causing us to desire it (since we do indeed naturally desire what's good). As such, this form of explanation is an alternative to explanations that appeal to good-independent,

non-rational desires that respond to pleasures and pains irrespective of our beliefs. Consider the following example. Children will typically choose to eat unhealthy sweet foods like pastries rather than less tasty healthy foods. How do we explain this? A natural explanation might be that since children's reasoning abilities are not sufficiently developed to guide them, they are prone to following their non-rational appetites instead, appetites that by nature aim at things like pleasant foods. However, an equally plausible explanation is that, since they have less developed reasoning abilities, children are likely to be guided by beliefs that are based on what's easily understood or immediately apparent, such as the salient fact that pastries taste good, while they are less likely to grasp more abstract, rationally-demanding facts, such as the long-term benefits of healthy foods. These two explanation are not incompatible, but they are individually sufficient. Most importantly, the latter explanation can sufficiently explain children's preferences for pastries even if the very existence of belief-independent, non-rational desires is denied.

Explanations of error of the kind we find in the *Gorgias*—errors arising from false beliefs about what's good or bad, acquired by uncritically accepting what appears good or bad—provide an account of a broad range of the practical mistakes people are liable to make, from childish foibles to the grave injustices of tyrants. This includes at least the majority of errors for which wayward desires are typically held responsibly, specifically those errors that are thought to be caused by the desiderative temptation of pleasure and pain. The most common non-rational desires that are thought to lead one astray are appetites: desires that are naturally fixed on the basic biological necessities of food, drink, and sex. It is no coincidence that Socrates takes the same basic pleasures of food, drink, and sex to be what have the most immediate appearance of being good. Wherever an error is blamed on a non-rational passion for some pleasure, Socrates can point to an alternative explanation, compatible with intellectualism, that appeals instead to the pleasure's compelling false appearance of goodness.

Thus far, then, there seems to be a neat story of the development of intellectu-

alism through the *Protagoras* and the *Gorgias*. The *Protagoras* deals with a narrow but pressing challenge to intellectualism, that of explaining the error of *akrasia*, and does so by appealing to an epistemic conflict caused by immediate pleasure and pain's 'power of appearance'. While *akrasia* poses a serious challenge to intellectualism, it is only one practical error amongst many. Most of the time when we err it is not because we act *contrary* to what we believe best but because our beliefs are themselves mistaken. The *Gorgias* fills this gap by developing an account of how our beliefs are easily led astray by pleasures and pains, not because our beliefs are corrupted by our lower desires but because pleasures and pains present an untrue and yet tragically plausible picture of what's valuable. Neat as this story is, it faces a serious obstacle: the *Gorgias* has elements that suggest it departs from intellectualism, particularly in the prominent place it gives to 'appetites' (*epithumiai*) that seem to be, as appetites are usually understood, non-rational desires for basic pleasures like food, drink, or sex.

2.3.1 *The two psychologies of the* Gorgias

Early in the dialogue we find an argument (466A–468B) that, while puzzling in other ways, includes a relatively straightforward defence of Socratic intellectualism. Specifically, it defends the claim that human desires (*boulêseis*) and actions aim only at good things, never at bad things or things that are neither good nor bad. As the dialogue proceeds, however, it seems that a more complex psychology emerges, containing elements that are not obviously compatible with what is defended in 466A–468B. We find the introduction of appetites—denoted, with near consistency, by *epithumia* rather than *boulêsis*—that seem to be a distinct subset of our desires that aim at pleasures, particularly bodily pleasures like food, drink, or sex. In itself this might not be a problem, since such desires could be a subset of the desires found in 466A–468B, being desires for pleasures *qua* apparent goods. But Socrates also argues that to be temperate we need to 'restrain' (*eirgein*) or 'discipline' (*kolazein*) our appetites; if they are undisciplined, then they become insatiable and will need to be arduously pandered to if we are

to avoid the pain of having large, unsatisfied desires (492E–494A). Restraining or disciplining would seem unnecessary if our appetites are simply a subset of inherently well-behaved desires for the good, desires that invariably fall in line with our better judgement—surely, one might think, the only desires that need to be restrained are those that are liable to conflict with our judgements? Finally, there are two striking similarities between the appetites in the *Gorgias* and the desires of the appetitive part of the soul in the *Republic*: first, appetites have their source in a distinct part of the soul—'that part of the soul in which the desires (*epithumiai*) are' (493A3–4)—and, second, we are told that having temperate, disciplined desires involves a certain 'structure and order of the soul' (504D1), suggesting that virtue in the *Gorgias*, as in the *Republic*, consists in the correct order between potentially conflicting parts of the soul.

One might reasonably suspect, then, that in the *Gorgias* (or at least in its latter half) Plato takes a step away from the intellectualism of dialogues like the *Protagoras* towards a moral psychology closer to that of the *Republic*. Compelling as this is at first sight, it faces some serious problems. The most pressing problem is that, as it stands, it leaves the *Gorgias* with a fairly blunt contradiction: it defends a non-intellectualist moral psychology in its second half (491D ff.) and yet also defends intellectualism earlier in the dialogue at 466A–468B. This makes the *Gorgias* seem an untidy sort of transitional dialogue, a dialogue that straddles Plato's early and middle moral psychologies not by defending a coherent middle-ground position but by employing both of these incompatible positions at once. It seems reasonable to take rendering the dialogue consistent, or at least giving an account of why Plato might have tolerated such an inconsistency, to be a criterion of a successful interpretation of the dialogue's psychological claims.

This leaves two avenues for resolution: either the *Gorgias*'s talk of appetite and discipline are not in fact incompatible with intellectualism or 466A–468B does not in fact defend intellectualism.[20] A defence of the latter might note that

[20] There is a third position, namely that Plato deliberately added inconsistencies to the *Gorgias*. This is the position taken by Cooper (1999). Cooper agrees that '[a]nyone who does suppose that Socrates adopts the Calliclean [i.e. non-intellectualist] moral psychology in his confronta-

Socrates uses a different word for desire in the claim that we only 'desire' (*boulêsis*) good things than he uses for the putatively good-independent appetites (*epithumiai*). It might be suggested, then, that 466A–468E defends not intellectualism but the more limited claim that we always have desires of a certain sort, *boulêseis*, for what we take to be good, even if we can in addition have a non-rational sort of desire, *epithumiai*, for objects we believe to be bad. I will have more to say on the difference between *boulêsis* and *epithumia*, but an initial objection to this reading is that 466A–468E is about action as much as it is about desire:

> It is because we pursue what's good that we walk whenever we walk; we suppose that it's better to walk. And conversely, whenever we stand still, we stand still for the sake of the same thing, what's good … it is for the sake of the good that those who do all these things do them. (468B1–8)

'These things' is of course a finite set of examples, but the examples Socrates chooses do not seem to carve out any special category of action. This strongly suggests that the argument is intended to be entirely general, to apply to all actions—could there be an end other than the good that we act for the sake of yet never *walk* for? Any doubt is allayed later in the dialogue when, reaffirming the conclusion of this argument, Socrates states unambiguously that what had been agreed between him and Polus was 'that the end of *all actions* is what's good' (499E8–9, my emphasis). (It should be noted that this explicit reaffirmation of intellectualism occurs during his discussion with Callicles, alongside the seemingly incompatible view of appetites and restraint.) Thus, if *epithumiai* were non-rational, good-independent desires, the argument of 466A–468E would entail that they are incapable of motivating even the most basic actions and, therefore, that they

tion with Callicles will have a difficult, not to say impossible, task in constructing a coherent overall interpretation of this dialogue on that basis' (p. 60). He also recognises that the majority of the putatively non-intellectualist elements in the dialogue seem, at least, to be introduced and developed by Socrates himself, not by Callicles. Thus, the reading Cooper offers is an ingenious attempt to free Socrates of any commitment to the claims that seem to undermine his intellectualism, while attributing them instead to an indirect conversation that Plato, through tensions in the position he gives to Socrates, has with his reader. As the arguments I make below will make clear, I do not think Cooper succeeds. If he were right, we would expect to find in the *Gorgias* some psychological position or claim that is sophisticated enough and sufficiently defended to challenge intellectualism. Unless we borrow from the *Republic*—a move which a number of Cooper's claims rely on—I do not believe we can find any such position or claim in the *Gorgias*.

are desires deprived of their *raison d'être*.

This evidence weighs in favour of resolving the inconsistency in the direction of intellectualism, questioning whether the putatively non-intellectualist elements of the *Gorgias* are as they seem. I've pointed to four of these elements: appetites, restraint, psychic partition, and psychic order. Each presents us with complex interpretive problems and a detailed examination of all four would be too lengthy an undertaking for this chapter. I will focus, therefore, on appetites and appetitive discipline. Appetites present the most important challenge to an intellectualist reading, which after all comes down to the question of whether in the *Gorgias* Socrates does or does not countenance the existence of non-rational desires. I will argue that the account of appetites he defends is in fact an intellectualist one, an account that is consistent with the view of appetites that we find in the *Protagoras* (and even, in a way, the *Republic*—a point I return to below). I begin by looking at the role appetites play in Callicles' view of the good life and then turn to Socrates' response, the so-called allegory of the leaky jar, which is often taken to be the best evidence that the *Gorgias* defends a psychology close to that of the *Republic*. Finally, I examine the account of 'restraint' (*eirgein*) or 'discipline' (*kolazien*) that Socrates offers.[21]

[21] Some brief comments on the two putatively non-intellectualist elements I won't examine. The evidence for *Republic*-like partition amounts to just one line: 'that <part> of the soul where the desires are' (*tês de psuchês touto hen hô epithumiai eisi*; 493A3–4). This is a line that introduces an allegory that itself has parts, in particular a part, a jar, that represents our desires. Why should this suggest the *Republic*? The only dialogue that argues for a partless soul is the *Phaedo* and even here this is achieved only by assigning part of our mental life to the body. We have no reason to think that intellectualism needs a partless soul. Like almost all theories of the mind, it makes more sense if it assumes parts of some kind, at the very least a conceptual division between desiring and reasoning. What is innovative about the *Republic* is not that the bare fact that it argues that the soul has parts, but the *kind* of parts it argues for: three parts that are each autonomous sources of motivation, with separate and potentially conflicting goals. Unless we look through the distorting lens of the *Republic*, 493A3–4 in no way suggests a similar theory. The same is true of the *Gorgias*'s claim that virtue requires a certain organisation (*taxis*) or order (*kosmos*) of soul. As with partition, there is nothing in the *Gorgias* itself that has any significant resemblance to the *Republic*. There is, for instance, no reference to either harmony or conflict between parts of the soul. Nor is there any attempt to link the brief mention of a desiring part with the idea of an ordered soul: the allegory refers to a desiring part but makes no mention of an ordered soul—only of an ordered life (*ton kosmion bion*)—and, likewise, in the discussion of order (503D–504D) we don't find a single mention of parts of the soul. It would be nearly impossible for a reader unaware of the *Republic* to conclude that Socrates has in mind a harmonious relationship between

2.3.2 *Callicles and appetites*

It is worth first noting that the word translated 'appetite', *epithumia*, can refer to a range of desires beyond what we would usually call appetites. In the *Republic*, for example, even the rational part of the soul's desires are called *epithumiai* (e.g. 431B9–D2, 580C7–8) and in the *Gorgias* itself we hear of a non-appetitive *epithumia* 'to do good in return' (520E8–9). That said, while it may not mean appetite, strictly speaking, in the *Gorgias* 'appetite' is often an apt translation of *epithumia*. What is often intended by the term is something that is co-extensive with what we typically take to be appetites: desires like thirst, hunger, or lust. Does this, in itself, pose a problem for an intellectualist reading of the dialogue? If appetites are, by definition, non-rational desires—if 'rational appetite' is an oxymoron—then of course the very mention of appetites would preclude an intellectualist reading. This, however, is too quick. Just as not all non-rational desires need be appetites, not all appetites need be non-rational. Appetites have many other, arguably more defining features that make them the sort of desires they are. Most importantly, they are desires that are commonly for bodily pleasures like food, drink, or sex. They are also desires that are often strong, impulsive, or urgent and they often arise spontaneously rather than after reflection.

It seems reasonable to think that having these or some set of these characteristics is sufficient for calling a rational desire an appetite: if one has a strong, urgent desire for sex, arising from little or no reflection, would we not call this an appetite even if it follows, and depends on, a hastily-conceived judgement that this is the best action? Indeed, Socrates can accept even the claim that one of the characteristic features of appetites is their role in psychological conflict, the same psychological conflict that the many understand, mistakenly, as motivational forces driving us in opposite directions. As we saw in the last chapter, accounting for this conflict within an intellectualist psychology is precisely what Socrates does in the *Protagoras*, so the major objection to the idea of rational

soul-parts. For two good discussions of order that look at the evidence within the *Gorgias*, see Woolf (2000, p. 32 ff.) and Carone (2004, p. 81).

appetites—that appetites can conflict with our judgements of what is best—has already been sufficiently dealt with, at least to the extent that we find the *Protagoras* plausible.

Consider, then, the desires that the *Gorgias* attributes to orators, tyrants, and the many. As we've seen, those who follow mere appearances give 'no thought' to the good and instead automatically accept what appears good or bad. Their beliefs arise with such little reasoning that they endorse pleasures as if 'the body itself makes the judgements about them, making its estimates by the gratification it receives'. The result is that they desire merely apparently good pleasures, such as the appetitive pleasures of food, drink, and sex. If we allow a broad conception of what an appetite is, such desires are, or at least have among them, appetites: they are aimed at the characteristic objects of appetites, basic pleasures; they are spontaneous and unreflective, mere automatic responses to what appears at first sight; and they can of course be as strong, urgent, and impulsive as desires for lower pleasures commonly are.

Importantly, then, to suggest that appetites in the *Gorgias* are rational is not to say that the *Gorgias* makes appetites something they are not: calm desires that follow careful deliberation about our overall good (as we might sometimes, mistakenly in this context, understand 'rational' to imply). Rather, it is to claim that appetites, while remaining a unique subset of our desires with distinctive features and co-extensive with what we typically think of as appetites, can *nonetheless* be analysable as one kind of desire for what we judge good. With this in mind, let us look at how and why appetites appear in the *Gorgias*.

Appetites first appear in Callicles' account of the good. He believes, in summary, that one should aim to satisfy every appetite since he believes, first, that pleasure and the good are identical and, second, that pleasure is the result of satisfying an appetite. Consequently, he sees the happiest life as one in which appetites are maximally indulged since this is how one can maximally experience pleasure and maximally experiencing pleasure is the good. In one sense Callicles' focus on appetites is unsurprising, given that he treats them simply as desires for

what he thinks best: pleasures and especially bodily pleasures, the characteristic objects of appetites. However, this does not fully explain why Callicles places so much importance on desires. Socrates' idea of the good life, like any, will equally involve its own desires for what's best, yet these desires do not play a prominent role in his account of the good. Why, then, are appetites so central in Callicles'? The answer, I believe, is found in the fact that pleasure is associated in particular with the *satisfaction* of an appetite.

To see why desire-satisfaction is important for Callicles it is helpful to draw a distinction between two kinds of object of desire. The pleasures that Callicles takes to be good are the sort of objects of desire, which I will call 'consumable', for which the activity of satisfying the desire is identical, at least extensionally, with the object of desire: hunger, thirst, or lust are both *for* and *satisfied by* the activities of eating, drinking, or sex. In general any desire for pleasure *per se* will be for a consumable object. Thus, Socrates observes that someone will 'stop having pleasure at the same time as he stops being hungry or stops having the other appetites' (497c5–6). Not all objects of desire are like this. For example, if I desire honour this can be satisfied by performing honourably in a battle, but the honour I desire does not last only as long as the battle; I will still have the honour I gained when I return home. Roughly, an object of desire will be consumable if it is something we do or experience—e.g. pleasure, dancing, contemplating, or listening to music—and non-consumable if it is something we achieve—e.g. having a just soul, knowing Greek, or owning a house. This distinction marks one of the fundamental differences between Socrates' and Callicles' views of the good.[22]

Since he takes the good to be pleasure and, *qua* consumable good, we experience pleasure only when satisfying our appetites for it, Callicles takes the extent to which one can achieve the good to be proportional to the extent to which one has appetites. One must, he believes, both constantly have appetites to satisfy and have appetites that are 'as large as possible', where presumably the larger one's

[22] I am grateful to Raphael Woolf for first suggesting to me that a distinction along these lines is needed and not, as I had believed, a distinction between desiring a limited and unlimited quantity of something. Berman (1991, pp. 123–125) defends a very similar account.

appetites, the more pleasure one derives from satisfying them. Unsurprisingly, then, Callicles rejects Socrates' reasonable-sounding rejoinder—the argument for which I'll consider shortly—that it is surely better to already have everything one wants, to have all one's desires 'filled up'. Reasonable as this sounds, Callicles rightly sees that *if* what's good is pleasure, it is false: 'the man who has filled himself up has no pleasure any more, and when he's filled up and experiences neither joy nor pain, that's living like a stone' (493A6–9). In other words, while it is true that if one believes a *non*-consumable object like honour is the good then having satisfied one's desire will indeed be the better state, since one will simply have got what one wants, the same is not true for a consumable good like pleasure. If pleasure is the good it will be better to be satisfy*ing* one's desires, but as soon as they are satisfi*ed* no good will remain to be enjoyed. This, then, is the reason why desires play such a prominent role in Callicles' conception of the good life: what he takes to be good, consumable pleasure, is a good that is only present while we are satisfying one or other of our desires for this good. The good life, then, is a life of constant desire and constant satisfaction.

So far, then, Callicles' appetite-driven conception of the good does not explicitly threaten an intellectualist reading of his position. What is distinctive about the desires he celebrates is the fact that their object is pleasure, which is a consumable object of desire and the characteristic object of appetites. Nothing suggests that it is important that they have the capacity to motivate us independently of our beliefs about the good. Far from it in fact: they are clearly desires rationally endorsed by Callicles. He desires what be believes to be good and he desires it *because* he believes it to be good. My claim is of course not that Callicles is purposefully developing an intellectualist account of the good life, but only that, by not raising anything that requires appetites to be non-rational, he presents a theory that Socrates can engage with from an intellectualist standpoint.

But does Socrates engage with it from an intellectualist standpoint? Many have thought that he does not. Socrates presents two arguments against Callicles: first, the so-called allegory of the leaky jars (492E–494B) and, second, a long

series of arguments aimed at driving a wedge between the good and the pleasant (494E–500A). The second, if anything, supports an intellectualist reading. Consider how Socrates sees its conclusion. Having shown that the pleasant and the good do not always coincide, he asks Callicles whether they must choose (*haireteon*) good pleasures and avoid bad ones:

> For Polus and I thought, if you recall, that we must do (*prakteon*) all things for the sake of what's good. Do you also think as we do that the end of all action is what's good and that we must do (*dein … prattesthai*) all other things for its sake, but not it for their sake? (499E6–500A1)

There are two reasons to take the verbal adjective *prakteon* (and the corresponding *haireteon* at E4) to imply necessity rather than prescription ('cannot but do' rather than 'ought to do'). First, this fits the purely descriptive 'the end of all action *is* what's good'. Second, the argument that Socrates is recalling here is 466A–468E and this concluded that we act for the sake of the good as a matter of psychological necessity. If this is right, then Socrates' conclusion seems to be that since, first, they now know that the good and the pleasant are not identical and, second, everyone acts solely for the sake of (what they take to be) the good, then we must—cannot but—pursue good pleasures and avoid bad ones, if we really believe that the former are good and latter are bad. If knowing that a pleasure is bad is sufficient to cease desiring it, this suggests that all our desires, even appetites for pleasures, are necessarily responsive to our beliefs about what is good or bad.

In stark contrast, however, his first response to Callicles, the allegory of the leaky jar, is often taken to be the very best evidence *against* an intellectualist reading. I believe that it has been greatly misunderstood.

2.3.3 *The allegory of the leaky jar*

The aim of the allegory is narrow. It doesn't directly challenge Callicles' view of *what* we should desire—that is covered in Socrates' second argument—but rather criticises, irrespective of their object, the *kind* of desires Callicles thinks we should have (notice that neither pleasure or any other particular object of desire is mentioned in the allegory). In Socrates' words, the allegory is an attempt to

persuade Callicles 'to choose instead of the insatiable (*aplêstôs*) and indisciplined (*akolastôs*) life, the orderly life that is satisfied and content with what it possesses at any time' (493c5–7).

Socrates gives us an allegorical representation of the soul and a description of the unique characteristics of this soul in unthinking or foolish people (*anoêtous*). The allegorical soul has two main elements, a jar representing 'that part of the soul in which the appetites (*epithumiai*) are' and a sieve simply said to represent 'the soul' of unthinking people (but far easier to understand if it represents another part of the soul). In foolish people both their jar and their sieve are leaky. The jar's leakiness represents the insatiability of unthinking people's desiring part and their sieve's leakiness represents a certain lack of intellectual retention, 'a lack of conviction and forgetfulness'.[23] The pedagogical force of the allegory comes from the twofold struggle unthinking people must endure to keep their jars full: they must constantly replenish their jars because they are constantly leaking and they can only carry water to their jars with a leaky sieve. Socrates' hope, then, is that the difficulty this leakiness causes—put with more force in the second part of the allegory: they must 'fill them constantly, both day and night, or suffer the utmost pain' (493e8–494a1)—will convince Callicles that having satiable desires is preferable to having insatiable desires.

There are a number of surface features of the allegory that fit an intellectualist psychology well. It is, first of all, an allegory in which motivational conflict never occurs—the allegory represents unthinking people's struggle to *fill* their appetites but they never struggle *against* them. Second, the allegory highlights faults in thinking and belief, rather than any non-rational defect, to explain why desires become insatiable: indisciplined people are unthinking; the sieve is leaky because it is 'unable to retain anything on account of lack of conviction and forgetfulness'; the jar (*pithon*) is so called because it is 'a suggestible (*pithanon*) and persuadable thing'. The mention of persuasion is especially interesting, given it is

[23] Presumably, as with the jar, the wise have an unleaky part corresponding to the sieve, perhaps a bowl.

the central theme of the *Gorgias*. As we've seen, most people—orators, tyrants, and the many—are 'unthinking' insofar as they fail to reflect on what's good or bad and, therefore, they are easily taken in by the compelling false appearance of goodness presented by pleasure. There is, then, a very simple link between our liability to be persuaded and our liability to desire pleasure: we are especially susceptible to believing that pleasure is good and, therefore, to desiring it. If the allegory's appetites are non-rational, on the other hand, it is more difficult to find a link: if anything, non-rational appetites would seem to be unique in *not* being responsive to persuasion, since they are capable of driving us towards pleasures even when we are convinced that this is something bad. Consider also the role of the sieve. Despite being said to represent 'the soul', it is easier to understand if it represents a second part of the soul—in the allegory it is a distinct part with a distinct function—and with the desiring part spoken for, a plausible suggestion is that it represents reason, especially since its leakiness represents a failure in *intellectual*, rather than desiderative, retention. If this is correct, the allegory presents the agreement of desire and reason as a permanent feature of the soul: the fact that the sieve supplies the jar with its water implies, first, that the fulfilment of an appetite requires reason and, second, that this is a task that desire and reason join in harmoniously.[24]

These, however, are just hints of intellectualism. The real battle-ground is the allegory's central claim, that the appetites of the unthinking are insatiable. Our

[24] Cooper comes to the opposite conclusion (1999, pp. 60–63). He agrees that the sieve represents reason, but he takes the fact that it merely serves to fulfil desires to suggest that reason is 'being co-opted by the enlarged appetites and serving their bidding', that is, that appetites are strong non-rational desires that distort our view of what we should desire. This is an odd conclusion to draw. It is in an *intellectualist* psychology that reason's role is to aid the fulfilment of our desires: our desires motivate us towards whatever is good and reason guides these desires with beliefs that 'tell' them what is and is not good. Intellectualism is, to this extent, a Humean psychology. On the other hand, it is only in moral psychologies like that of the *Republic* that it becomes necessary and possible for reason to deny some desires fulfilment—only when non-rational desires are introduced can what we judge best and what we desire come apart, creating something that reason might need to oppose. Cooper, then, has got it back to front: the jar's passive and dependent relation to the sieve neatly fits a psychology where appetites must necessarily follow our beliefs and it deprives them of the freedom that we would expect them to have if they were reason-independent, non-rational desires.

aim is to understand what this insatiability is. Let me first set out two straightforward, yet not always met, interpretive requirements of a successful reading. The first one is obvious: it must fit insatiability as it is represented in the allegory. The second, however, might be missed: it must not only accurately describe Callicles' view of the good—it is, after all, a comment on his position—but do so in a way that Callicles can accept. We should not lose sight of the fact that the allegory fails to convince Callicles and fails because he happily *agrees* that the life of pleasure is insatiable in the way that Socrates describes. In the imagery of the allegory, he *wants* to have a leaky jar. His response to the allegory (quoted in part above) makes this very clear:

> You do not persuade me, Socrates. For he who has filled himself up has no pleasure any more, but that … is to live like a stone, since to the extent he has filled himself up, he feels neither joy nor any pain. Rather, living pleasantly consists in this: having as much as possible flow in. (494A6–B2)

Guided by these two requirements, I hope to show that insatiability is a characteristic of the general desiderative behaviour of someone who believes, as Callicles does, that pleasure is identical to the good. As we saw, for Callicles happiness depends on constantly satisfying many and large appetites, since pleasure is found in the process of satisfaction. Pleasure is a 'consumable' object of desire. Thus, if the good is pleasure, we should devote our life entirely to satisfying appetites to the greatest possible extent, ceaselessly—insatiably—pursuing pleasure-giving activities like drinking or eating.

There are two controversial features of this reading. First, it does not attribute insatiability to token appetites of hunger or thirst. What can be called insatiable is the most general Calliclean desire, the guiding desire for pleasure, or the *set* of individual appetites—hungers, thirsts, lusts, and the like—that each contribute to a pleasant life. In contrast, a token hunger or thirst is satisfied after a certain amount of eating or drinking. Second, far from being a characteristic of non-rational desires, insatiability is a characteristic of one's desiderative behaviour that arises if and only if one holds a certain belief about what's good: believing that the good is pleasure (or, more precisely, that it is any consumable

object). Let us turn, then, to how insatiability is represented in the allegory.

Having an insatiable desiring part of the soul is likened to having a leaky jar that must be constantly replenished if one is to avoid the pain of having unfulfilled desires. What sort of insatiability does this suggest? A first thought might be that a desire is insatiable if it is a desire for an endless or infinite amount of something. This is certainly one way to have insatiable desires but it would be a poor way to challenge Callicles' position: he is not advocating futilely desiring infinite pleasure, but rather getting as much pleasure as possible. Nor does it fit the image of the leaky jar. If what it represented were wanting, insatiably, an endless amount of something then a fitting image would be having jars that are too big or infinite, so that the foolish will constantly labour to fill their jars without ever getting close to the brim, while the wise, having more modest desires, would have small jars that are easily filled.

Leakiness is fitting for a different type of insatiability: desiring not an endless amount of something but desiring something that is constantly trickling away. In other words, desiring a finite amount of something that diminishes over time and so can only be achieved through constant replenishment. If we take the goal of the water-carriers to be to have a full jar, both the wise and the foolish person can achieve this. As long as the foolish person has a constant supply, filling his jar at the rate it's leaking away, his jar will always be full. The problem with his desires, then, is not that they are futile, that is, that their object is unachievable. The foolish person can enjoy a *certain* sort of satisfaction, in the sense that he can achieve the optimal state that his desire demands: having a sufficient, constant supply of his ever-receding object of desire and, so, never *frustrating* his desire. But this state, the correlate of a full jar, disappears as fast as it is achieved and so it is maintained only through the labour of endlessly catering to an endlessly demanding desire. The foolish person can, then, be successfully satisfy*ing* his desires, but this process of satisfaction will never be complete, so he will never be satisfi*ed*.

It should be perfectly clear what feature of Callicles position this relates to: the fact that he desires, as the one and only good, a consumable object of desire.

The experience of pleasure lasts only as long as we are satisfying our desire for it, so as soon as we cease satisfying our desire it will disappear or diminish, just like the water in the leaky jar. Of course not just any desire for a consumable object will be insatiable; if I desire, say, to dance or read I am usually satisfied after a finite amount of dancing or reading. Such a desire becomes insatiable only if we want it in continuous supply: desiring something consumable (corresponding to leaking water) without cessation (corresponding to keeping the jar full) will lead to a desire that will never finish being satisfied. Now, in the life Callicles bids us to live the *only* desire that fits this description is the general desire for pleasure and the reason this desire fits is, perhaps surprisingly, *because* it is a rational desire. The reason Callicles desires to ceaselessly enjoy pleasure is that he believes it to be identical to the good. Whatever we judge to be our highest good, the source of human happiness, will *ipso facto* be something we desire to have without cessation throughout our whole lives. It simply makes no sense to want to take a break from the good. This is not true of any token appetite like hunger or thirst. (Consider the following comparison: on some interpretations of *Nicomachean Ethics* X.7–8 Aristotle identifies engaging in the activity of theoretical contemplation with the highest human good and, therefore, identifies it as what we should strive to engage in as much as is humanly possible, doing other activities only insofar as they are means to maximising our opportunities for contemplation. Irrespective of whether this is in fact Aristotle's view, this is a convincing example of a rational insatiable desire: contemplation is a consumable object of desire and, since it is desired *qua* the highest human good, it is clearly desired as something to be engaged in as much and as often as possible. This should certainly give those who associate insatiability and non-rationality pause for thought.)

This implies that hunger, thirst, or any other token appetite is satiable, but this is exactly how it should be. Callicles' position is that the happy man should satisfy 'whatever he may have an appetite for at a time' (492A2–3); he might get his pleasure from, say, eating at one time but sex at another. This makes perfect sense, since what Callicles counts as good is not food or drink *per se*, but

the pleasant life that such things as food and drink can contribute to. He is not, then, suggesting we eat as much and as constantly as possible, but rather that we eat the amount of food that is fitting to the most pleasant life possible. Furthermore, Socrates himself treats appetites as satiable, as is clear from his fairly detailed description of hunger and thirst (496D–497E). For example, he asks 'doesn't each of us *stop being thirsty* and stop feeling pleasure at the same time as a result of drinking?' (497B1–2; my emphasis). Callicles agrees that we do and, on reflection, it is hard to imagine anyone disagreeing: surely nobody can have a thirst that is unquenchable. Most decisively, in the allegory itself insatiability is not applied to token appetites. Rather, what is said to be insatiable is the *part* of the soul that contains appetites (493B1–3) and the *life* of the foolish person (493C5–6). It is not, then, any one of the foolish person's many appetites—his hungers, thirsts, lusts and the like—that is insatiable, but the *life* devoted to fulfilling appetites that he leads, that is, the endless task of satisfying a desire for pleasure.

Thus, the desire for pleasure *qua* good is the only viable candidate for the desire the allegory is an objection to: it is a desire that Callicles explicitly professes to, it straightforwardly fits the allegory's account of insatiability, and since it is the foundation of Callicles' position, it is a highly appropriate target for Socrates' objection. On the other hand, the idea that insatiability is a characteristic of non-rational desires finds considerably less purchase in the text. The claim, it seems, would be that the allegory is suggesting that encouraging or failing to restrain non-rational appetites causes them to become insatiable. If this were right, the allegory leaves a considerable gap, since it is a claim that is never explicitly made or defended (it certainly requires defence: I may have a very large hunger, satisfied only after eating a great deal, but is it plausible that by always feeding and never restraining my hunger it could become truly insatiable, so that I remain hungry no matter how much I eat?). Nor is it a claim that we would expect Callicles to accept. Having insatiable non-rational appetites would only hinder the pursuit of pleasure: insatiably desiring food *in addition* to insatiably desiring pleasure

would only lead one to constantly suffer the pain of a frustrated hunger when pursuing other pleasures. Most conclusively, I believe, the non-rational reading of insatiability seems to be redundant: it strikes me that the overriding purpose of the Calliclean life, the desire and pursuit of pleasure, is a sufficiently obvious candidate for insatiability for it to be highly surprising if Socrates were to choose any other candidate.

2.3.4 *Discipline and indiscipline*

What, then, of Socrates' claim that we must 'restrain' or 'hold back' (*eirgein*) and 'discipline' (*kolazein*) harmful appetites? I believe there is little doubt that Plato is well aware that talk of 'restraining' appetites will bring to mind repressing or holding back reason-resistant appetites, but when we look at what Socrates says about restraint and discipline we see that he silently rejects this conception by offering his own, very different one.

The first thing we should note is the fact, often lost in translation, that the verb *kolazien* that is first used for 'punishing' in the discussion with Polus (where it seems to refer to fairly standard legal punishments) is the same verb used for 'disciplining' bad appetites in the discussion with Callicles. Carone notes this and comments that 'it is conceivable that even at 505B [quoted below] Socrates thinks of "restraining" the appetites as a way of "punishing" them' (2004, p. 77). I believe there is a yet closer connection between the two uses of *kolazien*: the discussion of appetitive disciplining is just a more nuanced treatment of the same punishment discussed earlier.[25]

Socrates sought Polus's agreement to two claims: (1) suffering injustice is better than committing injustice and (2) if one does commit injustice it is better to suffer just punishment than to continue unhindered. This resulted in a hierarchy of worse to better states of the soul: being unjust and not 'suffering justice' (i.e. being justly punished or disciplined); being unjust and suffering justice; being

[25] For two discussion that similarly treat the earlier and later uses of *kolazien* as synonymous, and similarly take the later use to represent Socrates' considered view, see Rowe (2007) and Moss (2007a).

just and suffering injustice; and being just and not suffering injustice. Callicles begins his part in the discussion by expressing his indignation at these claims, accusing Socrates of turning life 'upside-down'. Nonetheless, Socrates slowly brings Callicles to the same conclusion, but via a different route and with new, finer-grained terms. 'Punishment' was earlier left undefined, apart from a few mentions of traditional punishments like imprisonment,[26] and the only explanation of its benefit was an application of the dubious principle that the qualities of an act are transferred onto what is acted upon, so if someone is *justly* punished, they are thereby made more *just* (476B–477A). Returning to the topic with Callices, with a more detailed account of our psychology, Socrates can spell out in more detail the benefit just punishment is supposed to have. Now the need to *kolazein* the unjust is introduced not as a response to criminality *per se* but to a soul in a state of *akolasia*, a lack of appetitive discipline (e.g. 505B11–12). Moreover, the *subject* of discipline or punishment is not only a whole person or soul, but more specifically one's appetites (492D5–6; 507E2). One might think that these two uses of *kolazien* have a merely homonymous similarity or at most flag an analogical connection—*prima facie* restraining a desire has little to do with, say, a flogging. However, there is good evidence for a more substantial connection, namely that instilling desiderative discipline—causing a person to desire the right things—is the benefit that Socrates believes just punishment should confer. Socrates makes this connection at the end of his discussion of discipline with Callicles:

> Each of us … must flee away from indiscipline (*akolasian*) as quickly as his feet will carry him, and even better must make sure that he has no need of being disciplined, but if he does have the need, either he himself or anyone in his

[26] Socrates mentions flogging, fines, imprisonment, exile, and execution (478C8–D3), although he doesn't quite commit to the idea that these are all just punishments. Each is introduced conditionally: for example, he says that *if* one's injustice merits a flogging, then one should present oneself to be flogged. In this earlier discussion, his unconditional conclusion is what's important: whatever it is that cures injustice, whether or not it is physical punishment, it is something that should be endured. Rowe (2007) argues convincingly that Socrates appeals to traditional punishments only for the sake of dialectical simplicity, while later introducing the true cure for injustice, which is a form of rational persuasion. While I will not argue for it separately, this is a natural implication of the account of discipline/punishment I develop below.

> house, either a private citizen or a whole city, then what's due must be applied and they must be disciplined (*epitheteon dikên kai kolasteon*). (507D1–5)

This mirrors almost exactly the conclusion Socrates reached with Polus (480A–D). The only exception is that previously it was simply 'injustice' that one must flee, but now this is replaced with the particular aspect of an unjust soul, *akolasia*, that discipline affects. Notice also that the final *kolasteon* blurs the line between desiderative disciplining and traditional punishment (accordingly, the phrase *epitheteon dikên* picks up on a phrase used earlier with Polus: *didonai dikên*, 'to pay the penalty'). Lest he has missed it, Socrates explicitly draws Callicles' attention to this connection and to the fact that these are the conclusions he poured scorn on earlier:

> These conclusions, Callicles, are all those previous ones, the ones about which you asked me whether I was speaking in earnest when I said a man should be his own or his son's or his friend's accuser, if he's done anything unjust, and should use oratory for that purpose. (508B3–C1)

As the connection with punishment would suggest, the various conditions of the appetites can be mapped on to the earlier worse-to-better hierarchy: giving free reign to indisciplined appetites; having indisciplined appetites that are *being* disciplined; and having temperate appetites.[27] The two extremes of this hierarchy, unrestrained indiscipline and temperance, are consistent with the understanding of 'appetites' I proposed in the previous section, since neither involve the sort of restraint that one might exert on non-rational appetites. Notice, for example, the distinction in 507D1–3 between undergoing the process of being disciplined and the better state of simply having 'no need of being disciplined'. These two extremes are the two desiderative states that we find in the allegory of the leaky jar: the foolish, indisciplined people who struggle to fill their appetites but never struggle *against* them and the wise who are temperate and live 'a life that is satisfied and content with what it possesses at any time' (493C6–7; see also 493E) and therefore have 'no need of being disciplined' because they have

[27] What about the fourth state, being just but unjustly punished? Presumably the truly just man's appetites cannot be corrupted by unjust discipline/punishment, so this fourth state drops out of consideration on a scale that measures only one's degree of discipline or indiscipline.

no bad desires that ought not to be satisfied.[28] What we need to look at is the middle state, *being* disciplined/punished. If anywhere, it is here that we will find restraint of the sort practised by, for example, Aristotle's enkratic man who has appetites urging him to act contrary to his better judgement but through an exertion of will-power manages not to act on them. At first sight, this seems to be exactly what Socrates has in mind:

> So long as it [sc. the soul] is evil, being foolish and undisciplined and unjust and impious, it should be held back (or 'restrained'; *eirgein*) from its appetites and not permitted to do anything other than what will make it better...Isn't holding it back from what it has an appetite for to discipline (*kolazein*) it? (505B2–9).

Out of context, this could easily be thought to describe a non-rational process: restraining one's wayward non-rational desires in the hope that, through repeated training, their force will diminish. In context, however, it is clear that Socrates has something very different in mind. The account of discipline just quoted describes for the soul what, in the previous few lines, the doctor does for a patient's body, namely keeping them away from appetites that harm his health (504E–505A). The patient's appetites are 'restrained' only in the sense that the doctor prevents him from satisfying them. What he is describing is not *self*-discipline or *self*-restraint, but discipline that, as with punishment, one person imparts on another.

In turn, the doctor is introduced as an analogue for the main topic of discussion in this part of the dialogue, namely the genuinely 'skilled and good' orator who, when writing his speeches, will 'pay attention to how justice may come to exist in the soul of his fellow citizens and injustice be gotten rid of, how temperance (*sôphrosunê*) may come to exist there and indiscipline (*akolasia*) be gotten rid of' (504D9–E3). A little earlier this kind of orator was distinguished from Themistocles and other famous orators by the fact that he satisfies only 'those

[28] Principe (1982) points out that this is implied by the image of the 'watertight' or 'restrained' (*stegein*) jars of the wise. For both the foolish and wise, how close their jars are to the brim represents how close their desires are to satisfaction. Only the wise, however, have jars that are constantly full, which implies that all their desires—including the appetites the jar represents—are satisfied. Thus, the wise can not have any appetites that ought to remain unsatisfied. Principe includes some illuminating diagrams to illustrate the point.

appetites that, once full, make a man better, and not those that make him worse' (503C7–D1).[29] The good orator can not, as a doctor might with a patient, physically hold his audience back from what they desire. The tool available to him is persuasion, so if he is to change his audience's desires he must do so through their beliefs. The simplest explanation, then, is that an orator 'disciplines' his audience's appetites by using persuasion to lead them to true beliefs about what they should and should not desire.

That this is the context is not in the least bit ambiguous. The passage flows smoothly from the improving effect that the genuinely 'skilled and good' orator has on his audience's desires, to the analogy of the doctor's response to his patient's harmful desires, to a general account of what disciplining desire is. This train of thought also leads neatly into Socrates' next comment. This comment has been taken to be in some way not serious or metaphorical,[30] but again it only seems so if we ignore the context. Immediately after the quoted passage, Callicles becomes uncooperative and elicits the following response from Socrates:

> This man won't bear being benefited and undergoing the very thing the discussion is about, being disciplined (*outos anêr ouk hupomenei ôpheloumenos kai autos touto paschôn peri hou ho logos esti, kolazomenos*; 505C3–4).[31]

What Callicles is undergoing is an argument in favour of the benefit of a just life. Socrates can only mean that being disciplined is the process of being turned towards correct beliefs about the good life and, therefore, about what one should and should not desire. This is certainly not metaphorical or playful. It neatly fits both the account and example of discipline that has been offered: the practice of the 'good and true' orator who, through persuasive speech, rids his audience

[29] *[H]oti hai men tôn epithumiôn plêroumenai beltiô poiousi ton anthrôpon tautas men apotelein, hai de cheirô, mê*. Zeyl translates: 'that a man should satisfy those of his appetites that, when they are filled up, make him better and not those that make him worse', taking *ton anthropon* to be the subject of *plêroumenai*. I take the subject of the sentence to be the true *arête* (503C5) that a good orator should possess. Socrates is claiming that orators like Themistocles were good only if Callicles is right to consider excellence to be 'to fill appetites, both one's own and those of others'. What he is saying in this line is that they are not in fact good, since the excellence of a good orator is in fact 'to fill those appetites that, once full, make a man better.…'

[30] For example, MacKenzie suggests that it is intended 'jokingly', (1981, p. 185).

[31] Cf 475D5–7: 'Don't shrink back from answering, Polus. You won't get hurt in any way. Submit yourself nobly to the argument, as you would to a doctor.'

of indiscipline and replaces it with temperance. That this is entirely serious is shown still more decisively by another comment Socrates makes, one that offers an important insight into his position in the *Gorgias*:

> I believe I am one of the few Athenians, not to say the only one, to undertake the true craft of politics and, in the present day, the *only* one to practise the political arts. This is because what speeches I make on each occasion do not aim at gratification but at what's best. (521D6–9, my emphasis)

The political arts, remember, are the two crafts of the soul, legislation and justice (464B4). Socrates, then, is himself the 'skilled and good' orator—or is at least the closest to this vocation[32]—and a clear implication of this is that if there is any orator or judge who can make a man better, through whatever form of discipline, Socrates is the only candidate. Not only, then, is Socrates serious when he says that he is disciplining Callicles, this is in fact the *only* form of discipline, of genuinely improving a person by making them more just, that is available in Socrates' time.

This brings disciplining a long way from flogging, but it is perfectly suited to Socrates' account of what injustice is. We saw that a recurring claim in the *Gorgias* is that having an unjust, indisciplined soul is 'the greatest evil' (*kakon ton megiston*). As Moss points out, precisely the same claim is made about having certain false beliefs:[33]

> For I count being refuted a greater good [than refuting], insofar as it is a greater good for oneself to be delivered from the greatest evil (*kakou tou megistou*) than to deliver someone else from it. I don't suppose there's anything quite so bad for a person as having false beliefs about the things we're discussing now. (458A5–B1)

This suggests an account of injustice that is familiar from the *Protagoras* and many other Socratic dialogues: being unjust and having false beliefs about what's good and bad are one and the same thing. Thus, Socrates makes the same claim about both refutation and punishment: if one has false beliefs, being refuted

[32] As Irwin notes: 'Socrates does not say that he *has* this craft but that he 'undertakes' (or 'attempts', *epicheirein*) it, looking for its principles; and so his remark need not conflict with his previous disavowal of knowledge' (1979, p. 240).

[33] Moss (2007a, pp. 233-234).

is a greater good than refuting; if one is unjust, being disciplined or punished is a greater good than disciplining or punishing. Similarly, Callicles is being refuted and this, Socrates claims, is 'the very thing the discussion is about, being disciplined'. Being refuted—being cured of false beliefs and so of injustice—is what Socrates means by just discipline or punishment. It is not surprising, then, that we find in the *Gorgias* a variant of another claim we saw in the *Protagoras*, the claim that no one errs willingly: 'no one wishes to do what's unjust, rather all who are unjust do so unwillingly' (509E5–7). Unwillingly, it seems, because they don't know that injustice harms and no one wishes to do what harms them. If they do injustice, then, they do so only because of their ignorance: unknowingly and, therefore, unwillingly. So to cure one of injustice one must be cured of the ignorance that causes the injustice. Given what we learned in the previous section about the cause of indiscipline, this is exactly what we should expect: one acquires indisciplined, insatiable appetites by believing, like Callicles, that all pleasures are good and should be desired indiscriminately, therefore one is cured of indiscipline by the sort of rational persuasion that can rid one of such false beliefs and replace them with true beliefs about what one should and should not desire.

Socrates does also say that an unjust, indisciplined man should be 'his own accuser' (508B5), which might indicate that there is also a form of self-discipline. This parallels what Socrates claimed earlier about punishment, that anyone who finds injustice in his soul should 'willingly go to where he'll pay his due as quickly as possible, to the judge just as he would to a doctor' (480A7–8). Given the account of what being disciplined is, being one's own accuser is to submit one's false beliefs to a form of correction—such as, for example, refutation, as Callicles' beliefs are refuted by Socrates—so that one will have true beliefs to guide one's desires. It is conceivable that this is something one can do without going to a judge or examiner such as Socrates. Socrates says that a person can not do just actions simply by wishing them to be just, rather one needs to 'learn and practise' a certain craft (509D7–E2). This is a craft that allows one to know on each

occasion what the right thing to do is. To the extent that one practises this craft, one is keeping one's desires disciplined. It seems plausible that this will include maintaining one's true beliefs in the face of temptations, that is, in the face of the false appearance presented by pleasures and pains, just as in the *Protagoras* one needs to employ the art of measurement to ensure one's beliefs are immune to the deception caused by immediate pleasure and pain.

Consider the account of courage as a kind of intellectual restraint or discipline in the *Republic*. Civic courage occurs when the city has 'the power to preserve through everything its belief about what things are to be feared … not abandoning [these beliefs] because of pains, pleasures, desires, or fears' (429B5–430B4); individual courage occurs when 'the spirited part of the soul … preserves through pains and pleasures the declarations of reason about what is and what isn't to be feared' (441C1–3). In the *Gorgias*, Socrates describes the temperate man as 'courageous' because he is such as to pursue or avoid 'what he should, whether these are practical matters, or people, or pleasures or pains, and to stand fast and endure what he should' (507B4–8).[34] One could interpret this as resisting acting on nonrational passions, but it is surely more appropriate to model our interpretation on Socrates' only explicit account of disciplining, a process that ensures that we have true appetite-guiding beliefs. The *Republic*'s description of courage shows that such an account is not unprecedented (the similarity, as we will see, is not a coincidence). Restraining or disciplining oneself, then, can be a matter of discovering and keeping correct beliefs even in difficult epistemic situations, where one is confronted with pleasures and pains that are liable to lead one's beliefs astray.

There are a number of *prima facie* reasons to doubt the intellectualist reading of the *Gorgias* I've offered. For example, Socrates explains to Callicles that be 'ruling oneself' (*heautou archonta*) he means 'nothing very subtle, just what the many mean: being temperate (*sôphrona*) and master of oneself (*enkratê*), ruling the pleasures and appetites within oneself' (491D1–E1). The appeal to the

[34] Compare what Socrates says about courage in the *Laches*: the courageous are 'not only those who are courageous against pains or fears, but also those who are clever at fighting against appetites and pleasures' (191D6–E1)

many and the claim that he is not saying anything subtle encourages the most straightforward reading of this claim, namely that he is saying that ruling oneself is holding back, habituating, or otherwise controlling non-rational appetites. This is the interpretation we would naturally give this line if it were spoken by the many of the *Protagoras* or even the Socrates of the *Republic*, and many similar lines can be found in the *Gorgias*.

This is not what we find, however, when we go beyond lines and examine the arguments Socrates gives, like the allegory of the leaky jar or his account of 'holding back' desires. This contrast between line and argument is, I believe, explained by Socrates' strategy. As with *akrasia* in the *Protagoras*, Socrates' aim in the *Gorgias* is not to spirit away common psychological phenomena, but to explain them in new ways. Socrates does not, implausibly, deny the existence of appetites or appetitive discipline and he rightly does not shy away from referring to them or engaging with his interlocutors on this shared ground. What he does do, however, is explain them in new ways and he uses arguments that draw his interlocutors into this new way of thinking about them. Thus, Socrates *does* believe, as his interlocutors do, that ruling oneself is 'ruling the pleasures and appetites within oneself', but he nonetheless differs in his account of what it means to rule one's pleasures and appetites.

There is another good reason to favour an intellectualist reading of the *Gorgias*. The alternative leaves us with nothing more than hints of a psychology of another dialogue. Within the *Gorgias* itself there is just not enough textual material to piece together a coherent psychological position that is anything like that of the *Republic*. The *Gorgias* does not have its own *Republic* book 4. Indeed, it does not present even one argument for any of the putatively non-intellectualist elements that readers have found, nor does it explain how these could do anything but undermine the psychological arguments that it does make. Consequently, to find a coherent non-intellectualist psychology in the *Gorgias*, one needs to use the *Republic* to fill in the many missing pieces. That would of course make it very easy to compare it with the *Republic*, but at the cost of making the comparison

meaningless.

There is, however, a substantial, meaningful comparison to be made between the psychologies of the *Gorgias* and *Republic* if the reading I have proposed is correct. There is no denying that intellectualism and the tripartite psychology of the *Republic* are at odds. In the previous chapter I said that intellectualism is reasonably summarised by the claim that there are no non-rational desires. In book 4 of the *Republic*, Plato argues that there are indeed non-rational desires. This is a substantial shift in Plato's thought—not a development, but a straightforward change of mind. However, there is more to both psychologies than their stance on the existence of non-rational desires. It is not, for example, simply the case that Plato decided in the *Republic* that the position of 'the many' in the *Protagoras* was right after all.

There is in fact significant continuity in the psychologies of the *Protagoras*, *Gorgias*, and *Republic*. While his view of the structure of the soul changes significantly, Plato's view of appetites and other lower passions remains, besides this, the same. What I aim to show in the next three chapters is that in the *Republic* Plato maintains, and develops further, the view of appearances we find in the *Protagoras* and *Gorgias*. Appetites, and other lower passions like anger or fear, arise through an unreflective acceptance of appearances. It is precisely because they are incapable of reasoning yet capable of apprehending appearances that the appetitive and spirited parts of the soul have their distinctive passions. This differs from the *Protagoras* and *Gorgias* in its view of the *subject* that judges by appearances—not the whole person, but a part of a person's soul—but not in its general account of how lower passions come about or, crucially, in its explanation of why such passions are often led astray. Thus, Plato's basic theory of practical error remains the same. Clearly this is an interpretation of the *Republic*'s psychology that requires defence. Providing this defence is the aim of the remaining chapters. I begin in the next chapter by examining an argument in *Republic* book 10, 602C–603A, that shows, I argue, that the non-rational parts of the soul have the cognitive abilities required to judge by appearances and are, due to their in-

ability to reason, confined to judging by appearances.

CHAPTER 3

Appearance & Non-Rational Belief: *Republic* 602C–603A

In book 10 of the *Republic* Plato offers an argument for the claim that the soul consists of parts. The argument's basic structure is the same as its better-known counterpart in book 4: Plato draws our attention to a case of opposition within the soul and appeals to the Principle of Opposites—that the same thing cannot at the same time do or undergo opposites in the same respect and in relation to the same thing—to conclude that different parts of the soul must be responsible for each side of the opposition. Beyond this similarity, however, the argument in book 10 is uniquely puzzling. The earlier argument begins with the widely accepted assumption that motivational conflict occurs, that is, that we can at the same time both desire to do and desire not to do the same thing. The argument in book 10, in contrast, centres on a far less palatable, purely cognitive sort of opposition: at the same time both believing and disbelieving the same thing.

It is generally thought that, unlike desire, reason abhors contradictions: we may unwittingly hold incompatible beliefs, but whenever we become aware of a conflict we are impelled, perhaps even compelled, to immediately resolve it (a Socratic elenchus is a good illustration of this compulsion to consistency). The argument of 602C–603A, however, relies on us accepting that in certain situations we knowingly hold simultaneous occurrent beliefs that contradict, seemingly with no avenue, or even wish, for resolution. We would hope to find a considerable attempt to make this plausible, but Socrates simply points to cer-

tain encounters with visual illusions as putative examples of such cognitive conflict: for example, believing a stick is straight but, at the same time, the opposite appearing to be the case because the stick is submerged in water. But is this really an example of conflicting beliefs? Surely, one might think, this is an opposition between what we believe and how things look, not an opposition between beliefs—surely we simply *don't* believe that the stick is bent, *despite* how it looks. Remarkably, Socrates shows no awareness of this concern; Glaucon accepts what he says without question.

Moreover, that Plato allows any cognitive conflict between parts of the soul is itself curious. Motivational conflict can occur between the appetitive, spirited, and rational parts of the soul because each have their own desires, but one might expect a cognitive ability like belief to be the preserve of just one part of the soul, the rational part, book 10 tells us otherwise: there are (at least) two believing parts of the soul. What is this second believing part? Unfortunately, in contrast to book 4, Socrates is not explicit about what parts he is dividing between. The rational part is named (*to logistikon*; 602E1) but the only other part mentioned is referred to vaguely as an 'inferior' or 'lower' (*phaulon*) part in us. This leaves space for two fundamentally different readings, each with evidence in their favour: some have argued that Plato is introducing a new division, subdividing the rational part of the soul into a higher and lower part; others maintain the earlier tripartition, arguing that the 'inferior' part is one or both of the non-rational parts we find in book 4.

In the following I offer solutions to these two problems—how visual illusions give rise to conflicting beliefs and what the two proposed believing parts are—with a reading of 602C–603A that pays particular attention to the use of the word 'appears' (*phainetai*) in the argument.

3.1 TWO BELIEVING PARTS OF THE SOUL

3.1.1 *The argument*

Showing that there are two believing parts of the soul is only a subsidiary aim of 602C–603A. It takes its place in a long and elaborate series of arguments defending the banishment of imitative poetry from the *kallipolis*, reinforcing the critique of poetry found earlier in book 3 with the help of the more sophisticated psychology that Plato has developed in the intervening books. Our passage's role in this series of arguments is to help identify the part of the soul on which imitation 'exerts its power' (602C5), revealing that it is an inferior part—'a part of us that is far from wisdom (*phronesis*)' (603A12–B1)—and thereby supporting the claim that imitative poetry's effect is corrupting. The argument does not consider poetry's effect on the soul directly, but rather makes a point about the effect of visual illusions, which are more closely associated with the imitative art of painting. But in the preceding discussion we learn that visual illusions and imitative poetry are connected: imitation's products are 'appearances' (*phainomena*; *phantasmata*) or 'images' (*eidola*) and these are the same kind of thing as reflections in mirrors, shadows, paintings, and visual illusions. The assumption, then, is that what is true of visual illusions—so long, at least, as it is true of them simply insofar as they are appearances—should also be true of imitative poetry.

The question that invites the argument is 'on which of a person's parts does it [sc. imitation] exert its power?' (602C4–5). Socrates begins his answer by drawing our attention to a variety of illusions:

> Through sight the same magnitude doesn't appear to us to be equal when near and far away … And something looks crooked when seen in water and straight when seen out of it and the same thing is seen to be both concave and convex on account of the eye's wandering anew around the colours (*dia tên peri ta chrômata au planên tês opseôs*). (602C7–12)

These are familiar and benign illusions. We typically see through them without difficulty: instead of trusting our senses we come to a correct belief by some more reliable means; Socrates mentions 'calculation, measurement, and weighing'. However, what is interesting about such illusions is their recalcitrance in the

face of our opposing belief: even if we know a submerged stick is straight, it will nonetheless *appear* bent. It is not hard to see why Plato thought that this tells us something interesting about the parts of the soul. If no amount of 'calculation, measurement, and weighing' will correct a visual illusion, this is a reason to think that it is not the calculating part of us, but some other part, that is responsible for the presence and persistence of the illusion. However, while this plausible thought may be its inspiration, the argument goes considerably further than this. The argument, as it is presented in the text, can be outlined as follows:

(1) 'Through sight the same magnitude doesn't appear to us to be equal when near and far away' (602C7–8) [This is the illusion most relevant to (2) and (4)].

(2) But 'measuring, counting, and weighing give us welcome assistance … so that we aren't ruled by something's appearing bigger, smaller, more numerous, or heavier, but by calculation, measurement, and weighing' (602D6–9).

(3) 'Calculating, measuring, and weighing are the work of the rational part of the soul' (602E1–2).

(4) 'But often to this [sc. the rational part], after it has measured and indicated that some things are larger or smaller or the same size as others, the opposite appears at the same time about the same things' (602E4–6).

(5) 'It is impossible for the same thing to believe opposites about the same thing at the same time' (602E8–9) [A variant of the Principle of Opposites (PO)].

(6) Therefore 'the part of the soul that forms a belief contrary to the measurements couldn't be the same part that believes in accord with them' (603A1–2).

(7) 'The part that trusts in measurement and calculation is the best part of the soul' (603A4–5).

(8) Therefore, 'the part that opposes it [sc. the best part] is one of the inferior parts in us' (603A7–8).

We can clearly see here the two problems I introduced above:

First, the argument as it stands has a conspicuous gap. Lines (1)–(4) establish that a certain belief-appearance conflict occurs: we believe a stick is straight but at the same time the opposite appears to be the case (where 'appears', occurring in (1) and (4), indicates that it *looks* to be the case, as I will argue). In (5) and (6), however, it is assumed that a belief-belief conflict has occurred: we believe a stick is straight and at the same time also believe that the stick is bent, in agreement with the appearance. What permits Socrates to move from belief-appearance conflict to belief-belief conflict? This is surely the most puzzling step in the argument and yet we find no attempt to justify it.

Second, the argument is not explicit about what parts it is dividing between: the part that trusts in calculation is described as the 'best' or 'calculating' part and the part which believes in opposition to this is described as an 'inferior' part of the soul. In (3) we learn that calculating, measuring, and weighing are the 'work' (*ergon*) of the rational part, as we would expect. It is possible, however, that the beliefs of the 'inferior' part are *also* the 'work' of the rational part—believing is, after all, an activity typically attributed to reason—and, therefore, that Plato is modifying his earlier tripartition, adding 'a new division, grounded on cases of cognitive conflict in which the reasoning part of the soul appears to be at variance with itself', as Burnyeat states his reading of the argument (1999, p. 223). Alternatively, if we can countenance the idea of non-rational belief, the argument of 602C–603A might simply be another route to the more familiar tripartite soul, in which case the 'inferior' part will match one or both of the non-rational parts established in book 4, the appetitive and/or spirited part.

I begin by addressing this second problem. I will argue that there are very strong reasons to think that the 'inferior' part is one or both of the non-rational parts. However, despite these reasons, we will see that a residual worry remains. Commentators on both sides of the debate have generally agreed that the text

presents one significant obstacle to taking the inferior part to be anything other than a sub-division of the rational part of the soul. Line (4), 602E4–6, seems to say—and one might think it says it quite explicitly—that the rational part alone is the subject of *both* sides of the conflict: it has the correct opinion through calculation and also has the opposite 'appear *to it*'. I will argue that until we have solved the first problem I raised—how Socrates moves from belief-appearance conflict to belief-belief conflict—we cannot assume that the part appeared *to* believes what appears. I will offer a solution to this problem through an account of what 'appearance' (*phainomenon*) means in 602C–603A: Socrates takes appearances to be themselves a (kind of) belief, allowing him to assume that the presence of a conflicting appearance entails the presence of a conflicting belief.

3.1.2 *What is the partition between?*

One source of resistance to the suggestion that the 'inferior' part is the appetitive or spirited part is the thought that belief is too cognitive an addition to parts of the soul that seem nothing more than seats of brute passions: surely it is the *rational* part's responsibility to perform such cognitive functions as thinking and believing, while the non-rational parts are responsible only for conative states like appetite, anger, or shame.

However, it is not at all clear that Plato takes our cognitive and conative functions to be divided so neatly. Our passage is not the first place where belief is attributed to a non-rational part of the soul. For example: moderation occurs when all three parts 'believe in common (*homodoxôsi*) that the rational part should rule' (442C11–D1);[1] in the soul's decline from oligarchic to democratic, it is not just the appetitive part's desires that take over the 'citadel' of his rational part but also false 'words and beliefs rush up and occupy this part of him' (560C2–3; see also 574D–E); in dreams a man's appetitive part can, while his rational part

[1] Although it requires greater exegesis, a similar point can be made about courage. A comparison of 429B8–D1 and 442B11–C3, where Plato describes civic and psychic courage respectively, suggests that psychic courage involves the spirited part preserving its correct beliefs in the face of temptations and fears.

is inactive, 'suppose' (*oietai*) that it is sleeping with his mother (571C9–D1); and the tyrannical man's decline is marked by being overcome by opinions that 'used only to be freed in sleep' (574D5–E2).[2] These passages are isolated and never receive any elaboration (unless, of course, this is what we get in 602C–603A), but they certainly upset the idea that the appetitive and spirited part are purely conative. One might dismiss them as misleading metaphors, but it is far from obvious that the psychology of the *Republic* requires us to. The soul in the *Republic* is not partitioned along Humean lines, a partition that aims to separate out the the soul's basic abilities: desire in this part, belief in that part, and so forth. Rather, what primarily distinguishes the parts of the soul are their distinct and potentially conflicting goals and, therefore, each part can be endowed with whatever cognitive or conative abilities allow it to effectively pursue its goal. Thus we find that all three parts of the soul have their own desires, including the rational part (e.g. 431B9–D2, 580C7–8). If all parts can have their own desires, what reason do we have to deny them, in principle, their own beliefs?[3]

One reason, it might be thought, is that a cognitive deficiency that Plato does attribute to the appetitive and spirited parts of the soul is, as our customary name for them suggests, that they are *alogiston*, non-rational. However, being *alogiston* is different from being incapable of belief. The most straightforward understanding of what it means for something to be *alogiston* is that it lacks the ability to engage in *logismos*: reasoning or calculation. It would surely be too strong a faith

[2] There are indications that the non-rational parts have other cognitive abilities too. The oligarch is described, in a seemingly critical tone, as holding his worse appetites in check 'not by persuading them that it's better not to act on them or taming them with arguments, but by compulsion and fear' (554D1–2; see also 586D–E), suggesting that the appetitive part is capable of, and best controlled by, responding to rational persuasion. The appetitive part is also often described as the 'money-loving' part since its 'appetites are most easily satisfied by means of money' (581E5–A1), which has led some commentators to think that the appetitive part is capable of rudimentary means-end reasoning, allowing it to recognise that money is a means to the satisfaction if its desires (e.g. Bobonich (2004, p. 244)). I agree with Lorenz (2006, p. 47 ff.) that the appetitive part's attachment to money is better understood as a habituated association between money and pleasure. For some discussion of the cognitive abilities this might involve, see pp. 174 ff. below.

[3] Burnyeat goes so far as to say that 'it is as mistaken to suppose the lower two parts of the soul incapable of thought or judgement as it is to deny desires and pleasures to the top part' (1976, p. 35, n. 22).

in the rationality of our beliefs to assume that all beliefs are the result of reasoning or calculating. Certainly there is nothing about beliefs *per se* that prevents them from arising not through reasoning but through some other, more causal means. As it happens, in 602C–603A we get clear evidence that Plato agrees: whatever interpretation of this passage one favours, there is no doubt that here he sets apart the ability to engage in calculation and the ability to form beliefs: the 'inferior' part is said to form the belief it does precisely *because* it lacks the ability to engage in calculation. There is, then, at least one part of the soul—whatever part it turns out to be—that is both incapable of *logismos* and capable of forming beliefs.

These observations should allay some of the worries about the very idea of attributing beliefs to the appetitive or spirited parts of the soul. I turn now to the reasons for thinking that this is exactly what Plato does in 602C–603A. I offer two arguments. The first argument centres on two claims, each of which enjoys convincing textual support: first, that poetry appeals to a non-rational part of the soul and, second, that painting—the form of imitation 602C–603A is concerned with—appeals to the same part of the soul as poetry.[4] The second argument looks at what 602C–603A is supposed to contribute to the broader argument of book 10. I argue that Socrates intends its conclusion to apply to imitation as such, with includes both painting and poetry, and not specifically to imitative painting.

The question 602C–603A aims to answer is, as we've seen, 'on which of a person's parts does [imitation] exert its power?' It is in fact the first of two arguments that answer this question. Since our argument reveals the part of the soul that visual illusions affect, Socrates takes it to illustrate most directly the part of the soul on which imitative *painting*, a visual imitation, exerts its power. However, what Socrates really wants to know is what part of the soul imitative *poetry* affects. While painting is visual, Socrates takes poetry, even in the dramatic form of tragedy, to be primarily auditory; its medium is 'words and phrases' that are given dramatic force through the 'musical colourings' of meter, rhythm, and har-

[4] Here I follow, and am indebted to, a number of authors who have offered a similar line of argument: Moss (2008), Lorenz (2006), and Singpurwalla (2011).

mony (601A4–B4). Thus he feels a second argument is needed, one that looks directly at the part of the soul affected by poetry (603C5–605C3).

This second discussion centres on the opposing tendencies of a person struggling with grief. On one side is a part of him that, following calculation and deliberation (604D4–5), bids him to tolerate his loss calmly and resist being overcome by grief. This side is again called 'the best' part of us (605B1) but it is now also referred to, apparently synonymously, as the *logistikon* (605B5), a word that has been used as a term of art for (the whole of) the rational part of the soul since book 4.[5] Pulling in the opposite direction is an 'uncalculating' (*alogiston*; 604D9) part of him that urges him to give in to his grief and that 'leads him towards recollections of his suffering and towards lamentation and is insatiable for these things' (604D8–9). This is the part affected by poetry. It is what leads us to enjoy, for example, a 'long lamenting speech' of a tragic hero recounting his suffering; it is a part of us that 'hungers for the satisfaction of weeping and wailing, because it desires (*epithumein*) such things by nature' (606A4–6) and, so, following a tragic tale gives it a vicarious taste of what it desires.

Now the emotion of grief and the 'hunger' for weeping and remembering suffering are new motivational states, at least in relation to the *Republic*'s psychology, but as strong, reason-resistant passions it seems highly likely that they find their home somewhere among the non-rational parts (even if it is hard to say exactly which of these parts they are most appropriate to). This is confirmed when Socrates turns from grief, which was introduced only as an example, to the broader impact poetry has on our desires and emotions. As well as low humour resisted by the rational part (606C2–9), poetic imitation is said to appeal to 'lusts and anger [or 'spirit'; *thumou*] and all that is appetitive (*epithumêtikôn*) and painful and pleasurable in the soul' (606D1–2) and to strengthen these elements in

[5] There are some especially clear examples at 439D5, 440E6, 550B1, and 553D1. Recall that in 602E1–2 Socrates says that calculating is the 'work' (*ergon*) of the *logistikon*. This suits Burnyeat etc. who need to say that the part that calculates is only a *sub-division* of the *logistikon* (and that the belief of the illusion-believing part is equally the *ergon* of the *logistikon*). So when Socrates implies here that the 'best' part *is* the logistikon, this suggests—assuming we can't have better and worse best parts—that the calculating part was simply the *logistikon*, not a sub-division of it.

a way that 'sets them up to rule in us when they should be ruled' (606D5). These are indisputably passions of the appetitive and spirited parts of the soul and these parts are, moreover, the only parts we would expect to be capable of challenging the rational part's rule.[6] Unlike the argument from visual illusions, then, there is little doubt about the parts of the soul poetry appeals to.

This gets us half way. If it is to help us identify the 'inferior' part in 602C–603A, we need a second claim: that poetry and painting appeal to the same part of the soul. This is generally the more disputed of the two claims. While few would wish to suggest that a part that is 'hungry for the satisfaction of weeping' is a division of our rational part, it is nonetheless possible that Socrates is employing two different divisions in book 10, apparent in two distinct situations: a division within the rational part revealed by cognitively deceptive visual appearances and a division between the rational and non-rational parts revealed by emotionally engaging poems. On this view, insofar as it relates to poetry, 602C–603A presents nothing stronger than a useful parallel, adding plausibility to the idea that, like painting, poetry's affect is local to *some*—though not one and the same—part of the soul.[7] There are, however, strong reasons to doubt this reading.

The first and clearest is a textual reason. Once Socrates has established that poetry appeals to our lamenting part, he concludes that, therefore, 'we'd be right to take him [sc. the poet] and put him beside the painter as his counterpart' (605A8–9). A few lines later he explains that they are counterparts because they appeal to the same part of the soul:

> [T]he imitative poet … by making images (*eidôla eidôlopoiounta*) far removed from the truth, gratifies the part of the soul that is thoughtless and doesn't dis-

[6] Socrates does not clearly distinguish here between the appetitive and the spirited part. I am inclined to agree with Nehamas that, rather than creating a simpler non-rational part amalgamated from appetite and spirit, 'the explanation of why he opposes reason to spirit and appetite together … is simply that he does not need to distinguish these two for his present purposes' (1999, p. 267). In particular, distinguishing them is unimportant if both parts are corrupted by poetry for precisely the same reasons, being equally sensitive to imitations. This is not to say they will respond equally to the same pieces of poetry: we might expect our spirited part, for example, to be corrupted by false accounts of what is honourable and the appetitive part to be more susceptible to false accounts of what is desirable.

[7] See Burnyeat (1999) and Nehamas (1999).

> tinguish greater from lesser, but believes the same things are at one time large and another time small. (605B7–C4)

This is undoubtedly a reference to the visual illusions used as the central example in 602C–603A: 'the same magnitude doesn't appear to us to be equal when near and far away' (602C7–8). What Socrates tells us, then, is that poet and painter gratify the very same part of the soul: both gratify the part that believes in visual illusions that we encounter in 602C–603A.

This leaves little room for doubting that the 'inferior' part in 602C–603A is one or both of the non-rational parts of our soul. We might still wonder, however, why painting and poetry affect the same part of the soul. Those who argue that painting and poetry affect different parts of the soul naturally argue that they do so by virtue of different characteristics: e.g. painting because it is cognitively deceptive and poetry because it is emotionally engaging.[8] Conversely, if they both affect the same part of the soul we should expect painting and poetry to do so by virtue of some characteristic they share in common. The answer, I believe, is that visual illusions, paintings, and poetry are all the same kind of thing, appearances, and so each equally appeals to a part of the soul sensitive to appearances.

Consider again the passage quoted above. Socrates says the poet affects the illusion-believing part of the soul '*by* making images far removed from the truth'. A maker of images far removed from the truth is Socrates' definition of imitation (*hôrisametha*; 599D2–4). An imitator is a craftsman who makes 'images' (*eidola*) or 'appearances' (*phainomena*) of what's real, 'at a third remove from the truth': e.g. a painter makes a kind of 'couch', an image-couch, that imitates the carpenter's couch, which is itself modelled on the 'true' couch, the Form of a couch.[9] The poet, then, affects the non-rational parts of the soul by making imitations. I believe we should take this to be sufficient. That is, a poem affects the inferior part of the soul not in virtue of something unique to a poem—say, its

[8] Nehamas (1999) takes the difference to be between the cognitively deceptive and the emotionally engaging. Burnyeat seems at times to suggest that there is also a significance in the fact that one is visual and the other auditory (1999, see 225 and 320).

[9] I look that this hierarchy in detail in the next chapter.

emotive content or the fact that it's auditory—but simply in virtue of being an imitation, an image or appearance of what's real.[10]

The same would seems to be true of painting. Painting is introduced in book 10 to serve as an especially clear example of imitation (a painting is, in a fairly obvious way, an image that imitates a subject). Thus, while poetry is what Socrates is ultimately interested in, he relies on the example of painting in his account of 'what imitation in general is' (595C7) and why it is something epistemically inferior (595C–598D). Since these arguments must apply to poetry too, the underlying assumption is that paintings and poems are, insofar as they are imitations, equivalent: while a painting of a couch and a Homeric epic are different in many ways, insofar as they are mere appearances of what's real they are the same kind of thing and so can be studied from the perspective of their shared characteristics. So in the same spirit, when Socrates asks what part imitation exerts its power on, he turns first to painting as his clearest example of imitation. If this is right—if painting is considered simply as an example of imitation—then the argument considers visual appearances not insofar as they are *visual*, or any other way unique, but insofar as they are *appearances*, the kind of thing an imitator makes. On this reading 602C–603A aims at a conclusion that applies to both painting and poetry: since painting *qua* imitation affects *x* part of the soul, then imitation *per se*, including poetry, affects *x* part of the soul. Likewise for the subsequent argument from poetic imitations: both isolate a part of the soul affected by imitation as such, a part that is insensitive to calculation yet sensitive to appearances. Only on this reading can we explain why Socrates is able to apply the argument from painting to both 'painting and *imitation as a whole* (*holôs hê mimêtikê*)' (603A10; my emphasis) and to draw the similarly comprehensive conclusion that 'imitation is an inferior thing consorting with an inferior thing to produce an inferior thing' (*phaulê ara phaulô suggignomenê phaula genna hê*

[10] Of course this is not to say that a poem's emotive content doesn't affect our lower parts. It does, but it does so *because* the emotive content is the content of an imitative appearance and appearances are the medium through which poetry communicates with soul.

mimêtikê; 603B4).[11]

There are, then, two good reasons for taking the inferior part of the soul in 602C–603A to be one or both of the non-rational parts: first, the poet undoubtedly appeals to this part of the soul and Socrates clearly tells us that the part the poet appeals to is the illusion-believing part and, second, it allows for a plausible, unified explanation of the psychological effect of both painting and poetry that is grounded in their most relevant characteristic, the fact that they are imitative arts. However, I have so far ignored the text that is often thought to be the strongest evidence in favour of the contrary view, that the division is internal to the rational part of the soul. In the next section, I aim to show that what this text tells us cannot in fact be decided until we solve the first problem I raised, the argument's undefended move from belief-appearances conflict to belief-belief conflict.

3.2 BELIEF-APPEARANCE CONFLICT: 602E4–9

Those who argue that the division is within the rational part of the soul take their cue from line (4), 602E4–9. On its standard and the most natural translation this line states that the rational soul suffers by itself the opposition between our belief about what is the case and what appears to be the case. To quote the line again:

[11] It is true that he goes on to express reservations about the generality of the argument, and so recommends a second argument that looks directly at poetry:

> S: Does it apply only to the imitations we see, or does it also apply to the ones we hear—the ones we call poetry?
> G: It is likely (*eikos*) it applies to poetry too.
> S: Then we must not rely only on a likeliness (*eikoti*) drawn from painting, but also go directly to the part of our thought that poetic imitations consort with [...] (603B6–C1)

How 'likely' it is surely depends on nothing other than the likeliness of the general conclusion he reaches, which is that the argument tells us what part *all* imitations appeal to. So the worry he expresses is about the *certainty* of this conclusion, which obviously does not imply that the conclusion he has reached is, while held with reservations, anything other than what he says it is. Notice also the sentences not-only-but-also (*mê … monon … alla kai*) structure: 602C–603A does support the general conclusion about all imitation, but for *additional* support they must look directly at poetry.

(4) But often to this [sc. the rational part, *to logistikon*; 602E1], after it has measured and indicated that some things are larger or smaller or the same size as others, the opposite appears at the same time about the same things (*toutô de pollakis metrêsanti kai sêmainonti meizô atta einai ê elattô hetera heterôn ê isa tanantia phainetai hama peri tauta*; 602E4–6).

There has been general agreement that *if* the rational soul suffers this belief-appearance conflict then it must also suffer the belief-belief conflict and, therefore, applying the Principle of Opposites, it must be divided into two parts, yielding a tetrapartite soul consisting of two rational parts and two non-rational parts.

Naturally, then, those wishing to avoid this conclusion have offered alternatives to the standard translation of 602E4–9, translations that allow the rational part of the soul to be the subject of the belief but not of the opposing appearance. Many alternative translations have been attempted. Some try to find an alternative subject for the *toutô*. Barney (1992, 286, n. 4), for example, suggests that *toutô* refers to the rational part's *conclusion* and should be governed by *tanantia*, giving us 'often the opposite of this [i.e. of what the rational part concludes]—when it [the rational part] has measured and indicates that some things are greater or less than or equal to others—appears at the same time, about the same things'. Others have suggested that the *toutô* refers to whole soul, not just the *logistikon*. Both of these alternative translations fall foul of the fact that by far the most likely subject of *toutô* is the *logistikon*, as is very clear from the exchange leading up to 602E4–6: 'S: But this [calculating, measuring, and weighing] is the work of the rational part of the soul (*touto ge tou logistikou ê an ein tou en psuchê ergon*) G: Yes, of this [part of the soul] (*toutou yar oun*) S: But often to this …' It is surely most natural to take the *toutô* to have the same reference as the *toutou* in the previous line. A quite different approach, found first in Adam (1902, App. II) and more recently in Lorenz (2006, pp. 67-68), is to take the *tanantia* not to refer to the opposite of what the rational part of the soul concludes, but to signal that *its* conclusion is the opposite of the false appearance. On this reading, however, the *phainetai*

would need to refer not—as it does in all other uses in the passage—to the false appearances but to the view the rational soul takes of the matter (I discuss the passage's use of *phainesthai* further below).

Both sides of the debate, then, take 602E4–9 to be decisive, believing either that it shows that it is the rational part that is partitioned or that it *would* show this on its standard translation, and so requires a less obvious translation. It is only decisive, however, if their common assumption is correct: if the rational part suffers the belief-appearance conflict then it must also suffer the belief-belief conflict. This assumption has not received much defence and yet I think there is little doubt that, at least where its validity is concerned, the argument's transition from belief-appearance conflict to belief-belief conflict is its most puzzling step. Why does Plato think that having something appear to one entails that one believes what appears? Until we have answered this question, we are not entitled to assume that the part of the soul that is appeared to—the rational part of the soul on the standard translation of 602E4–9—is the same part that believes what appears.

So, to understand the implications of the standard translation of 602E4–9, which strikes me as certainly the most likely translation, we must first try to understand this step in the argument. That is, we need to understand how it is valid for Socrates to claim in (4) that belief-appearances conflict often occurs and then in the continuation of the argument to assume that he has shown that belief-belief conflicts often occur:

(5) 'It is impossible for the same thing to believe opposites about the same thing at the same time' (602E8–9).

(6) Therefore 'the part of the soul that forms a belief contrary to the measurements couldn't be the same part that believes in accord with them' (603A1–2).

What is assumed, it seems, is that the 'appears to' in (4) implies the presence of a concurring belief. Accordingly, the simplest way to render the argument valid is to take the following premise to be implicit between (4) and (5):

P If it appears to one that *p*, then one believes that *p*.

Prima facie, **P** is highly questionable. Most of us will readily accept that what appears to be the case can be—in at least some sense of 'appears'—the opposite of what we believe to be the case, as is illustrated well by the familiar false appearances Socrates lists and well-canvassed modern examples like the Müller-Lyer lines. But there are two fairly strong reasons for not positing an opposing belief in such situations: first, that it seems to be not only irrational but psychologically impossible to simultaneously hold occurrent beliefs that contradict (cf. *De Anima* 428B5–9) and, second, that in such cases one does not seem to be aware of having an additional belief that assents to the appearance. Thus, a quite natural conclusion to draw from such cases is that having a belief and having something appear to one are different and independent mental states. In particular, as Aristotle saw, such cases can be taken to be *counter-examples* to **P**, on which Socrates' argument seems to turn.[12]

It should surprise us, then, that there is no explicit attempt to defend the argument's move from belief-appearance to belief-belief conflict. Indeed, even the occurrence of belief-belief conflict is not explicitly stated. Rather it is simply assumed in (5), the statement of the Principle of Opposites. The fact that Socrates is so blasé about the move from appearance to belief certainly makes interpreting the argument difficult, but it also gives us a clue about what sort of assumption is at play. It suggests that Plato thought there was something plausible, perhaps even obvious, about taking belief-appearance conflict as a form of belief-belief conflict, so much so that it doesn't require mention. If this is right, it places a constraint on acceptable accounts of what background assumptions allow Plato to take **P** for granted in the argument: what we are looking for is not an elaborate philosophical defence of **P** but some intuitive, relatively theory-light reasons

[12] Aristotle, in the *De Anima*, uses this line of argument to object to the view—he is likely to have Plato's position in the *Sophist* in mind—that appearances are a 'blend' (*sumplokê*) of perception and belief. Such a view would require that we believe that *p* when it appears that *p*, but he points out that 'often false things appear while at the same time we hold true beliefs about them, e.g. the sun appears a foot across but is believed to be larger than the inhabited part of the earth' (3.3.428B).

why someone could take it to be true.

There are two places where one might look for Plato's reasons for believing **P**: in his psychology, namely some feature of it that implies invariable assent to appearances, or in something internal to his conception of what an 'appearance' is, something which implies a conceptual link between having it 'appear' to one that *p* and having a belief that *p*.

Hendrik Lorenz (2006, p. 68) is an example of someone who takes the first approach. Lorenz recognises that **P** is just assumed in 602C–603A and that this requires justification, but he says 'a moment's reflection on Plato's psychological theory should make it clear how natural it is to assume that the parts of us below reason accept sensory appearances'. The reflections he has in mind are the following:

> [The lower parts of the soul] could never begin to perform [their characteristic] functions without being supplied with tolerably good information about the person's environment … The text before us suggests that, just as one would expect, one way they get the information they need is by sensory appearances. Moreover, the lower parts of the soul cannot do what we do, namely resort to measurement, arithmetic, and the like, so as to discover how things really are. For these are the resources of reason. Unlike us, then, the lower parts are at the mercy of how things appear to the senses (cf. 602 D 6-9). They cannot help being taken in by sensory appearances. (p. 68)

While this is a plausible story, much of which I would agree with, Lorenz is mistaken to identify this as the assumption the argument requires. Lorenz's suggestion is that the belief-belief conflict is established by bringing into play three claims about the cognitive abilities of the non-rational parts of the soul: they have access to appearances; they lack the reasoning abilities needed to doubt appearances; they have the ability to, and are constrained to, be taken in by appearances (e.g. to form desires on the basis of these appearances). Together these claims make it natural to assume that a non-rational soul will be taken in by the appearance that the rational part is said to reject in (4). The problem is that this would make the argument circular: it makes a premise of the argument rest on assuming that there are two different believing parts of the soul when the argument is intended to *establish* such a partition between believing parts. We should not for-

get that even though Socrates has already argued for the division of the soul, what he offers in book 10 is a new argument for this claim, starting from scratch. So if the argument is to be valid, the possibility of belief-belief conflict must be made plausible without taking for granted either partition itself or the existence of two believing parts of the soul. This precludes any appeal to Plato's tripartite psychology, especially to features of the as-yet-unestablished non-rational/believing part of the soul. Once the argument is concluded, of course, it is possible to match up the illusion-believing part to its corresponding part in book 4, but insofar as both are equally arguments for partition they must remain independent.

The argument's deductive direction, then, must be from belief-appearance conflict, to belief-belief conflict, and finally to the existence of separate believing parts of the soul. An attempt to explain the former in terms of the latter will lead to circularity. With appeal to Plato's broader psychology precluded, we should turn to the second option: that Plato's understanding of 'appearance' entails, in and of itself, that we believe what appears.

3.2.1 *What is an appearance?*

The question of what Plato means when he says something 'appears' (*phainetai*) to the soul should be broken into two smaller questions. First, there is a linguistic question. 'Appears' is an ambiguous word, with at least two senses, so our first question is which sense of 'appears' is Plato using. This is not a question about what theory of appearances Plato might have, but simply about which of the common Greek senses of *phainesthai* he employs. Second, once we have settled on a sense of 'appears', we can then consider what, if any, conception or theory of such appearances Plato held.

As we saw in chapter 1, the first question arises because the meaning of *phainesthai*, like the English 'appear', ranges from a judgemental sense, which indicates little more than that someone holds a certain belief, to a phenomenal or sensory

sense, indicating something that is presented to us, often in perception.[13]

The judgemental use of *phainesthai* is by far the most commonly found use in Plato. In this sense, saying something 'appears' to one to be so-and-so simply indicates that one believes it to be so-and-so, although often provisionally or tentatively (e.g. 412E10–11; 470B4–6; 559D5). A fairly ubiquitous example is when one of Socrates' interlocutors registers his assent with the familiar 'it appears so' (*phainetai*). In this sense, the use of *phainesthai* involves no specific commitments about what evidence supports one's judgement, perceptual or otherwise—one is simply stating one's view of a matter.

The phenomenal sense of *phainesthai* is very different. In this sense, saying something 'appears' so-and-so to one is to say it is presented to one as so-and-so, irrespective of the opinion one actually has of it. The most common and clearest example of the phenomenal use of *phainesthai* is its use to describe how something is presented to the senses—how something looks—and for present purposes it is safe to take phenomenal appearances to be a type of perceptual experience (thus, I will at times refer to them as sensory appearances). Thus, if we say that 'the moon appears larger when lower in the sky' we are saying that this is how it looks to us, without implying that this is how we believe it to be: the moon will appear larger—will be presented to us as or will look larger—irrespective of what we believe. Aristotle took great pains to carve out this sense of *phainesthai* and distinguish it from believing, telling us for example that:

> Often, false things appear (*phainetai*) while at the same time we hold true beliefs about these things, e.g. the sun appears a foot across, but is believed to be larger than the inhabited part of the earth. (DA 3.3. 428B2-4; see also EE 1235B25-29)

As these examples illustrate, in contrast to seeing that *p*, saying that it appears that *p* does not imply that *p* is true. This makes the phenomenal meaning of 'appears' particularly apt in cases where perception is suspected of failing to represent how things really are, perceptual illusions being a prime example.

[13] In a third use, *phainesthai* with a participle has a veridical meaning: to be manifest or evident. Naturally, this is not a sense we find in the passages I discuss here.

While they are related, the judgemental and phenomenal senses of *phainesthai* are different and unmistakably so. They pick out what seem to be very different mental states: beliefs and, for example, certain perceptual experiences. Thus it can phenomenally 'appear to' one that *p* and also judgementally 'appear to' one that not-*p*. I say *seem* different because, as we will see, I will argue that phenomenal appearances are in fact a type of belief—not according to the *sense* of the word, but according to Plato's *theory* of what they are—but a non-standard sort of belief unconnected to the judgemental sense of 'appears' (and so making it no less easy to distinguish the judgemental and phenomenal appearances).

Which of these senses of 'appears' is Socrates using in *Republic* book 10 and specifically in our passage, 602C–603A? There is an obvious advantage to taking Plato to be using the judgemental 'appears': this would give us exactly the sort of relation between appearance and belief we are looking for, rendering 602C–603A valid. On the other hand, while phenomenal appearances do seem, *prima facie*, to be independent of our beliefs they also seem a better description of the sorts of appearances, perceptual illusions, that Socrates is concerned with.

Rachel Barney takes book 10's sense of 'appears' to be judgemental, claiming that the apparent 'is in general what I, or some constituent part of me, initially and unreflectively takes to be the case on the basis of perception' (1992, p. 287). If Barney is right about Plato's use of 'appears' in book 10, then the move from belief-appearance to belief-belief conflict is simple and valid, indeed no more than a rewording. To say that often when the rational part calculates or measures to a certain judgement, the opposite appears to be the case is, on Barney's reading, just a way of saying that the opposite is—'initially and unreflectively' and 'on the basis of perception'—believed to be the case. This way of stating this interpretation, however, reveals its weakness. It allows for a smooth transition from premise (4) to (5) at the cost of passing onto (4) precisely the same difficulty we had with (5): if 'appears' indicates a certain sort of belief, then (4) already assumes belief-belief conflict occurs and, again, we need to ask why. This also fails to give a clear place in the argument for the undoubtedly phenomenal appearances that

are the argument's intuitive starting point, the fact that sometimes something can look (i.e. sensibly appear) one way and be believed to be another. Thus, even if we take a judgemental reading of 'appears', the original question remains, if requiring a restatement: why does Plato think that when we judge a (phenomenal) appearance, like a perceptual illusion, to be false, the opposite at the same time (judgementally) appears to us to be true?

While this explanatory failure is a reason in itself, book 10 furnishes us with plenty of additional evidence to allay any doubt about the sense of 'appears' Plato is using. Imitators are described as craftsmen of *phainomena* but also, and seemingly synonymously, of *phantasmata* (semblances) and *eidola* (images), words not easily given a judgemental construal. Furthermore, the examples Socrates gives of what he calls 'appearances' are predominately typical perceptual appearances: paintings, reflections in mirrors, and perceptual illusions.

To take a specific example, consider one of the opening lines of the argument: 'Through sight (*dia tês opseôs*) the same thing appears (*phainetai*) to us not to be of equal size when near and far away. – No indeed.' (602C7–9). The reference to sight makes it natural to take this to be a claim about sensory appearances, about how something's size is presented to sight, how it looks when near and far. Such a reading also makes the claim comprehensible and true, at least to the extent that the same object will take up a larger space in our visual field when near than when far away. On a judgemental reading, on the other hand, what Socrates says, insofar as it is his *opening* claim, is absurd. The claim is not a qualified one, it is not about what sometimes appears to us. So on a judgemental reading Socrates is claiming that we generally believe the same thing is a different size when near and far away (with the 'through sight' presumably adding that this belief is based on perceptual evidence). This is not something we would expect Glaucon to agree to so readily; *prima facie*, at least, it seems false.

Of course while it seems *prima facie* to be false, as the argument develops it becomes clear that it is, in some way, something Socrates believes. The 'in some way', however, is important. To understand this we need to turn from the first,

linguistic question, to the second question: if 'appears' has a phenomenal sense in 602C–603A, what conception of phenomenal appearances allows Plato to assume **P**?

Before answering this question, consider first an implication of taking 'appears' in a phenomenal sense. If the rational part of the soul is itself sub-divided into a lower, uncalculating rational part and a higher, calculating rational part, the suggestion is that it 'appears to' the lower rational part that the submerged stick is bent and, at the same time, the higher rational part believes that the submerged stick is straight. Now, on the phenomenal reading to say that something 'appear' so-and-so to a part of the soul, is to say that it is presented to that part of the soul as so-and-so, without—at least according to the meaning of the word—necessarily entailing that that part believes it is how it appears. It can, in other words, be used in a similar way to 'looks': if I say the stick looks bent to me, I am not committing myself to the belief that it is bent. Rather, I am simply stating that I am aware of how the stick is presented to my senses. On this understanding, would we want to say that, whatever appears to the lower rational part, the stick doesn't also appear bent to the higher, calculating part? That it is not even *aware* of how the stick looks?

On reflection, it would in fact be very strange, on any reading of the passage, if the calculating part were *not* appeared to. Knowing how the visual illusion is presented to the senses is surely a prerequisite for questioning it. The rational part employs calculation to reach a correct judgement about the illusion, but it must start by figuring out what is awry with the way things appear. For example, what strikes me as the most plausible candidate for the 'measuring' used to conclude that contrary to appearances a submerged stick is straight is that while it appears bent to it, the rational part of the soul knows enough about optics to realise that this is exactly how one would expect a straight but submerged stick to appear. In such a scenario, the rational part is not blind to the false appearance but actually requires the rectifiable false information imparted by the appearance to know that there is straight stick there.

This might seem to put so much distance between being 'appeared to' and believing that there is very little space left for **P**. I will argue that the seeming absurdity of claiming that we believe, in an unqualified way, near things are larger than far away things is avoided by using the phenomenal 'appears' but nonetheless taking the appearance itself to be a belief or belief-like state. That is, being cognisant of the appearance does not entail belief, even though the representational mental state that is the appearance is itself a kind of belief.[14]

3.2.2 *Appearance as belief*

My suggestion, then, is that Plato takes phenomenal appearances to be a type of belief. An initial problem with this suggestion is that phenomenal appearances were characterised as independent from belief: it will appear to us that *p* even if we believe that not-*p*. I suggest we modify the claim as follows: phenomenal appearances are independent of *rational* beliefs, being themselves beliefs in a non-standard way, namely a form of non-rational belief.

By 'rational' belief I mean simply what most people would call a belief. Rational beliefs tend towards mutual implication, abhor contradiction, and are typically responsive to evidence and reasoning. So if we employ some reasoning, such as 'measuring, counting, and weighing', to conclude that *p*, we will, other things being equal, form a rational belief that *p*. Moreover, it seems impossible to knowingly holding simultaneous occurrent rational beliefs that contradict. Clearly Plato's conception of belief goes beyond this: in our passage he allows contradicting occurrent beliefs and beliefs that are unresponsive to reasoning. It also seems safe to assume that the beliefs Plato would take one to, in normal circumstances, report as one's own are rational beliefs: a person who reasons to the

[14] I use the cautious 'belief-like' in recognition of the fact that someone might agree with the understanding of appearances I attribute to Plato in the following and yet be unwilling to call them beliefs. Subsequent to writing this chapter, I have found that Todd Ganson, in an excellent discussion of 602c–603a, similarly argues that we should identify the sensory appearance itself with the opposing belief (2009, pp. 185–187). Ganson appeals to the plausible similarities between sensory appearances and beliefs and the fact that it makes the beliefs appropriately non-rational. My discussion differs in focusing on the textual support for this claim and on the implications it has for our understanding of the structure and validity of the argument.

conclusion that a visual illusion is false would surely report this as their belief, not the opposite belief associated with the appearance.

I will argue, then, that Plato takes this non-standard, non-rational belief to be the appearances itself. I will appeal first to evidence from within the *Republic* and then to the quite different, but nonetheless enlightening, evidence from the *Sophist* and *Theaetetus*.

Evidence from the Republic

Since what we are dealing with in 602C–603A are sensory appearances—how certain things look to us—we should expect to learn something about them from Plato's view of perception. In the *Republic*, the place to find such a view is 523A–525A, the so-called 'finger' passage. Plato here draws a distinction between perceptions that summon understanding and perceptions that do not.

> Some perceptions don't summon the understanding (*tên voêsin*) to investigate them, because the judgements of perception (*tês aisthêseôs krinomena*) are themselves sufficient, while others encourage it in every way to look into them, because perception seems to produce no sound result. (523A10–B4)

Perceptions that summon understanding (or 'calculation', *logismos*, 524B4, or 'thought', *dianoia*, 524D3) occur 'whenever sense perception doesn't declare one thing any more than another' (523C2–3). Socrates' example is perceptions of magnitude: sight declares one's ring finger to be both large (in relation to one's little finger) and small (in relation to one's middle finger). Such perceptions are said to 'compel' us to summon the understanding: if perception doesn't declare something to be *F* any more than not-*F*, understanding must be used to know what *F* is. In contrast, if what we look at is simply a finger, not its relational properties, sight is sufficient on its own: 'an ordinary soul isn't compelled to ask the understanding what a finger is, since sight doesn't indicate to it that a finger is at the same time the opposite of a finger' (523D3–6).

In normal circumstances, then, perception is said here to be capable of issuing sufficient judgements to the soul—presumably sufficient to, for example, act on them—without calling on understanding. One might object that 'judgements' is

an over-translation of *krinomena*, which could be less suggestively translated as 'discriminations', but this would not be in keeping with the similarly judgemental language we find throughout the passage: perception is said to give 'reports to the soul' (*paraggellei tê psuchê*; 524A3); 'indicate' or 'declare' (*sêmainei*) things to be a certain way (524A7; 534A10); give 'descriptions' or 'interpretations' (*herêmeiai*) (524B1); and make conflicting 'announcements' (*eisaggellomenên*; 524B5) that the understanding must unravel. Throughout the passage, then, perception is presented as a faculty that *tells* the soul what is being perceived, for example that this is a finger.

This way of talking about perception is not mere metaphor. It is importantly germane to the distinction the passage draws. The passage contrasts perceptions that require work on the part of understanding with those that 'do not awaken understanding' (524D5), those cases where 'the judgements of perception are themselves sufficient (*hikanôs*)' (523B1–2). Plato, then, characterises perception as, in normal circumstances, sufficient on its own, that is, capable of giving us information about the world without requiring help from the understanding, calculation, or thought. Plato's assertoric account of perception allows us to understand this: in normal circumstances, perception presents the soul with a fully-formed judgement or report about the relevant feature of the world, so there is no extra work for reason to do—it is not necessary, for example, to *infer* external objects from our sensory experience. (Contrast *Theaetetus* 184B-186E, discussed below, where the labour of seeing external objects is divided between perception, which only gives us awareness of proper sensibles—colours, sounds, and the like—and the mind's work 'itself by itself', which applies concepts to the information the senses gather.)[15]

[15] Thus, I am largely in agreement with Burnyeat: '[F]or Plato the misleading appearance already involves judgement. Throughout the *Republic* perception is treated as a judgment-maker independent of reason, but much less reflective. A modern Idealist philosopher wrote: " 'Sense-perception' is a form of 'knowledge,' a 'cognizant experience,' in which the mind thinks sensuously. There is 'thought' in sense-perception, but not thought free and explicit - not 'thought' which the percipient controls, or of which he is even aware as 'thought'." With much of this Plato could agree. Especially the last clause.' ((1999, p. 288); see also Fine (1988)). While this strikes me as correct, Burnyeat nonetheless believes that the 'appeared to' in 602E4–9 implies that it is

This is in fact quite a plausible view of perception. The idea that perception announces or asserts something to us captures an important fact about perceptual experiences. This is the fact that the content of perception does not seem to be simply laid before us to be coolly entertained. If I ask you to simply consider the proposition 'the flowers in the vase are red', you have no inclination to take this to be true or false. All you can be said to do is to entertain it or imagine it as true. As such, a proposition like this in no way purports to represent the world as being a particular way—it could just as easily be a sentence from a work of fiction (or example in a philosophy paper) as a description of something real. Perceptual experiences, however, are not similarly neutral; they do purport to represent the world as being a certain way. Indeed, that is their function. To do this they must present their content with a certain attitude towards it, presenting it as true and thus inviting us to—and giving us a reason to—take it as true. This makes it possible for perception to be representational, to be *about* something in the world, in a way that content that is simply entertained cannot be.[16] I find the idea that perception makes assertions about the world both compelling in itself—I will give further reasons below—and appropriate to Plato's descriptions in the finger passage: perceptual experiences can be seen to report or announce that this is the way the world is, to assert their content, leaving us to accept or deny their testimony.

On this assertoric reading, then, when we see that *p*, perception asserts to the soul that *p*, where this means simply that it presents *p* to the soul as a fact it has discovered about the world, as something true. This alone might make a sensory appearance sufficient to oppose the calculating part of the soul's conclusions in 602A–603C (insofar as it implies that perception presents a proposition, such as 'this stick is bent', as something true when the calculating part believes it to be

---

the rational part that believes what appears. Contrast Burnyeat's earlier view, (1976, pp. 34-36), which is closer to what I argue here.

[16] This has been widely recognised in modern discussions of perception: Heck talks of perceptual experiences having, like beliefs, 'assertoric force' (2000, p. 508); Martin of how perceptual experiences 'directly coerce our beliefs about the world' (2002, p. 391); and Gluer of perceptual experiences making 'claims about how the world actually is' (2005, p. 6).

false) but it is also not implausible that it is sufficient to make it a *bona fide* belief. Certainly on Plato's own account of belief, unfortunately not given in the *Republic* itself, the principal characteristic of a belief is that it is something we assert to ourselves.[17] Belief in the *Sophist* is defined as an assertion (*phasis*) or denial of a proposition (*logos*) in thought and, in turn, thought is defined in comparison with speech as the soul's silent 'conversation with itself' (263E–264A).[18] Similarly, in the *Theaetetus* a belief is 'a proposition asserted not aloud to someone else, but silently to oneself' (190A). If Plato thought something similar about belief in the *Republic*, his description of perception in the finger passage would be an unmistakable example of belief.

Consider also the implications this has for 602E4–9. When perception at the same time 'reports to the soul' that the middle finger is both large and small (since it is perceived alongside the little finger and ring finger) we should take this to mean that the middle finger *appears to* the soul to be both large and small.[19] There is surely no other way in which the soul can become aware of the content of a perception, other than being perceptually aware of it. Now clearly what is implied is not that the part of the soul reported *to* believes the report. This is, first of all, the part that rouses understanding to help unravel the illusion. Moreover it is described as the *recipient* of the report: it has been told by perception what *perception* itself has concluded. Similarly, if you assert something to me, this does not imply that I believe what you say, although it is at least good evidence that

[17] While Plato says a lot about belief in the *Republic*, in particular about how it differs from knowledge, he never goes so far as to define it. I take it that there is nonetheless nothing incompatible—nor, admittedly, any explicit agreement—between the *Republic*'s and Plato's later remarks on belief.

[18] I keep the more literal translation of *phasis*, 'assertion'—rather than, for example, 'assent' or 'affirmation'—to better reflect the analogy between speech and thought: unlike assenting, one asserts something to someone and, so, a belief is a *logos* the soul asserts to itself, while in 'conversation with itself', in direct analogy to asserting something to another person in conversation. As to *logos*, given the Sophist's description of it as a combination a noun and a verb to make sentences such as 'Theaetetus sits' (262C), I see no danger in taking the relevant *logoi* to be statements or propositions.

[19] Note that two separate assumptions I make here are justified within the finger passage, at 524E2–4: that the conflicting perceptions are concurrent and that it is appropriate to use the language of appearances (*phainesthai*; 524E3) to describe this perceptual experience.

you believe it. Thus, the finger passage tell us both that perception—the faculty of perception—forms judgements about the world and that other parts of the soul can be aware of perception's judgements without thereby believing them.

Evidence from later dialogues

In the *Republic* we are given no reason not to take even an illusory sensory appearance to be a kind of perceptual experience in a fairly straightforward way—to say a stick 'appears' bent in water is simply to say that this is how it looks. However, in later dialogues Plato draws a clear line between perception and appearance and he does so at least in part because he believes perception lacks the conceptual sophistication of belief. The later evidence, then, is double edged: it makes it very clear that, at least at some stage of his thought, Plato took appearances to involve belief, but it also presents a very different view of perception than the one found in *Republic* 523A–525A.

In the *Sophist* appearances are seen as partly constituted by belief. They are described as beliefs arising through (*dia*) perception (264A4–6) or 'the blending (*summeixis*) of perception and belief' (264B1–2). Comparison with the *Republic*'s view of appearance is invited by strong similarities between the two discussions of appearances. In the *Sophist* the same range of terms as book 10 of the *Republic* are used—*phainomena*, *phantasmata*, and *eidola*—and appearances are also the product of a dubious craftsman, namely the sophist, who is an 'imitator of real things' (*mimêtês … tôn ontôn*; 235A1) and is, as an imitator, explicitly grouped with the painter (236B9–C1). They also phenomenal appearances, including, again, such natural copies as reflections and, in addition, shadows and 'things in dreams' (266B–C). Thus, at least by the time he wrote the *Sophist*, Plato saw it as consistent to define as 'the blending of perception and belief' the same set of things that are called 'appearances' in the *Republic*, an account which would give us precisely the conceptual relation between appearance and belief needed to explain Socrates' move from belief-appearance conflict to belief-belief conflict in 602C–603A.

However, there is also an important shift in Plato's thought between the *Republic* and *Sophist*. Whereas in the *Republic* perception is construed as a kind of belief, in the *Sophist* belief and perception are sharply distinguished, as is evinced by the fact that being a perceptual experience is not enough to make an appearance judgemental, rather it must be a blend of perception *and* belief. The reason, it seems, is that Plato's view of perception changed. While the *Republic* allows perception broad, propositional content, in a passage of the *Theaetetus*, 184B–186E, we find a very narrow view of the content of perception, according to which we perceive only the proper sensibles corresponding to each sense—colours for sight, sounds for hearing, and so forth. Although not without detractors, a common reading of 184B–186E has it that this passage excludes even the most basic propositional content from perception by denying perception access to the concept of being and so denying it access to the 'is' required for propositions of the form 'x is F' (e.g. this is red).[20]

Nonetheless, in a surprising way the *Theaetetus* also makes it clear that Plato is very aware not only of the alternative view that perception can have broad propositional content but also of the view that perception is a type of belief. Consider Theaetetus's proposal that knowledge is 'nothing but perception' (151E2–3). Not a judgement *based on* perception, but perception itself; Theaetetus is making an identity claim. As it turns out, of course, Socrates rejects this view, largely because it is precluded by his narrow account of the content of perception. But we might reasonable wish to speculate on Theaetetus's own reasons for proposing this definition. First of all, they are not the reasons Socrates gives, the elaborate theory supposedly woven from Protagoraen and Heraclitean ideas: this is a theory Socrates, acting as midwife, offers as a gift to Theaetetus to strengthen his proposal. We should expect the modest Theaetetus to begin with far simpler reasons, especially as it is a tentative, impromptu definition of knowledge that he

[20] Both Michael Frede (1987, pp. 5-6) and Burnyeat (1976, p. 44), for example, argue that being is necessary for any statement of the type 'x is F', Burnyeat on the grounds that *ousia* should be construed predicatively to mean 'being F' and Frede on the grounds that 'in assuming that A is F [Plato] attributes being both to A and to F-ness'. For a sceptical view of such approaches see Holland (1973).

offers. What Theaetetus actually says is that 'it seems to me that one who knows something is perceiving the thing he knows and, so far as I can see at present, knowledge is nothing but perception' (151E1–3). Theaetetus's definition seems to be a generalisation—and this generalising is its most immediately evident fault—from what many would accept as a simple case of knowing: I know the apple is on the table because I see it on the table. What is curious, however, is that Theaetetus does not take perception like this to ground knowledge, but to *be* knowledge.

Some accounts of perception, like that of 184B–186E, focus on the qualities of sensory experience, its 'raw feels'. With this in mind, one might take vision, for example, to yield an unconceptualised colour mosaic that the mind needs to interpret to make contact with the objects that have these colours. So characterised, perception is very different from knowledge, belief or any other intentional state, as it is in the 184B–186E account. However, another account of perception takes it to be intrinsically representational, without any additional interpretive work from a faculty outside of perception. This account is a response to the fact that what we see at least *seems* to be something already fully-formed: apples, people, fingers and the like. This is a more accessible, less theoretical view of perception, at least insofar as it captures how we normally describe our perceptions. It seem highly likely that it is by thinking of perception in this way, as something that represents the world to us, that Theaetetus comes to the conclusion that perception is knowledge. When we see something, in normal circumstances we are not aware of needing to derive from this perception a separate judgement that interprets or confirms what we're seeing: to reach for an apple I do not need to consciously affirm that the apple is, as I see it, on the table. Rather, I simply reach for what I see. At least on the level of conscious awareness, then, there is usually no gap between perception and judgement—perception, as a faculty that represents the world, is sufficient without the aid of thought. Nor is this a particularly implausible suggestion: representing the world to be a certain way is just as credible a characterisation of the function of belief as it is of perception.

Now, of course the idea that perception is knowledge is anathema to Plato's

philosophy, in any period of his thought, but nonetheless it strikes me that by giving this view to Theaetetus, one of his most gifted interlocutors (143E–144B), Plato shows awareness of the fact that there is plausibility in the idea that perception is an intentional state—belief if not knowledge. So if in 184B–186E Plato is introducing his new theory of perception, it seems reasonable to speculate that we see something of his earlier view of perception in the position that Theaetetus defends. Michael Frede recognises this connection, seeing the *Theaetetus* as a dialogue that attempts to 'untangle the conflation of perception, belief, appearance, and knowledge' (1987, p. 5), a conflation that he takes both Plato himself to have made—excluding, of course, knowledge—in the *Republic* and equally a conflation to be found generally in Greek thinking prior to the *Theaetetus*. In ordinary usage, he argues, *aisthanesthai* is broader than mere sense-perception and encompasses belief:

> [B]ecause one does not see a radical difference between the way the mind grasps something and way the eyes see something. Both are supposed to involve some contact with the object by virtue of which, through a mechanism unknown to us, we become aware of it. (1987, p. 4)

In contrast to Frede, however, I do not think Plato's view of perception in the *Republic* should be seen as a confusion, stemming from a failure to fully grasp the difference between belief and perception. I would rather characterise it as a view of what perception is—a view that may be a first impression, and lack the deeper theoretical foundation of his later view, but one that is a response to real features of perception and is, in fact, defensible. Indeed, we might say that in the *Sophist* he retains a large part of his earlier view by moving the doxastic work that perception used to do entirely onto appearances, using the latter to explain, for example, both false perceptual beliefs and the power over our beliefs held by 'appearance-makers' like sophists, poets, and painters.[21]

[21] See Silverman (1991) for an interesting argument for a broader role for appearances in Plato's later account of perception. Silverman argues that appearances allow Plato to be faithful to the conflicting intuitions behind the two accounts of perception in the *Theaetetus*. As he summarises his position:

> We actually sense only the special sensibles, and yet we think we sense colored objects and

Back to 602C–603A

There is excellent evidence, then, that in the *Republic* and later dialogues Plato saw a conceptual link between appearances and belief: appearances are (*Republic*) or are partly (*Sophist*) constituted by beliefs. If this is right, the move from belief-appearance conflict to belief-belief conflict in 602C–603A ceases to be puzzling: the fact that a submerged stick 'appears' to us to be bent, and continues to do so even when we have concluded that it is straight, shows that a part of us believes what appears simply because appearing is itself a kind of believing. Thus, the move from claim (4) to claim (5) is rendered valid by the implicit premise **P**, a premise that is justified by the arguably pre-theoretic assumption that perceptual experiences are representational states that are sufficiently similar to rational beliefs to be consider another kind of belief: both can be characterised as ways in which we represent the world as being a certain way. The fact that we experience visual illusions that we rationally disbelieve shows, then, that we can at the same time experience opposing representational states, one that affirms *p* and one that denies *p*, offering sufficient opposition for the PO to apply (note that even in book 4 'assent and denial' (437B1) are included among the oppositions to which PO can be applied).

We also saw compelling evidence that the 'lower' or 'inferior' believing part of the soul is one or both of the non-rational parts of the soul. Given that the appearance-language in 602C–603A must be read in a phenomenal rather than judgemental sense, claim (4), 602E4–9, no longer poses an obstacle to this reading. In fact, it tells us only that the rational part has, as we would expect, access to what appears to be the case, without entailing that it believes what appears. As we see in the finger passage, being appeared to can be likened to someone asserting a claim to us. So a part of the soul appeared to, the part of the soul that receives perception's 'report' or 'announcement', need no more be the believer

their properties. I believe that Plato posits the faculty of *phantasia* to fill this gap. *Phantasia*'s role is to link the limited irrational findings of *aisthêsis* to the conceptual faculty of belief. In keeping with its Protagorean heritage, *phantasia* becomes for Plato the faculty responsible for the rich tapestry of what is apparently sensible. (p. 133)

than someone who has a claim asserted to them in conversation. A part of the soul can be *aware* of the content of the appearance, without in any way *agreeing* with it.

3.2.3 *Concluding observations*

Many modern theorists have been struck, like Plato, by the fact that believing and perceiving are both ways in which we represent something about the world to ourselves. This is not a superficial similarity: representing the world to be a certain way would seem to be the primary function of both perception and belief. Furthermore, if they are also both propositional, then their content can be given the same description: if I see that an apple is on the table and I believe it is on the table, then what I see and what I believe is the same thing—that an apple is on the table. These parallels between perception and belief led Armstrong to propose the influential 'belief theory' of perception, according to which perception is constituted by the acquisition of a belief.[22] The belief theory has generally been thought most compelling as an explanation of normal, veridical perception. In normal circumstances, we do not need to infer a separate belief to be able to act on our perceptions. If I open a door, I see the handle and reach for it. I do not need to consciously infer that the handle is, as it looks, in front of me. In such cases, the burden of proof is on the side of those who suppose that perception is not sufficient, that there must in addition be some unnoticed inference from the perception to an affirming belief if acting on a perception is to be possible. However, the belief theory is thought to be weakest when explaining cases of belief-appearance conflict, when one explicitly disavows belief in what one seems to see. This has sometimes led to retreats from the belief-theory, for example towards the view that perception is constituted by an inclination to believe. Others have bit the bullet, claiming, like Plato, that in cases of belief-appearance conflict we

[22] See Armstrong (1968) Ch. 10. He calls perception the *acquisition* of a belief, rather than simply a belief, because he sees perception as an event and belief as a state (213-16). Nonetheless, as an acquisition of a belief, a belief is still seen as constitutive of perception insofar as the content of perception will be identical to the content of the acquired belief.

are suffering from conflicting beliefs. Smith (2001), for example, argues that such a conflicting belief is indicated by the fact that in some cases of belief-appearance conflict we nonetheless act on the appearance.[23] Smith gives the example of aversive behaviour to what I know to be a mere hallucination of a spider crawling on my arm:

> Although my linguistic and betting behaviour fully entitles the attribution of non-belief to me, my aversive behaviour as the spider seems to encroach calls for a quite different explanation. Such behaviour surely need not be a mere reflex requiring no belief-involving reason. I am not at all sure that we shouldn't follow this inclination, and say that I both perceptually believe that there is a spider there, and theoretically believe that there is not. We are, indeed, enjoined so to attribute beliefs as to render persons intelligible to us; but possession of such a pair of contradictory beliefs is far from unintelligible. Doubtless a subject who has both these beliefs is in an irrational state; but *the senses are irrational*.[24]

We notice here a misalignment between what we typically call beliefs, an example of which would be Smith's 'theoretical belief', and the putative perceptual beliefs that occur in belief-appearance conflict. Unlike the belief that constitutes a veridical perception, which will be in tune with one's 'linguistic and betting behaviour', the perceptual belief that Smith educes here is clearly not what we would standardly call a belief. The best indication of this is the fact that the person who suffers the hallucination in this example would deny holding it. There seems, then, to be a sense of ownership over the theoretical belief that is lacking in relation to the perceptual belief.

One can justly demand an explanation for this. Since Smith attempts to assuage worries about positing conflicting beliefs with the claim that the senses are irrational, he might be gesturing at the idea that perceptual beliefs are outside of the rational structure that we take rational beliefs to fall into (i.e. standard beliefs tend towards mutual implication, abhor contradiction, and are responsive to evidence and reasoning). Thus, a belief theory of perception coupled with the

[23] Armstrong is less explicit, but seems to entertain a similar view. He talks of 'half-believing' in such cases and of having an inclination to believe that he cashes out as having a 'belief that is held in check by a stronger belief' (1968, p. 221). See also Gluer (2005), p. 6.

[24] Smith (2001, p. 292). An influential exploration of such 'belief-discordant' behaviour is Gendler (2008a), which I consider in the final section of chapter 5.

fact that we encounter perceptual illusions will suggest that we have a divided believing faculty: one set of beliefs that fall into a rational structure and that we feel ownership over and a second set of perceptual beliefs that can lie outside this rational structure and sense of ownership. As Gluer puts it, such an idea makes perception 'an autonomous little homunculus telling you what *he* believes' (2005, p. 6). However, unlike modern theorists I've mentioned, Plato is very well placed to explain this. Since he defends a partite psychology, conflicting beliefs can be held by different parts of the soul—indeed the Principle of Opposites implies that they must be—and so the fact that they behave in different ways, and even conflict, need not be surprising. Contrary to first impressions, then, we can take a very positive view of the argument at 602C–603A: not only is it valid but also, by marking out two believing parts of the soul, it offers a solution to a well-recognised difficulty faced by any belief theory of perception including its contemporary variants.

CHAPTER 4

Three Removes From the Truth: *Republic* 595C–598D

The previous chapter looked at the basic role appearances play in non-rational cognition: the non-rational parts of the soul are sensitive to appearances, but insensitive to calculation. In this short chapter, I look at another passage from book 10 of the *Republic*, 595C–598D, a passage that specifies appearance's position—a distinctly lowly position—in the *Republic*'s metaphysics and epistemology. I look especially closely at a puzzling argument that has been taken by some commentators to imply that appearances are a part of normal, veridical perception (597E10–598B8). I argue that the lesson we should draw is rather that the deceptiveness of appearances derives ultimately from certain essentially deceptive features of perception.

The passage I wish to examine explains why it is a bad thing that imitative appearances affect the non-rational parts of our soul. At 603B4 Socrates gives a succinct summary of his criticism of imitation: 'imitation is (1) an inferior thing that (2) consorts with an inferior thing to (3) produce an inferior thing' (*phaulê ara phaulô suyyiynomenê phaula genna hê mimêtikê*). Claim (2) was established by the argument that was the subject of the last chapter, 602C–603A, which showed that imitative appearances appealed to an inferior part of the soul. On its own, however, this is not sufficient to draw the conclusion, (3), that imitations effect on our non-rational parts produces something inferior. One might argue, and indeed many think, that poets like Homer make imitative appearances that are

true and edifying and, therefore, likely to have an improving effect on our non-rational parts. What I examine here is Socrates' response, his defence of claim (1): imitative appearances are, necessarily and by nature, misleading.

4.1 THREE COUCHES: 595C–597E

In the previous chapter we saw how Socrates appeals to the psychology developed in book 4 to explain imitation's effect on the soul. In a similar way Socrates turns to the metaphysics of the *Republic*'s central books to give an account of 'what imitation in general is' (595C7), a question that was broached solely through examples in book 3. He suggests that they begin their answer by applying 'our customary method' (*eiôthuias methodou*; 596A6), which is described as to 'posit a single form (*eidos*) for each of the [sets of] many things to which we apply the same name' (596A6–8). Socrates chooses the example of a couch (*klinê*).[1] Thus there are the many material couches that share the name 'couch' and a single Form of couch set over them.[2]

Socrates' next step is to add to these two kinds of couch yet a third kind, an *appearance* of a couch. He tells us that one way in which one can make a couch—or indeed 'in a way' (596D3) make any material thing—is by wielding a mirror to

[1] Or 'bed'. A *klinê* is what, for example, guests would recline on at Greek symposia. Burnyeat gives an interesting overview of the significance this example might have had for a contemporary reader of the *Republic*, noting, for example, that a symposium was a setting for 'food, drink, sex—and poetry' (1999, pp. 232-236).

[2] The oddity of the idea of a Form of couch has led some to believe that Plato did not intend this to be a serious piece of metaphysics, since typically, and elsewhere in the *Republic*, we find Forms only for a range of relational or evaluative properties, like largeness or justice, not for objects that are material and man-made. See Fine (1993, pp. 110-119) and Halliwell (2002, p. 57). While a Form of couch is unusual, I find it hard to believe that Socrates characterises 'our customary method'—a vital step in his thought—in a misleading way simply to permit this example: he says that they 'posit a single form for *each* of the many things to which we apply the same name' (Fine suggests we avoid this by borrowing a narrow sense of 'name' from the *Cratylus*, but with no indication from within the *Republic*, this strikes me as too subtle). In the *Parmenides* Plato presents Socrates as having wavered about what kinds of Forms there are—he has often been 'in doubt' about a Form of man or fire but is confident that hair or dirt are 'just what we see' (130B7–D9)—which strikes me as an indication that Plato was at least not consistently wedded to the idea that there are only Forms of properties. For a detailed defence of the claim that Plato is serious about the Form of couch see Burnyeat (1999, pp. 245-249)

catch its reflection (596C5–596E3). Another way is by making a painting of it:

> S: For I think the painter too belongs to this class of makers … [He] doesn't truly make the things he makes. Yet, in a certain way, the painter does make a couch, doesn't he? G: Yes, he too makes the appearance of one. (596E5–11)

The imitative painter, then, falls into the same category of maker as a mirror-wielder because he produces the same kind of thing: 'appearances' (*phainomena*), 'images' (*eidola*), or 'semblances' (*phantasmata*) of real things.

Socrates' aim is to show that this kind of product is 'by nature third from the king and the truth' (597E6), that is, that it is a product that is at three removes—according to Greek inclusive counting—from the nature of each thing, the Form, and suffers a degradation of 'truth' and 'reality' (in some sense) at each removal. He first considers the carpenter who looks to the Form to make a material couch that is second from the truth. The carpenter 'isn't making that which is (*to on*), but something like (*hoion*) that which is, but is not that which is', something that is not 'completely real' (*teleôs … on*; 597A4–5). Socrates concludes from this that we shouldn't be surprised 'if the carpenter's couch, too, turns out to be a somewhat dark affair in comparison to the true one' (579A10–11). The 'too' (*kai*) invites us to compare the carpenter's couch's predicament with that of the painter's: it is like the carpenter's couch, but not it, and correspondingly even less 'completely real'. These three kinds of couch, then, are related by likeness and distinguished by being (in some sense) to a greater or less degree 'real' or 'true': the completely real couch, the Form; the many material couches that are like the Form but less real; and the apparent couches of the mirror-wielder or painter that are like material couches but even less real.

As with similar descriptions in books 6 and 7—for example, in the allegory of the Cave the puppeteer's originals are said to be 'more real' (*mallon onta*; 515D3) than the shadows they cast—this failure to be 'completely real' presents us with the difficult idea that the carpenter's and painter's couches are real only in some partial way. They become less cryptic, however, if we place them in the context of Socrates' more detailed descriptions of the differences between the three couches. Socrates tells us that the carpenter 'doesn't make the form, which

is our term for the being of a couch (*autên ekeinên ho esti klinên*), but rather one particular couch (*klinên tina*)' (597A1–2; see also 597D2). The painter, in turn, makes a particular couch only 'in a way' (*tropô ge tini*; 596E10), but in another way he doesn't make a couch (cf. 596D3–4)—a confusion that Glaucon clears up to Socrates' satisfaction by describing it as an appearance of a couch but not such things 'as they are in truth' (596E4). There is no suggestion here either that the carpenter's couch is not a completely real *particular* couch or that the painter's couch is not a completely real *appearance* of a couch.[3] Nonetheless, it does indeed suggest that both of them are, each in their own way, 'removed' from the nature of a couch—one by being just a particular couch and the other by being just an image of a couch. That is, both are couches only of a qualified sort, representing only one small way of being a couch and so giving us a less than complete picture of what it is to be a couch. Only the Form, the nature of the couch, presents us with a complete, unqualified instance of a couch, making it the only truly reliable guide to understanding what a couch is.[4]

Turning to the painter's couch, then, we see that it is removed from the Form of a couch—is a less 'real' or 'true' example of a couch—for two separate reasons: first, it shares the material couch's predicament of being just one particular couch and, second, it has its own peculiar shortcoming of being a mere appearance of a particular couch. Thus a painter's product is 'by nature third from the king and the truth' because it presents a doubly obscured example of the nature of what he paints—we learn little about couches by looking at just one example and even less by looking at a painting of one example. Socrates takes this to be sufficient answer to the question of 'what imitation in general is' (597E3–10). Being a maker of images, third from the truth, is later described as how they 'defined' the imitator: *eidôlou dêmiourgos, hon dê mimêtên hôrisametha* (599D2–4; see also 601B9).

[3] Compare the following exchange in the *Sophist*: 'V: So you're saying that that which is like (*to eoikos*) is not really that which is, if you speak of it as not true. T: But it *is*, in a way. V: But not truly, you say. T: No, except that *it really is a likeness*.' (240B7–11; my emphasis).

[4] Thus I am sympathetic to Vlastos' (1965) interpretation of the degrees of reality claims we find in the *Republic*.

4.2 IMITATING APPEARANCES: 597E–598D

What Socrates describes as third from the truth is not just the product of imitation (597E3–4) but also the imitator himself (597E6–7). Thus when he turns to poetry he imagines asking Homer:

> Homer, if you're not third from the truth about virtue, a maker of images, which we just defined the imitator to be, but even if you're second and capable of knowing which ways of life make people better in private or in public, then tell us which cities are better governed because of you… (599D2–6)

Socrates believes, then, that an imitator does not need—and in fact does not possess[5]—knowledge of what he imitates. An imitator's state of knowledge is, like his product, third from the truth, which means that his knowledge of what he imitates extends only as far as what will be revealed by appearances. As Socrates summarises his view: 'we say a maker of images—an imitator—knows nothing about that which is but only about its appearance' (601B9–10).

Socrates recognises that this is not at all obvious. Many people think that 'if a good poet produces fine poetry, he must have knowledge about all the things he writes about, or else he wouldn't be be able to produce it at all' (598E3–4). In fact the imitator appears to be exceptionally knowledgeable, being versed in all crafts and every subject, so as to appear almost 'omniscient' (*passophos*, 598D5–C6; in this respect, he is comparable to the orator of the *Gorgias*). Moreover, Socrates himself seems to have implied that the imitator is closer to the truth than his imitations are. What the imitator looks to when imitating is what is *second* from the truth—a painter of a couch imitates material couches, not other paintings of couches—so he would seem to be a step above someone whose knowledge of

[5] Thus there are two claims here, a weaker and stronger one. The weaker claim is that an imitator does not *by virtue of being an imitator* have knowledge of what he imitates. This implies only that insofar as one imitates one does not need or apply any knowledge of what is imitated, even if one does in fact possess this knowledge. Second is the stronger claim that any actual imitator does not in fact *posses* knowledge of what he imitates. From 598D–601C Socrates defends this second claim, at least with regard to the poet, with a variety of counterfactuals: for instance, if an imitative poet were to know how to perform the sort of heroic, virtuous actions he appears to represent in his characters, then he would cease poetry and become a hero himself, being 'more eager to be the eulogised than the eulogiser' (599B6–7).

a subject is confined to paintings and poems. Socrates' claim, then, requires defence and he offers a series of arguments for it. The first and most comprehensive of these arguments claims that in fact an imitator does not, strictly speaking, imitate what is second from the truth, but something that is, like his product, third from the truth: to imitate a couch he needs to know 'only about its appearance' because what he is imitating is not really the couch, but just another appearance of it. His argument for this surprising conclusion reveals a lot, and has added some confusion about, the nature of appearances.

Having defined the imitator, Socrates asks Glaucon to consider a 'further distinction' (*eti diorison*): does the painter imitate a couch 'as it *is* or as it *appears*' (*hoia estin ê hoia phainetai*; 598A5)? He illustrates what he means with the following example:

> If you look at a couch from the side or the front or from anywhere else is it different from itself, or does it only appear different, without being at all different? And is that also the case with other things? (598A7–9)

Our perception of a couch, as of anything else, is necessarily perspectival. One might conclude from this that from any one perspective I can see only a *part* of the couch, but that it is certainly the couch itself that I'm seeing—if I see a couch from the side, surely I see the side *of the couch*. Socrates comes to a different conclusion. He brings out the distinction he wishes to draw more precisely as follows:

> Now consider this very point: what does painting do in each case? Does it imitate the existing thing, as it is (*to on, hôs ekei*) or does it imitate the appearance, as it appears (*to phainomenon, hôs phainetai*)? Is it an imitation of a semblance (*phantasmatos*) or of the truth? (598B1–3)

Note that the distinction is not between 'the existing thing, as it is' and 'the existing thing, as it appears', but between two different *subjects* of imitation, an existing thing and an appearance.[6] It seems, then, that from the occurrence of

[6] What would it be to imitate 'the existing thing, as it is'? In the *Cratylus* Socrates says that 'an image cannot remain an image if it presents all the details of what it represents'—not just the shapes and colours a painter captures, but all of the physical characteristics of the object—because it would then produce an exact replica of the object, rather than an image of it (432A–C). Perhaps, then, to imitate a couch 'as it is' is to make another couch rather than a painting of it (and, therefore, would not be a genuine case of imitation).

perspectival variation Socrates concludes that what is imitated cannot be the existing thing, the material couch, but must be something else, an appearance of a couch.

This is a sufficiently surprising conclusion to warrant double-checking that this is indeed what Socrates intends here. What offers the most interpretive latitude is the *to phainomenon* at 598B3. This might be translated 'what appears' and this, in turn, might be taken to refer not to an appearance, but to the material couch *that* appears.[7] On this reading *to on* and *to phainomenon* are locutions with different emphasis but the same reference, the material couch, so what the painter imitates would not be an appearance construed as something separate from the couch, but simply the couch, as *it* appears. However, the text gives two good reasons to resist this reading. First, what Socrates calls *to phainomenon* he also calls a 'semblance' or 'phantom' (*phantasma*), a word that implies some lack of reality, making it a very awkward description of a real couch. Even a part of the couch could hardly be construed as a 'semblance' of the couch. Second, and most decisively, Socrates' concluding statement very strongly favours the first reading: 'it is because of this that it [sc. painting] can make everything: it lays hold of only a small part (*smikron ti*) of each thing *and that part is only an image* (*eidôlon*)' (598B7–8; my emphasis). With the addition that the part a painter imitates is only an 'image', it would seem that Socrates is deliberately emphasising that the part he refers to is *not* a material part of a couch—a section of a couch that one might cut off—but a new item in addition to the couch: an appearance, semblance, or image of the couch.

Consequently, as a number of commentators have noticed, Socrates' argument bears a striking resemblance to the argument from perspectival variation appealed to by empiricist and sense-date theorists: the couch itself doesn't change when we vary our perspective, but what we see does change, therefore the couch

[7] While a possible translation, I am not aware of any translator who renders this particular line like this, although Barney (1992) does argue that occurrences of *to phainomenon* in book 10 refer either to the *thing* that appears so-and-so or the belief that it is so-and-so. However, she does not discuss 598B1–3.

and what we see must be two different things.[8] What perception makes available to the imitator is not the couch itself, or even a perspectivally circumscribed part of it, but a different kind of thing altogether: an image of the couch that, it seems, mediates our apprehension of it. Thus, the imitator's product and what he imitates to make this product are of the same kind: both are mere appearances of things that are second from the truth. The imitator, then, is not someone who tries to capture the truth in his product, but someone who merely *reproduces* in a painting or a poem how things appear to him, placing the imitator himself and his product at a third remove from the truth.[9]

4.3 ARE APPEARANCES INVOLVED IN VERIDICAL PERCEPTION?

We have seen that Plato uses 'appearances' to refer to a range of perceptual experiences: paintings, reflections, and various optical illusions. The argument from perspectival variation seems to make a surprising extension to this range. Appearances seem now to play a role in normal, veridical perception: if what I encounter when I see a couch is an appearances, then surely *everything* I see can be called an appearance? Plato typically uses the word 'appearance', in both this and related philosophical contexts, to refer to a *mere* appearance, that is, an appearance that is in some way misleading or false and that will, therefore, lead us to a false belief:

[8] Paton (1922) and Hamlyn (1958) draw comparisons between Plato's views—both in this passage and his account of *eikasia* in books 6–7—and, respectively, the empiricism of Berkeley and Hume and the sense-data theories of the early 20th century. Compare, for example, Hume's argument in the *Treatise*, XII.1: '[P]hilosophy teaches us that images (or perceptions) are the only things that can ever be present to the mind.... The table that we see seems to shrink as we move away from it; but the real table that exists independently of us doesn't alter; so what was present to the mind wasn't the real table but only an image of it.'

[9] Moss (2007b), Paton (1922), and Nehamas (1982) offer similar readings of this passage. Nehamas agrees that Socrates 'is thinking of the object of imitation and the product of imitation as the same object—if not in number, at least in type', but he goes on to characterise the object and product quite differently: 'the image is both the surface of the subject and the product of the painter' and it is as if 'the painter lifts the surface off the subject and transplants it onto the painting' (1982, p. 62; see also 1988, p. 220). Although it may not be Nehamas' intention, the surface of a couch—which is a three-dimensional, material part of a couch—cannot be the same in type as a painter's image of a couch.

pleasure appears good but is bad; a submerged stick appears bent but is straight; and a painting or reflection appear to be, but are not, what they represent. As I hope to show, Socrates' argument from perspectival variation is not in fact an exception to this rule. The lesson we should draw from it is not that appearances are entailed in all perception, but that Socrates takes very little perception to be veridical.

Consider Socrates' view of the optical illusions listed at the beginning of the argument considered in chapter 3:

> Through sight the same magnitude doesn't appear to us to be equal when near and far away … And something looks crooked when seen in water and straight when seen out of it and the same thing is seen to be both concave and convex on account of the eye's wandering anew around the colours, and every similar sort of confusion clearly exists in our soul. It is by exploiting this condition in our nature that shadow painting, wonder-making, and many other such contrivances have powers that are little short of magical. (602C7–D4)

Socrates lists a number of naturally occurring illusions and two man-made illusions.[10] Naturally occurring illusions are said to be a confusion that 'exists in our soul' (*enousa … en tê psuchê*) and man-made illusions succeed by taking advantage of this 'condition in our nature' (*pathêmati tês phuseôs*). In other words, what is responsible for these false appearances is not, or not only, something outside of us—the appearances of a couch is not an external object like a material couch—but is at least in part explained by facts of our psychological. This fits what we learned about appearances in chapter 3: a straight, submerged stick 'appears' bent because of a limitation of our perceptual faculty, which mistakenly reports that it is bent. The error is a result of the non-rational (perceptual) parts' inability to apply 'calculation' to compensate for the limitations of the perceptual faculty, leaving us to rely on the rational part of the soul to come to a second, correct conclusion.

Now, consider the first illusion: things of equal size look larger or smaller de-

[10] *Skiagraphia*, shadow-painting, is often compared to *trompe l'œil* painting, but seems rather to have been a sort of impressionism that can be deciphered only from a distance. It is more difficult to know how to understand *thaumatpoiia*—literally 'wonder-maker'—although it is notable that the same word is used in the Allegory of the Cave to refer to puppetry (see Ch. 5, n. 16 below)

pending on how near or far from them we are. This is only an optical 'illusion' in a very attenuated sense. It is a normal result of properly functioning perception; variation in size is just one way in which perception represents distance to us. We would be more likely to call the opposite an illusion. That is, if something were to look the *same* size despite the fact that it was becoming more distant—if we judged by how perception normally functions, this would lead us to falsely believe that it was either not moving further away or was growing in size. J.L. Austin, responding to the argument from illusion, applied the same reasoning to the appearance of a straight, submerged stick: light passing from water to air is subject to refraction, so a straight stick *should* appear bent when it is submerged. An optical illusion, in a stricter sense, would occur if a straight stick continued to appear straight even when it was submerged: if we applied our knowledge of optics, this would lead us to falsely conclude that it is an apparently straight stick that is in fact bent.

In what sense, then, might Socrates take these to be illusions? Despite being attempts to deflate the idea that they are optical illusions, what the above responses could be taken to show is that we need to employ basic reasoning, however rudimentary and instinctively applied, to successfully navigate even the simplest perceptual information. For example, we can judge that a person is about ten feet away only by reasoning that an average person at a distance of ten feet will take up about this much space in our visual field. The non-rational parts of the soul are incapable of such reasoning. As we saw in the previous chapter, Socrates says that the images of the imitative poet affect 'the part of the soul that is unthinking and doesn't distinguish greater from lesser, but believes the same things are at one time large and another time small' (605B8–C3). Likewise, it will presumably believe that the same couch is at one time a narrow L-shape, when seen from the side, and at another time rectangular, when seen from the front. Moreover, it will presumably be unable to distinguish a painting of a couch from a real couch.

Thus, much of what we might think of as veridical perception, Socrates treats

as non-veridical perception that can easily be rectified by basic reasoning applied by the rational part of the soul. So perception is not only misleading when it goes wrong in some way, rather perception is a faculty that, even when functioning properly, often presents false information (where a measure of whether or not the information is false is whether or not it would lead one to a false belief if judged *solely* by what perception reports). Whether perception presents true or false information is largely determined by the kind of content it can have. As we saw in the previous chapter, unlike Plato's later dialogues, in the *Republic* perception is given very broad, representational content. For example, we can see that a finger is a finger without the aid of the rational part of the soul ('an ordinary soul is not compelled to ask the understanding what a finger is'; 523D3–5). In such cases the reports of perception are sufficient to ground true belief. Perception fares worse, however, when it comes to properties like large or small, hard or soft. For example, we can perceive relational differences in size—we can see that the ring finger is smaller than the middle finger and larger than the little finger—but we can't perceive something's size *per se*, that is, we can't see that something is ten centimetres long, or any consistent size, but only that it is sometimes small (when distant) and sometimes large (when close) or both small (in relation to the middle finger) and large (in relation to the little finger). The qualifications in parenthesis are not available to perception, so by sight alone it appears that something changes from small to large or is both small and large.

We can give a similar analysis of a couch's varying perspective. We can't see a couch's shape *per se*, so perception can only indicate that it has at one time a narrow L-shape (when seen from the side) and at another time a broad rectangular-shape (when seen from the front). Since it has no non-perspectival conception of a couch's shape, perception has no way to unify these two perspectives into a single, unchanging couch. Consequently, from varying perspectives a couch's shape will, as Socrates describes it, 'appear different, without being at all different'. This is a mere appearance—thus Socrates calls it as a 'semblance'—and so it is misleading in much the same way as our perception of the shape of a submerged

stick. To know the truth about a couch's shape, and to know that the couch itself does not change shape, we must use cognitive abilities that are unavailable to perception and, thus, unavailable to the non-rational parts of our soul.

Socrates' argument from perspectival variation takes advantage of this very narrow conception of veridical perception. Thus, despite their striking similarities, the *aim* of this argument is quite different from the arguments from illusion that we find in, for example, Hume or Ayer: whether or not it is in fact a consequence of his argument, his aim is not to refute direct realism. Socrates wants to show that what the painter needs to know if he is to be able to paint a couch is something different from and less true than—third from the truth—what the carpenter needs to know to make a couch. As elsewhere in the *Republic*, his approach is to show that what they know are two different objects: an appearance of a couch for a painter and a material couch for a carpenter. The argument emphasises the epistemic inferiority of the former—sight reveals only a mere appearance of a couch, a 'couch' whose shape changes with perspective—not the fact that it mediates our apprehension. It would be anachronistic, then, to conclude as Paton (1922) and Hamlyn (1958) do that Plato thought all perception is mediated by something analogous to sense-data. This is not to say that Plato is a direct realist, but simply that he would not call all perceptions mere 'appearances'. He calls the couch the painter imitates an 'appearance' to signify that there is something about it that is false. This is a subtle but important point. Widening the sense of 'appearance' to include all perceptual experiences, and permitting true appearances, would have broad implications. In particular, it would jeopardise Socrates' claim that appearances are corrupting. On the reading I have offered, appearances are essentially misleading with respect to relational properties, such as largeness or equality, because these are beyond the reach of perceptual representations.

As I will argue in the next chapter, this is equally true of evaluative properties. While book 10's discussion of painting tells us a great deal about the imitator and his product, its purpose is to prepare us for his discussion of poetry. We are

not greatly harmed by paintings of couches, let alone by reflections or optical illusions, but the poet's product is very different in this respect: his form of imitative appearance 'is likely to distort the thought of anyone who hears it' (595B5–6). The difference is that what the poet imitates is not, like a couch or stick, something evaluatively neutral. He imitates 'images of virtue' (*eidôlôn aretês*; 605E5). Nonetheless, the lessons we learn from looking at the appearance of a couch still apply: the poet's 'images of virtue' appeal to the non-rational parts of the soul and they are, like all appearances, essentially deceptive. Thus, they pose two closely connected dangers: they stir and strengthen harmful passions in the lower parts of our soul and, unlike submerged sticks, they threaten to corrupt the whole soul, in belief, desires, and action.

CHAPTER 5

Appearances & Poetry in the *Republic*

What emerged from my analysis of 602c–603a in chapter 3 is an account of non-rational cognition. The non-rational parts of the soul are insensitive to calculation, but sensitive to appearances—to phenomenal, sensory representations provided by the perceptual faculty—and this sensitivity to appearances is for Plato a form of non-rational belief: just as rational beliefs represent the world to the rational part of the soul, appearances represent the world to the non-rational parts of the soul. As we saw in chapter 4, the result is that the non-rational parts of the soul are guided by perceptual appearances that are, by their nature, 'third from the truth' and, therefore, deceptive.

Compare this with the role of appearances in the *Protagoras* and *Gorgias*. In these dialogues, I argued, Socrates introduces a kind of judging, judging by appearances, that doesn't employ any reasoning or calculating—the 'art of measurement'—but instead consists entirely of unreflectively accepting what appears to be the case. This led to an intellectualist account of appetitive desire: since what appears best is lower, bodily pleasure, unreflectively accepting appearances leads one to blindly pursue appetitive pleasures like food, drink, and sex. The similarities with the *Republic* are striking, as is the most obvious dissimilarity: in the *Republic* judging by appearances is not just something that a whole person, if not well educated, can lapse into, it is a permanent feature of two parts of our soul. Unsurprisingly, these are the parts of the soul responsible for our lower, non-rational passions.

This suggests the following connection between non-rational cognition and non-rational conation: the appetitive and spirited parts of the soul have their distinctive passions, passions that are often at odds with the calculated aims of the rational part, precisely because they are compelled to judge by mere appearances. Put another way, the cognitive conflict described in 602c–603a explains the motivational conflict described in book 4. The conflict in 602c–603a is of course *purely* cognitive, but this is because the perceptual illusions it considers, such as the apparently bent stick, happen not to stir the passions of any part of the soul. Other appearances, appearances that are relevantly similar to perceptual illusions, do stir the passions, representing something that is, for example, desirable, angering, or frightening. Like the apparent crookedness of a submerged stick, these appearances can be illusions: what appears to be worthy of anger can, after calculation, be revealed to be something that should be responded to calmly. When we encounter this kind of appearance, motivational conflict can ensue. Our spirited part is incapable of calculation, so it is compelled to believe that it is as worthy of anger as it appears; our rational part, if it is in a healthy condition, can use calculation to conclude that remaining calm is more appropriate. This is, in the first instance, cognitive conflict of the kind we found in 602c–603a, but now the opposing beliefs warrant different emotional responses: our spirited part gets angry and our rational part must try to restrain it.[1]

Crucially, this shows that motivational conflict need not entail that the parts of our soul hold a different general conception of what is desirable, angering, or frightening. Rather, in certain circumstances they can reach different conclusions, due to their different cognitive abilities, about whether something is a *bona fide* example of something desirable, angering, or frightening. The spirited and rational part can agree, for example, that what is frightening is expected harm, but the spirited part may fear apparent harm even when the rational part knows that this appearance is deceptive.[2] Appearances account both for this cognitive dis-

[1] In this view of the relationship between non-rational motivation and non-rational cognition, I have been greatly influenced by Jessica Moss; see especially her (2007a) and (2008).

[2] As we would expect if motivational conflict has its root in cognitive dissonance, a virtuous

sonance and, relatedly, for how a person as a whole can be misled. The *Republic* is, like the *Gorgias* and for similar reasons, pessimistic about the general state of ethical education. When it comes to evaluative facts, most people assume that things are simply as they appear, while this is in fact where we find the sharpest disparity between the apparent and the real. The majority, consequently, are like the prisoners in the Allegory of the Cave, confined to an awareness of mere shadows of the good or just. Plato now has a sophisticated psychological account of how people are deceived by such appearances. Appearances exert their power on the non-rational parts of a person's soul. They can, therefore, circumvent our better judgement, persuading a lower part of us even when our rational part is fully aware of the deception. This sets our non-rational parts at odds with our rational part, causing motivational conflict and threatening *akrasia*, but it also has the more serious effect of creating the conditions needed for the appetitive or spirited part to assume 'rule' of the soul. What it means for a person to be 'ruled' by, say, the appetitive part of the soul is for all three parts of his soul to devote their cognitive and motivational capacities to appetitive ends. One result of this is that even the rational part of his soul assesses evaluative facts as the appetitive part would—uncritically accepting mere appearances of the good or just. Appearances, then, corrupt a person from the bottom up, persuading first the lower parts of a his soul and then, through their corrupting influence, the person himself.

---

soul, free of internal conflict, is characterised by cognitive harmony. We are courageous when 'the spirited part … preserves through pains and pleasures the declarations of reason about what is and what isn't to be feared' (441C1–3). As the account of civic courage explains (429B5–430B4), this 'preservation' is to stably hold on to a true beliefs in the face of pleasures or pains that encourage one to believe the contrary—an account that should remind us of the belief-changing 'power of appearance' exerted by immediate pleasure and pain in the *Protagoras*. Similarly, temperance occurs when all three parts of the soul 'believe in common (*homodoxôsi*) that the rational part should rule and don't engage in civil war against it' (442C11–D1). In book 9 Socrates argues that since the appetitive and spirited parts are deceived about their own ends, it is better *for them* to 'follow knowledge and argument and pursue with their help the pleasures that reason approves' (586D5–7). Thus, in a harmonious soul they are not forced to settle for less or inferior pleasure, but rather guided, since they are incapable of finding them by themselves, to 'the truest pleasures possible for them … and the ones that are most their own' (586D8–E1). Motivational harmony, then, is achieved once there is cognitive harmony, which is characterised as all parts of the soul agreeing to the rational part's correct conception of what is, for example, fearful or pleasant.

In this chapter I will explore and substantiate the claims made in the previous few paragraphs. They are, for the most part, developed in Socrates' criticism of imitative poetry in book 10. As a paradigmatic source of ethical miseducation, poetry serves as a model for the corrupting effect of appearances. It also highlights the fact that the pervasive influence of appearances is not a purely natural phenomenon, but something that is greatly amplified by certain 'image-makers' (*eidôlou dêmiourgoi*): poets and also orators (examined in the *Gorgias*) and sophists (examined in the *Sophist*). Plato sees the role of these image-makers to be sufficiently important to make them, or so I'll argue, the sole source of the shadows in the Allegory of the Cave, represented by the shadow-making 'puppeteers'.

My argument is divided into three sections. I begin in §5.1 with an account of the lowest level of the Divided Line [Line] and Allegory of the Cave [Cave] in books 6 and 7. I argue that the cognitive level of *eikasia* introduced in the Line and exemplified by the prisoners in the Cave is the *Republic*'s account of judging by appearances—precisely the kind of limited cognition that we find in poetry's audience in book 10. In addition I point to a number of further continuities between books 6 and 7 and book 10, finding in particular a place for the 'image-maker' in the Cave. I conclude that we should see book 10 as an examination of a literal example of what is metaphorically represented by the lowest level of the Cave.

In §5.2 I turn to a problem that is shared by books 6 and 7 and book 10: how can Socrates justify the move, which is made without comment, from sensory appearances like shadows and reflections to evaluative appearances like 'shadows of justice' (517D8–9) or 'images of virtue' (605E5)? I argue, first, that since non-rational cognition is confined to sensory information and imitative appearances exert their power on our non-rational parts, evaluative appearances must themselves be sensory. This raises a problem: since poetry is linguistic, its imitations cannot be literally perceived. The solution I propose follows a suggestion made by Lorenz (2006): that we extend non-rational cognition beyond occurrent acts of perception to include remembered and imagined quasi-perceptual appearances.

I argue that there are very compelling reasons why this is a plausible, even necessary, addition to the *Republic*'s psychology. Second, I draw attention to a feature of appearances in book 10 that, while close to the surface of the text, has often been missed: they derive their evaluative content—content that is meaningful and motivating even for the non-rational parts of the soul—through representations of pleasure or pain.

Finally, in §5.3 I look at book 10's account of the conflict poetry causes in the 'decent man' (*epieikês*) who recognises that a poem's content is shameful but, at the same time, finds it emotionally engaging and enjoyable (605C–606B). This is analogous to the conflict caused by painting examined in the last chapter. I argue that it is strongly analogous: as in the argument from visual illusions, it is a cognitive conflict between opposing beliefs held by different parts of the soul. This time, however, the appearance is evaluative, so it elicits an emotional response from the non-rational parts of the soul. This has the potential to be seriously corrupting: strengthening our lower passions, setting them up to 'rule' in us, and so 'destroy[ing] the rational part' (605B3–6; 606D4–5). By drawing on the conception of psychic 'ruling' developed in books 8 and 9, I argue that the rational part is destroyed by being compelled to believe, in agreement with the lower parts of the soul, that poetic imitations are true. I conclude with a survey of some recent work on similar questions; I argue that Plato's approach to seemingly irrational, belief-discordant desire and behaviour is defensible in its own right.

5.1 THE LINE AND THE CAVE

The connection between the images of the Line and Cave in books 6 and 7 and book 10's criticism of poetry is not often discussed, but the similarities are unmissable. Importantly for our purposes, the move from the Line to the Cave involves exactly the same progression from evaluatively neutral appearances like mirror images to ethical appearances like images of virtue, and they similarly warn of the danger of the latter. I hope to show that the relation between these

discussions is very strong indeed. I argue, first, that Socrates' account of *eikasia* is an account of a kind of cognition that is confined to appearances—precisely the kind of cognition possessed by the non-rational parts of the soul in book 10 (§5.1.1). Second, I argue that the image of the Cave itself anticipates and prepares us for Socrates' criticism of poetry: what we find in book 10 is a literal example of what is metaphorically represented in the lower half of the Cave (§5.1.2).

Consider first the image of the Line. The hierarchy of appearance, original, and Form in book 10 is unmistakably reminiscent of the hierarchy represented by the Line. As in book 10, we find three objects of awareness related by likeness and contrasted by epistemic clarity and degree of reality: 'likenesses' or 'images' (*eikones*) (L1), sensible originals (L2), and Forms (L4).[3] The section I've labelled L2 contains such things as animals, plants, and 'the whole class of manufactured things' (510A6). The lowest level of the line, L1, contains 'shadows, then reflections [or 'semblances'; *phantasmata*] in water and in all close-packed, smooth, and shiny materials, and everything of that sort' (509E1–510A3). A couch will clearly be placed in L2, but whether or not paintings can be placed in L1 depends on the scope of 'everything of that sort'. Shadows and reflections are a small set, but the characteristic they share is quite clear. The word Socrates chooses to describe them is *eikôn*, an image that is a likeness (a cognate of the verb *eoikenai*, to be like), and he describes the relation between the objects of L1 and L2 as one of 'likeness to the thing it is like' (*to homoiôthen pros to hô hômoiôthê*; 510A10). At one point, he even describes it as 'imitation' (510B4). Being a likeness or imitation is precisely the characteristic that defines appearances in book 10. We can be confident, then, that the *eikones* of books 6–7 pick out the same class of thing as *phainomena* in book 10, a connection that is surely hinted at by the inclusion of reflections in both the *phainomena* of book 10 and *eikona* of books 6–7.[4]

[3] The line of course has a further section, L3, corresponding to *dianoia*. This is absent from book 10's hierarchy, but this is surely not surprising. It's hard to see how *dianoia*, which relates in particular to mathematical knowledge, would be relevant to book 10's discussion.

[4] This is explicit in the *Sophist*, where reflections, shadows, the imitator's product, and in addition 'things in dreams' are all objects of the same type (e.g. 266B–D). Halliwell (2002, pp. 57–58) doubts that there is meaningful parallel between the lowest rung of the hierarchies of book 10

So what do we learn from the Line? While it quite neatly reinforces the account of the relationship between image, original, and Form found in book 10, the actual discussion of the lowest section of the Line is extremely brief by comparison. However, it does make one important addition, namely that the four sections of the line correspond to four 'conditions in the soul' (*pathêmata en tê psuchê*). These are, it seems, cognitive states or faculties that are each set over the objects that populate a section of the line (511D6–E4). Corresponding to the top two sections of the Line are two kinds of knowledge: 'understanding' (*noêsis*) at L4 and 'thought' (*dianoia*) at L3; corresponding to the two lower sections are two kinds of belief (*doxa*): 'conviction' (*pistis*) at L2 and *eikasia* at L1.

Again, the Line says little about what *eikasia* is. What we learn about it comes from the allegory of the Cave, on the assumption that the prisoners in the Cave, who see and believe in nothing but shadows, are a metaphorical representation of the cognitive level of *eikasia* (an assumption defended in §5.1.2).

5.1.1 *What is* eikasia*?*

The word *eikasia* can be translated 'liken', 'conjecture' or, drawing on its proximity to *eikôn* and its use in books 6–7, 'image-thinking' or 'imagination'. Despite a variety of views about the details, there is a near-consensus about what *eikasia* is in general terms: being taken in by appearances; failing to recognise the dif-

and books 6–7. He cites, first, a lack of shared vocabulary. This does not strike me as true to any significant extent. Halliwell himself notes their shared use of *phantasmata*, *eidola*, and even *mimêsis*. It is true that *eikôn* is not used in book 10, but with three similar words already in use—*phainomena*, *phantasmata*, and *eidola*—a fourth would add nothing, except possibly confusion. Moreover, *eikôn* is indeed used to describe the product of imitation in book 3 (e.g. 401B2). (We should also remember that there are only twenty or so lines of text that refer to images in books 6–7, and half of them do so only allegorically.) Halliwell's second argument is that 'the physical images [sc. shadows, reflections, 'and everything of that sort'] in the bottom section of the Line stand in a necessarily one-to-one relationship to particular objects, whereas neither *Republic* 10 nor any other Platonic text suggests anything of the kind for the mimetic products of painting or poetry'. Why should the images in the Line necessarily bear a one-to-one relationship with their originals? Halliwell doesn't offer a theory. I find it far more likely that the important characteristic of shadows and reflections is that they are an unfaithful likeness of the the objects of L2, just as painting and poetry are. As far as I can see, this is the only characteristic that does philosophical work.

ference between an appearance and the object it is merely an appearance of.[5] I think this is for the most part correct: it compares well with L2 and L3, both of which also correspond to a form of cognition that fail to make appropriate use of the objects of higher sections of the Line, and it is a good description of the error made by the prisoners in the Cave, mistaking shadows for the objects that cast the shadows.

However, there is one important point on which I disagree with some elucidations of this view of *eikasia*. A number of accounts portray the images or appearances that *eikasia* is directed towards as themselves beliefs, so that *eikasia* becomes an unreflective acceptance of other's opinions.[6] This reading does not find any direct textual support, but is rather motivated by an interpretive problem: the prisoners in the Cave are taken to represent, primarily, the majority's *eikasia* with respect to ethical concepts (a claim defended below), but if this is so what are the 'images' of the just and good that they mistake for their originals? Since it is quite likely that Plato thought most people uncritically fall in line with the prevailing opinion of what is good or just, one answer is that these 'images' are common, unquestioned ethical opinions. Reasonable as this claim sounds by itself, it struggles to fit the actual examples of *eikones* that Socrates gives: shadows, reflections, 'and everything of that kind'. There are two ways one might respond

[5] To take a selection of examples. Sedley: 'basing one's experience on mere images of sensible particulars that are themselves mere images of Forms' (2007, p. 262); Wilberding: 'an inability to distinguish between appearance and original' (2004, p. 129); Fine: a failure to 'systematically discriminate between images and their objects' (2003, p. 102); Hamlyn: 'the state of mind of him who holds that sense-data or appearances are all that there is, who is unaware that or does not acknowledge that there are also material objects' (1958, p. 23); and Paton : 'mistaking the image for the thing, the unreal for the real' (1922). A recent exception is Dominick , who sees *eikasia* as 'the state in which one can view an image as an image' (2010, p. 1).

[6] Two of the more explicit examples are Malcolm (1981) and Wilberbding (2004). Malcolm takes appearances to be false beliefs that are a copies of true beliefs (p. 61). See also Malcolm (1962, p. 44). Wilberbding takes them to be the beliefs of the many that are imitated by politicians (e.g. p. 120). Annas gives an example of a more common version of this understanding of *eikasia* in her reading of the Cave: 'So within the metaphor the shadows are literally shadows and metaphorically are any ordinary opinions about justice, taken over unreflectively and based on acquiescence in the ways things appear rather than an effort to find out what they really are' (1997, p. 155).

to this. First, reflections and shadows could be dismissed as metaphors.[7] This strikes me as unlikely. It would be an unannounced and quite arbitrary move into metaphor. Socrates' description of the lowest level gives every indication of being delivered as fact and it occurs in a passage in which it is very clear what is just a helpful image (the Line, its sections, and their proportions) and what is to be taken literally (the objects and forms of cognition corresponding to the sections). Without any textual hint to the contrary, we should take Socrates at its word. The second option is to attempt to place common ethical opinions within the scope of 'everything of that kind'. This is equally difficult. How can shadows, reflections, and prevailing ethical opinions form a non-arbitrary set, such that we can call them all 'images' without using the word homonymously? Ethical opinions are simply not sufficiently appearance-like to be compared to shadows and reflections.[8]

A more promising alternative is that ethical appearances are not common ethical beliefs but the sort of thing that these beliefs are *about*: an action or type of action, a kind of person, a character in a poem, and so forth. Unlike beliefs, this is something that can plausibly be said to appear a certain way. Agamemnon might plausibly be said to appear just, but it sounds awkward to say that the *belief* that he is just itself appears just (although it might appear true). Likewise, pleasure might appear good, but we would hardly say that the belief that pleasure is good itself appears good. This also fits well with the shadows in the Cave: since they are shadows of people and things, it is natural to assume that the relevant contrast is between how these appear and how they really are.

This way of construing ethical appearances is not without its own difficulties. In particular it anticipates a problem we find in book 10: Socrates moves from

[7] This is the line taken by Wilberbding (2004, p. 131).

[8] This is not to say that common beliefs have no role to play. As Irwin says: 'If we simply accept, without question or criticism, the views we are brought up with or have absorbed from our social environment, we cannot distinguish appearance from reality in this area' (1995, p. 276). This is true, but it does not entail an identity between the views we have uncritically followed and the appearances that we, therefore, uncritically accept. Rather, both one's own unexamined opinions and the social norms that shaped them can each be supported by an uncritical acceptance of how things appear.

the painter's product, which is easily understood to be a kind of image, to the poet's product, 'images of virtue', which might be thought to be 'images' only in some attenuated sense—they do not, at first sight, seem to be a *perceptual* copies. I will argue in §5.2.2 that there is in fact significant continuity between ethical and non-ethical appearances, but for the present the following two advantages over the suggestion that ethical appearances are beliefs will suffice: first, as noted, it fits more naturally with the language of appearances and, to this extent at least, can be placed alongside shadows and reflections; second, it fits the descriptions of ethical appearances we find in the *Republic*, which are not beliefs but rather misleading representations that we find in, for example, paintings or poems.

To take another example, consider pleasure. In chapters 1 and 2 we saw that pleasure and pain are deceptive. Pleasure presents a *prima facie* appearance of being good and pain presents a *prima facie* appearance of being bad. In book 9 of the *Republic* we find a more nuanced and in many ways quite different account of pleasure, but one that nonetheless presents a comparable view of the false appearance of bodily pleasures. Socrates now explicitly rejects the idea that a tyrant's life, which revolves around pleasures of the appetitive part of the soul, is the most pleasant. Rather, the most pleasant life, and the happiest, is the philosophic life of the guardians who enjoy the pure, 'true' pleasures of the rational part of the soul, the pleasures of knowledge and virtue. In contrast, the bodily pleasures that tyrants and the many pursue are, for the most part, 'false' pleasures: they are in truth either the mere absence of pain or pleasures that are mixed with pain. However, false pleasures *appear* to be true pleasures: a relief from pain falsely appears pleasant and pleasure in close proximity to pain falsely appears intensely pleasurable. The language Socrates uses to describe these false pleasure recalls the images in books 6–7 and anticipate book 10: there is a certain 'magic' in the absences of pain such that they become 'semblances' (*phatasmata*; 584A10) of true pleasure; mixed pleasures are mere 'images and shadow-paintings of true pleasures' (586B8) (compare 602C10–D4, where 'shadow-paintings' are among the false appearances that employ a certain 'magic'). Consequently, if someone does

not know what true pleasures are, they will be 'deceived' by false ones (584E7–585A3) and indeed the majority of people pursue these mere appearances of true pleasures as if they were the real thing (586A1–B6). There can be no doubt that such bodily pleasures make an excellent candidate for one of the 'shadows' the prisoners in the Cave are fooled by and, therefore, they are at least a good model to work from when we think of more explicitly ethical appearances, shadows of virtues and vices.

This gives us a clearer idea of the object of *eikasia*. I wish now to turn to the details of *eikasia* itself. It has so far been defined as an inability to distinguish image and original, but there is a passage in book 7 that suggests that a broader definition than this is needed. This is a passage that we looked at in chapter 2, where Socrates describes the attitude of the freed prisoner to what is considered 'wisdom' (*sophias*) in the Cave, a wisdom that is found in a sort of shadow-spotting competition that plays out among the prisoners that are still bound:

> If there had been any honours, praises, or prizes among them for the one who was sharpest at identifying the shadows as they passed by and who best remembered which usually came earlier, which later, and which simultaneously, and who could thus best divine the future, do you think that our man would desire these rewards or envy those among the prisoners who were honoured and held power? (516C8–D4)

In chapter 2 I compared this with the way in which oratory 'guesses (*stochazetai*) at what's pleasant with no consideration for what's best' (465A2); as Socrates described it: 'through routine and knack it merely preserves the memory of what usually happens and that is how it procures pleasures' (501A7–B1). I argued that both amount to an explanation of what thinking is like if one's information is confined to appearances: one takes one's experiences at face value and then applies memories of these experiences, perhaps combined into a general picture of what 'usually' happens, to guide one's future actions.

I suggest that we identify this cognitive activity with *eikasia*. The obvious connection is that it is the cognitive activity of people in a predicament that is generally thought to represent *eikasia*. This does not make it a foregone conclusion, since *eikasia* might be understood to be only the prisoner's bare appre-

hension of the shadows, not their way of thinking about them. But even on the standard understanding of *eikasia*, it seems to be more than just a kind of apprehension. It also involves a certain attitude towards images: mistakenly thinking that images are their originals.[9] As Fine argues, the prisoners are 'at L1 not because of the object they are confronted with, but because of the ways in which they reason about them' (2003, p. 102). She cites the fact that when the freed prisoner returns to the Cave, he sees the images again, but he does not cognitively return to L1, making the errors he was previously prone to, because he now knows them for what they are, mere images.[10] The description of the shadow-spotting activity makes it clear that mistaking image for original is only part of a wider appearance-confined cognitive activity that the prisoners engage in—so why limit *eikasia* only to this part?

A further reason to identify *eikasia* with the prisoners' shadow-spotting is the word itself. As noted earlier, in the *Republic* it is usually translated along the lines of 'image-thinking' or 'imagination', but this is motivated, for the most part, simply by the fact that it reflects Socrates' descriptions of the Line and the Cave.[11] There are very few contemporary uses of *eikasia* to guide translation and not a great deal more in later writing.[12] Nonetheless, it is fairly clear why Plato

[9] Contra Dominick (2010, e.g p. 2 n. 6). He takes *eikasia* to be the *experience* of images, on the grounds that it is defined as one of the four *pathêmata*, a word suggesting a passive state. However, it is not easy to maintain the same claim in relation to the other *pathêmata*. For example, *dianoia* is described as a certain method of investigation and, unlike *eikasia*, it is at least not obvious that it corresponds to any special object that it might be the experience of (510b-511a).

[10] Fine also argues that apprehending images is not *necessary* for being at L1. She points out that the freed prisoner is at first confused when he sees the artefacts that cast the shadows, which she attributes to a continued inability to distinguish images and their objects, despite apprehending the latter.

[11] Grube and Reeve translate 'imaging'; Jowett 'perception of shadows'; Cornford 'imagination'; Shorey translates 'picture-thinking and conjecture', capturing both its apparent meaning in the *Republic* and the connotations I will argue for below.

[12] I find two near-contemporary uses. In Xenophon's *Memorabilia* Socrates describes painting as *eikasia tôn horômenôn*, 'representation of the visible' (3.10.1.5). The second is perhaps closer to its use in the *Republic*: in the spurious Platonic dialogue the *Sisyphus*, it is used to describe a certain bad type of practical deliberation; Gallop translates 'guesswork' (390C7). An example of later uses is Plutarch's *Themistocles*, 29.3, where it has a clear meaning of likeness or comparison and, consonantly, in Demetrius' *On Style*, a manual on rhetoric, it is used in an apparently technical sense for 'simile' (2.80.5).

chose the word *eikasia*. The objects of L1 of the Line are *eikones*, 'images' or 'likenesses', and Plato wishes to name a level of cognition corresponding to these. The natural place to look is *eikazein*. The noun *eikasia* would seem to be related to the verb *eikazein* as, for example, *doxa* is to *dokein*. Now, *eikazein* can mean to represent by a likeness but also, and more appropriately for a cognitive activity, to liken or compare, infer from comparison, conjecture, or guess. These latter senses do an excellent job of capturing the thinking we find in the prisoners' shadow-spotting competition: the prisoners recognise new shadows by comparing them with memories of previous shadows and they make conjectures about what is or will be the case—for example, whether the presence of one shadow implies that another shadow will be along shortly—by drawing on their memories of previous correlations they've seen between certain shadows ('which usually came earlier, which later, and which simultaneously').[13] This is a cognitive level that avails of purely empirical means of investigation, such as comparison and inductive conjecture, and makes no use of more abstract rational enquiry, such as an explanatory account of the shadows and their correlations. As its central skill is the ability to draw speculative judgements from comparisons with previously experienced shadows, the connotations of comparing, conjecturing, and guessing that *eikasia* carries make it an excellent word to describe the prisoners' shadow-spotting. This strongly suggests that it is precisely this cognitive activity that Plato intends *eikasia* to describe and not just the fundamental error underlying it, the failure to distinguish image and object.

One final piece of evidence for this understanding of *eikasia* comes again from a passage that was discussed in chapter 2, a passage that bears strong similarities to both *Gorgias* 501A4–B1 and *Republic* 516C8–D4:

S: If someone were to take away all counting, measuring, and weighing from

[13] For a similar view of the connection between *eikasia* and *eikazein* in the sense of 'conjecture' see Wilberding (2004, p. 130), although he presents a very different view of what the prisoners' conjectures concern. Adam believes that 'the translation "conjecture" is misleading, for conjecture implies conscious doubt or hesitation, and doubt is foreign to *eikasia* in Plato's sense.' (1902). It is true that *eikasia* is largely defined by a failure to doubt specifically that what appears to be the case is the case, but it is not defined by a failure to doubt *as such*.

> the arts, the rest might be said to be worthless. P: Worthless, indeed! S: All we would have left would be conjecture (*eikazein*) and the training of our perception through experience and habit (*empeiria kai tribê*). We would have to rely on the ability to make guesses (*stochastikês*) that many people call art, once it has acquired strength through practice and hard work. (*Phil.* 55E1–56A1)

This is a description of a low form of cognition that bears a number of similarities to the lowest form of cognition in the *Republic*. First, recall that in book 10 'counting, measuring, and weighing' were the calculative abilities the rational part of the soul used to see through false appearances and, conversely, that it was the non-rational part's inability to apply 'counting, measuring, and weighing' that leads them to accept false appearances. Second, *eikazien* in this passage describes something very similar to the kind of thinking that I argue its cognate *eikasia* refers to the *Republic*. It describes a conjecturing or guessing that is based on a habitual 'training of our perception', which presumably aims to improve one's perceptual recognition and ability to predict further perceptual experiences (for example, a musician training their ear to recognise and reproduce the right harmonies, 56A3–7).[14] In other words, it is a form of cognition that functions entirely without calculations, instead drawing purely empirical or inductive conclusions from perception and, we can safely assume, perceptual appearances.

Through a wide range of dialogues, then, Plato presents a more or less consistent view of a low form of cognition that, in the *Republic*, seems to warrant the name *eikasia*: a form of cognition that doesn't require reasoning, is primarily perceptual or empirical, is acquired through habit rather than understanding, and that is incapable of going beyond what appears to be the case. If we reflect on this description, I think it naturally leads to the following proposal: that *eikasia* is a non-rational form of cognition. This claim can be understood in two ways, both of which I think are plausible. First, that it is non-rational in the sense that it is a

[14] The discussion this passage occurs in is also reminiscent of the Divided Line. *Philebus* 55C–59D considers a hierarchy of kinds of arts or sciences in order of 'purity and certainty'. The quoted passage describes the lowest, which includes oratory and the musical arts—*musikê*, which includes both music in our sense and poetry—and the highest is the philosophical art of dialectic, with mathematical arts lying in between. While the parallel is not perfect—it ranks arts rather than objects and forms of cognition (explicitly, at least) and it has a greater variety of distinctions—it is certainly close enough to warrant comparison.

way of acquiring beliefs that doesn't require reasoning or calculation (*logismos*). In this sense it is what throughout this thesis I've called 'judging by appearance'. Second, that it describes the cognition of the non-rational parts of the soul, parts of the soul that, as I argued in chapter 3, are largely defined by the fact that they are insensitive to calculation, but sensitive to appearances. This would expand the abilities of the non-rational parts to include not only occurrent perceptions, but also memories of, and associations between, previous perceptions, as well as the ability to use these to envisage possible future perceptions. But these, or at least comparable abilities, are necessary if the non-rational parts are to be able to perform some of their basic functions. I look at this again in §5.2.1, but for now consider just two simple examples: first, some way of representing the past and future is necessary to have past- or future-related passions and, second, some kind of learning, such as making sensory associations, is necessary to be capable of developing, becoming better or worse. Both of these are abilities the non-rational parts undoubtedly have in the *Republic* and they can not be explained by occurrent perception alone.

In the next section I take a closer look at the lowest level of the Cave. Its principal message is that while ordinary people may go beyond the level of *eikasia* when it comes to shadows and reflections, they fail to do so when it comes to ethical appearances. I will present a case for a connection between the lowest level of the Cave and book 10's account of poetry: what the Cave presents in metaphor, we get a literal example of in book 10. It is obvious that both concern how ordinary people are misled by appearances, but I will argue that there is a stronger connection: the shadow-show that fools the prisoners, orchestrated by certain 'puppeteers', is largely a representation of the effect poetry has on ordinary people. Conversely, the account of the psychological effect of poetry in book 10, how it corrupts the whole soul through its influence on our non-rational parts, can be seen as an example of the dangers of the prisoners' *eikasia*.

5.1.2 *Poetry and the Cave*

The allegory of the Cave is at once one of the most significant passages in the *Republic* and one of the most opaquely metaphorical, with the result that it has inspired a great variety of interpretations. Nonetheless, these interpretations are usually within the confines of three claims, which together might be said to form a loose orthodox view of the allegory. As these claims have been argued for extensively elsewhere, I will offer only a cursory defence of them here.[15] The first claim, (1), is that the four divisions of the Line should match the four stages of the Cave: the chained prisoners who see only shadows and hear only echoed voices are at L1, confined to the level of *eikasia*; the first step in the freed prisoner's ascent, turning to see the statues and other objects that cast their shadows, corresponds to L2, conviction; as he leaves the Cave and views the outside world through their shadows and reflections, he is at L3; and finally when he is able to look directly at everything in the outside world, and ultimately the sun itself, he is at L4. This parallel is justified not only by the good sense it makes of the allegory, but also by a direction Socrates gives: '[t]he whole image, Glaucon, must be fitted together with what we said before. The visible realm should be likened to the prison dwelling … [and] the upward journey of the soul to the intelligible realm' (517A8–B6). Claim (2) concerns what Socrates means when he says that the prisoners in the Cave are 'like us' (515A5): while Socrates and Glaucon may not be the most typical representatives, the 'us' is generally taken to be ordinary citizens of the polis, perhaps extending to those within the *kallipolis* at an early stage in their education (more discussion of this claim below).

The third point of agreement can be seen as a solution to an apparent consequence of accepting the previous two, one that particularly troubled early interpreters of the allegory. If the prisoners represent the cognitive level of ordinary people and this level is to be compared with L1 of the Line, then it seems we reach the unlikely conclusion that most people mistake shadows and reflections for the material objects that they are like. This conclusion can be avoided if we

[15] Their earliest defence, as a single, unified reading, was given by Malcolm (1962).

pay attention to the intended scope of the allegory. Socrates tells us that what the allegory represents is 'the effect of education and of the lack of it on our nature' (514A1–2). Recognising ordinary material objects is surely not within the scope of what one might either need or lack an education in. Thus, claim (3): since the aims of the educational program of the *Republic* are primarily ethical, the prisoners' beliefs about shadows primarily represent ordinary people's ethical beliefs—just as the prisoners mistake shadows for real objects, most people mistake what appears just for what is just. This finds support in Socrates' description of the prisoners as fighting over 'shadows of justice' (517D8–9) and merely apparent goods (520C7–D1) and most decisively, although often overlooked, Socrates tells us that the freed prisoner's ethical knowledge is *all* that he needs to recognise the shadows for what they are: 'because you've seen the truth about fine, just, and good things, you'll know each image for what it is and also that of which it is the image' (520C4–6).

The interpretation I will offer is confined to the first half of the allegory: the prisoners and the parade of artefacts that cast the shadows that fool them. My aim is to give a compelling defence of a claim that has been made more or less in passing by a number of authors, namely that the puppeteers principally—although not exclusively—represent the poets.

Socrates' decription of the life of the prisoners is as follows. They are doubly shackled, at their neck and legs, so that they can only look towards the wall in front of them. Behind them is a fire, and between them and the fire is a wall 'like the screen in front of puppeteers above which they show their puppets' (514B5–7).[16] Below and behind this wall are people carrying all kinds (*pantodapa*) of artifacts—including 'statues of people and other animals, made out of stone, wood, and all sorts of material' (514C1–515A1)—and they hold them above the wall so that their shadows are cast onto the wall in front of the prisoners. Further-

[16] The words translated 'puppeteers' and 'puppets' are *thaumatopoiois* and *ta thaumata*; a more literal translation would 'wonder-worker' and 'wonders'. Puppetry is suggested both by its good fit with what Socrates describes and the use of *thauma* in the *Laws* to compare us to 'puppets' of the gods, controlled by strings (644D–645B).

more, some of the carriers talk, and their voices echo so that the prisoners hear them as if they come from the shadows in front of them.

The prisoners' predicament leads them to make two errors: first, when they talk to each other, they assume the words they use apply to the shadows passing before them rather than the artefacts that cast the shadows; second, when they hear the echo of the carriers' voices, they assume they are the voices of the shadows of the statues (515B4–9).

There are, then, no less than three sets of 'people' in the Cave: the prisoners, the puppeteers,[17] and the statues of people. All three are related by the auditory illusion created by the puppet show: the prisoners believe the echoed voices come from the statues' shadows, when in fact they come from the puppeteers. There is, then, agency behind the deception the prisoners suffer. The play of shadows and the words that accompany them are deliberately orchestrated by certain puppeteers. Who, then, do the puppeteers represent? And who are the 'speaking' statues they carry?

Since according to (2) and (3) the prisoners' predicament represents the low ethical education (or miseducation) of ordinary citizens, a natural answer to the first question is that the puppeteers represent one or all of groups that Plato commonly accuses of misguiding people's ethical beliefs: poets, orators, and sophists.[18] Each of these professions is an example of what Socrates calls an *eidôlou dêmiourgos*, a maker of images, and making images is precisely what the puppeteers do when they cast shadows of their artefacts (artefacts that are themselves copies of the real objects outside the Cave—the shadows, then, are at three removes from the truth). This is a plausible suggestion and, with one important qualification, I believe it is correct. The qualification is needed to address a problem raised by Wilberding (2004). He argues that the shadow-spotting of the prisoners is a competition not between ordinary citizens but between those who are vying for power: the sharpest shadow-spotters are rewarded with 'hon-

[17] Strictly speaking, they are said to be 'like' puppeteers, but their activity is sufficiently close to puppeteering to justify the name

[18] See, for example, Burnyeat (1999, pp. 140-141) and Malcolm (1962, p. 44)

ours, praises, and prizes' and are those who hold 'power' (516C8–D4); it is compared to the sort of competition that plays out 'in the courts and elsewhere' (517D8); and most cities are governed by those who 'fight over shadows and struggle against one another in order to rule, as if that were a great good' (520C7–D1). Wilberding takes the political focus of the shadow-spotting to be a reason to reject (2), the idea that 'like us' refers to ordinary citizens. Rather, he argues, the *prisoners* represent the orators (and also artists and poets; p. 120), while the puppeteers are the ordinary citizens whose beliefs and desires the orators must try to guess if they are to persuade them. While I cannot give Wilberding's reading its full due here, I will argue that we can accept his claim that orators, poets, and sophists are among the prisoners without rejecting either claim (2) or the claim that they are, in addition, to be identified with the puppeteers, in a qualified sense.

The political nature of shadow-spotting does indeed suggest that orators are among the prisoners, placing them at L1 rather than L2. This makes good sense: Plato never portrays orators as being ethically wise in comparison to ordinary people. As we saw in chapter 2, orators are as fooled as the majority of people by what merely appears just. In book 10 Socrates makes the same claim about the poets: they not only make mere appearances for their audiences to enjoy but they themselves are 'third from the truth', having no knowledge beyond the very appearances they deceive others with. We have to wait until the *Sophist* for the same claim to be explicitly made about sophists, but there is little reason to think Plato held them in higher regard in the *Republic*. If this is right, then poets, orators, and sophists should be placed among the prisoners at L1, not among the puppeteers at L2, since the latter would suggest that they have achieved a higher level of ethical understanding than the masses—after all, the puppeteers not only see the 'statues of justice' but are controlling them. For precisely the same reasons I do not believe we should identify the majority with the puppeteers, placing them at the L2.[19] I do not think the political features of the shadow-spotting

[19] Wilberding is of course aware of this response, but he argues that orators, poets, and soph-

are, as Wilberding claims, a reason to exclude ordinary citizens. It is true that the courts are where orators ply their trade, but there would also have been up to 500 ordinary citizens in attendance, the jury, each of whom would have to consider 'shadows of justice' before they cast their vote. We should not forget that in the direct democracy of Athens—which is likely the society Socrates has in mind—every citizen is a politician: every citizen was entitled to membership of its institutionally most powerful body, the Assembly, and the majority would have at some point in their life taken their turn in the law courts or the Council. Debates over 'shadows of justice', then, are part of the everyday life of a citizen. It is true that only the exceptional shadow-spotters—no doubt the likes of the orators—would have received 'honours, praises, and prizes', but this still leaves room for the unexceptional among the prisoners.

These considerations suggest that both the ordinary citizens and the poets, orators, and sophists are among the prisoners—who, then, are the puppeteers? I think there are two good reasons not to identify them with any group of people *per se* but rather with the *professions* of poetry, oratory, and sophistry. Consider first the strange life the puppeteers lead. The puppeteers devote their lives entirely to the carrying back and forth of artefacts, delivering a constant shadow-show for the prisoners. Nothing in the allegory explains why they do this. Psychologically, their behaviour is opaque: there are no beliefs or desires that we could attribute to them that might make their behaviour psychologically intelligible (unlike the prisoners, whose behaviour is explained by their mistaken belief in the reality of the shadows; their shackles; and the 'prizes and honours' bestowed by their peers). They are also oddly isolated from the prisoners and the affairs that play

ists have lower level of education than the majority because they reject the idea that there is an objective morality, seeing morality as simply a set of conventions that the majority follow (pp. 132-135). On the other hand, he believes that the majority do believe in an objective morality and hold ethical beliefs that are 'by and large true' (p. 133), justifying their position at L2 (which he equates with true belief). The first claim could certainly be challenged—e.g. Callicles, and arguably Thrasymachus, believes in an objective morality *despite* the fact that he thinks that the many's view of justice is mere convention—but I find the second the most difficult to accept. To take just one example, consider Socrates' long discussion (377a-392c) of the mistaken content of the poetry that informs the ordinary child's ethical views, which are 'for the most part opposite to the ones we think they should hold when they grow up' (377b7-8).

out among them. They only interact with them from a distance and indirectly, through the shadows and echoes they create. Even when one prisoner is freed, they appear to be unaware or unmoved by the drama that unfolds and, similarly, the prisoner ignores the puppeteers both when he is first freed and when he returns to govern the Cave (nor is there any suggestion that they have the potential, like the other prisoners, of partaking in the journey out of the Cave—if they are real people, they are incurably confined to L2). In short, the puppeteers do not seem like real people at all. This is because, I suggest, they represent something more abstract than a particular group of people's fate, namely the image-making industries that sustain and spread misbelief.

A second reason to identify the puppeteers with the image-making professions is that it is in fact a very accurate way of representing the image-maker's role in society. As we noted, while a poet, orator, or sophist presents images of justice that persuade his fellow men, he is himself equally persuaded by these images—like the painter, he makes appearances of appearances. Consequently, poets, orators, and sophists do indeed have a double role in society: they are both craftsmen and consumers of images, both puppeteer and prisoner. However, where their state of education is concerned, their role as consumers of images is what counts. In this sense, they themselves are victims of the image-making professions: each will have been raised on the same poetry as the ordinary citizen, been persuaded by speeches in the assembly, and many will have mistaken the teachings of the sophists for wisdom. In other words, with respect to their ethical education, they are ordinary citizens. Thus, the allegory will represent 'the effect of education and of the lack of it' most accurately if it does not separate poets, orators, and sophists from their fellow citizens, but nonetheless recognises the influence their professions have on education, by embodying them in the puppeteers.

So far, then, we have seen two reasons to think that the puppeteers represent all three of the professions of poetry, oratory, and sophistry: these are the professions that Plato commonly blames for the public's misguided ethical views and

they are each a type of image-making, which is precisely what the puppeteers do. I wish now to point to a number of ways in which poetry in particular is represented in the allegory.

The clearest sign of poetry in the allegory is simply the fact that it looks very like a dramatic performance. In the *Laws* Socrates describes puppetry as a low form of theatre—a form of theatre that would be the favourite of 'the little ones' (*ta smikra*) if it were entered into a theatre competition (658C2–3)—and the image in the Cave can easily be seen as a piece of theatre: the puppeteers put on a show, complete with props and talking characters, to a seated audience. Consider in particular the statues of people whose shadows the prisoners believe to be talking. This element of the allegory is usually neglected, but this is surely a mistake. Some elements of an allegory are necessary only to aid a coherent and plausible narrative, while having no metaphorical significance of their own. For example, the wall the puppeteers walk behind is included only to facilitate the mechanics of the shadow-show and is unlikely to have any independent meaning. The apparently talking shadows are clearly not one of these auxiliary elements. The prisoners see each other's shadows and hear the echo of each other's voices. This is the allegory's representation of ordinary people and ordinary speaking. The puppeteers' echoed voices represent something else and it is something important enough for Socrates to make this one of only two misapprehensions the prisoners suffer. To consider only the shadows and not the echoed voices is to consider only half of the prisoners' predicament.

If the shadow-show is compared to a poetic performance, the echoed voices have a clear function: the statues' shadows represent the characters of a poem and the echoes represent the voices of these characters. Like fictional characters, the shadows of people are merely apparent speakers; speakers that the audience listens to, but doesn't converse with; and speakers that form part of a more general performance, orchestrated by puppeteers and enjoyed by an audience. In book 3 Socrates draws a distinction between 'simple narrative' and 'narrative through imitation', which he illustrates with the *Illiad*:

> [At first] the poet himself is speaking and doesn't attempt to get us to think that the speaker is someone other than himself. After this, however, he speaks as if he were Chryses and tries as far as possible to make us think that the speaker isn't Homer but the priest himself, an old man. (393A6–B2)

The statues are an example of 'narrative through imitation': the puppeteers speak as if they are the statues, leading the prisoners to think that it is the shadows of these statues that they are hearing. This gives the allegory an appropriate scope that it would lack—or at least not clearly represent—if it were confined to shadows. In book 10 Socrates makes a distinction between two mediums of appearances, the visual medium of the painter and the auditory medium of the poet. The visual will be used in theatre too, but it's the speeches of the characters that carry the main content of the poem, allowing it to present compelling 'images of virtue' to its audiences. Thus, by using both shadows and echoes the Cave is able to capture both mediums through which the craftsman of images will produce appearances of what's good or just. Arguably, it is the echoes that are the more important of the two, since words will allow more complex and compelling appearances, especially when it comes to more abstract concepts like virtue (thus poetry is more dangerous than painting). We can also quite neatly, if more speculatively, use this reading to explain the role of all three 'speakers' involved in the echo: the puppeteers who speak represent the poets, the statues they carry represent the actors or rhapsodes (the poet's puppets), and the statues' shadows that are mistaken for the speakers represent the characters the actors are playing.[20] If this is correct, the prisoners' error is to mistake the characters for real people. This may seem implausible—although no doubt some people did believe the contents of poems were true, like Euthyphro (5E ff.)—but this is a further point that is clarified in book 10, where we get a quite detailed account of the kind of errors involved in enjoying a dramatic performance. As we will see, this involves a 'belief' in the performance's content that might be described as suspension of

[20] Note that all three of these—writer, actor, and character—are discussed in detail in the *Republic*: the discussion of imitation in book 3 focuses on the actor and the harm of impersonating bad characters (395C ff.); as we have seen, book 10 aims to characterise the poet himself; and the characters in poetry are discussed in both books 3 and 10, but are given their most detailed examination in book 10, as we will see in the next section.

disbelief: being carried along by the drama and empathising with its characters' fates. This in turn is explained by a literal belief, but a belief of a non-rational part of the soul.

On reflection, we should not be surprised that poetry is represented so prominently in the Cave. The allegory represents 'the effect of education and of the lack of it' and the *Republic* leaves no doubt that Plato considered poetry the most important source of miseducation, worthy of two lengthy examinations. In book 10 Socrates mentions those who give Homer the title of 'the poet who educated Greece' and who believe that one should model one's 'whole life in accordance with his teachings' (606E1–5). This does not appear to be an unusual view. While oratory and sophistry have quite specific spheres of influence—for oratory, the assembly and courts and, for sophists, those willing to go or send their children to them—poetry was ubiquitous in Greek culture. Poetry shapes citizens' characters at every stage of the their life, beginning with the fairy tales told by nurses and mothers (377B11–C5). It was also a central subject of the Greek educational curriculum, shaping every citizen's moral outlook at a time when their soul is 'most malleable and takes on any pattern one wishes to impress on it' (377B2–3). In a description of a typical Greek child's education, Protagoras—of the dialogue—tells us that: 'they are given works of good poets to read at their desks and have to learn them by heart, works that contain numerous exhortations, many passages describing in glowing terms good men of old, so that the child is inspired to imitate them and become like them' (325E4–326A4). We are explicitly told that the prisoners are bound 'since childhood' (514A5)—a child would have little contact with orators or sophists, but would be steeped in poetry, and it is likely that much of the moral dye will have become indelible by the time he first encounters an orator or a sophist.[21] The influence of poetry continues solidly into adulthood.

[21] Equally, Socrates compares the shadow-spotting to the sort of political activities that would take place in the courts, where there is unlikely to be much poetry. My claim is not that the shadow-show represents poetry exclusively. Part of the appeal of an allegorical form of argument is the breadth of its explanatory scope. The imagery of an allegory can represent more than one thing, and there are no shortage of examples in the Cave. It would be unwise, therefore, to try to confine the allegory to an interpretation that is more specific than necessary. My claim is only

The lion's share of Greek popular culture was poetry, both the hugely influential epics of Homer and its popular theatre culture. The Cave, then, rightly represents poetry as a major influence on ordinary citizens' opinions. However, its explanation of its harm, of how it misleads, is vague: it presents mere shadows of what's real. We have to wait until book 10 for substance to be added to this claim.

5.2 'SHADOWS OF JUSTICE' AND 'IMAGES OF VIRTUE'

To illustrate what appearances are, Socrates points to straightforwardly *sensory* appearances: shadows, reflections, paintings, and optical illusions. These are intended to be paradigmatic: Socrates appeals only to sensory appearances when defining both the lowest section of the Line and 'what imitation in general is'. Yet the appearances that most concern him are ethical or evaluative appearances, which might seem to be a quite different sort of thing—how can something, literally, *look* good or just? There is no sign that this is a problem that Plato recognised; both books 6–7 and book 10 move freely between sensory and evaluative appearances and take any claim that applies to one to *ipso facto* apply to the other: from shadows and reflections in the Line to 'shadows of justice' in the Cave and from images of couches in painting to 'images of virtue' in poetry.

There are two related questions that need to be answered: how are evaluative appearances related to *sensory* appearances? And how can appearance have *evaluative* content? I offer an answer the first question in §5.2.1 and to the second in §5.2.2.

5.2.1 *Painting to poetry: images and words*

I aim to show that imitative painting and imitative poetry share substantially the same kind of appearance, and that they do so precisely because they are imitative. This is, broadly speaking, sensory appearances. This conclusion is entailed by

that a majority of the imagery of the allegory suggests poetry and that this should be taken to be proportionate to the harm Plato thought poetry causes.

the following claims, both of which we have already seen substantial evidence for: first, poetic appearances exert their power on the non-rational parts of the soul and, second, the cognitive abilities of the non-rational parts of the soul are limited to the sensory.

However, a *prima facie* reason to think that poetry does not, at least as a rule, involve sensory appearances is that the majority of it is presented through language. While theatre does involve visible action, and audible additions like music and thunder, the majority of poetry is in the form of 'words and phrases' (601A5). One doesn't see but either reads or hears a rhapsode recite Homer's *Illiad*. Thus, poetry's content is at least not literally perceived. (Socrates does characterise poetry as an auditory medium in contrast to visual painting (603b6-7), but he presumably does so loosely, in reference to the fact that poetry is, for the most part, presented in language and language is, for the most part, heard.)

There is one excellent reason to think that the story must be more complex than this. In chapter 3, we saw that poetry affects the non-rational parts of our soul and does so 'by making images' (605B7–C4). But surely the non-rational parts of the soul don't understand language. Our non-rational parts are capable of *some* form of low-level understanding. Crucially, their ability to apprehend content-rich perceptions endows them with belief-like representations of the world. Moreover, as previously mentioned, the *Republic* hints at other cognitive abilities too, most notably in relation to communication and agreement between the parts of the soul: for example, in moderation all three parts of the soul 'believe in common that the rational part should rule' (442C11–D1), suggesting they are capable of agreement that is based on more than just non-rational conditioning, and Socrates criticises the oligarch for disciplining his wayward appetites 'not by persuading them that it's better not to act on them or taming them with arguments, but by compulsion and fear' (554D1–2; see also 586D–E), suggesting that the appetitive part is best controlled by some form of rational persuasion. The *Republic*, unfortunately, never explicitly reveals the process by which such communication is achieved. Nonetheless, it seems highly unlikely that the ra-

tional part speaks to the non-rational part in any literal sense; certainly it does not persuade the appetites that 'it's better not to act' by presenting them with the arguments we find in the *Republic*.

Moreover, in many imitative poems characters will make speeches that involve reasoning, engaging in deliberation or presenting arguments for their actions. These more rationally demanding aspects of a poem, at the very least the more complex of them, will not be accessible to the non-rational parts of the soul, whether we have direct or indirect access in mind. Here we need to draw a distinction: while poems can and often do present arguments or deliberations, they do not do so by virtue of being imitative. A character could present a philosophical theory or even mathematical proof, if the poet so wished, but this would not make such content imitative (although it is presented *in* an imitation). Imitations have been defined as images at three removes from the truth, but an argument will suffer no degradation if it is spoken by a character who is himself an imitation. If a character presents good reasons to be, say, courageous, these are still good reasons—they are not images of good reasons and they are certainly not at three removes from the truth. What Socrates takes to be the defining subject of imitative poetry—in fact, what he says is its only subject—is people in *action*, expressing evaluations of these actions and experiencing pleasure or pain through them (603c4–9; I examine this and related passages below). This is not to say that poetry cannot or does not venture beyond this remit, but to the extent that it does, it is going beyond pure imitation. Indeed, the *Republic* is itself a work that is written in what was defined in book 3 as narrative through imitation. If what I've said here is correct, it is nonetheless the case that very little of its content could be described as imitative.

But how does even the purely imitative content of a poem exert its power on the non-rational parts of the soul? While imitation concerns actions, doing well and badly, and experiencing pleasure and pain (content which has a more emotional, less rational impact of a kind that we would expect the non-rational parts to be responsive to) the original problem still remains: if the non-rational

parts lack the cognitive resources needed to grasp language—if they cannot, as it were, read over one's shoulder—then how can they understand *any* part of a poem? The solution I wish to propose is that poetry's imitative content affects the non-rational parts of the soul in a way very similar to the effect of painting and other visual appearances: both represent only the sensory aspects of what they imitate. The 'images' that poetry creates are not its words, but what its words represent. While a painter uses paint to make images that can be perceived, a poet uses words to convey a vivid image of his story in the minds of his listeners. My proposal is that it is this image that the non-rational parts of the soul are responsive to, rather than any intellectual grasp of the words that are used to convey them.

As in the *Gorgias*, those who are taken in by imitation do not only lack knowledge of what the poet represents—which might lead one to a humble profession of ignorance—but form judgements about what they don't know with appeal to the wrong sorts of reasons, namely how something appears. Thus, while a painter knows nothing about cobblery, he 'can make what seems to be a cobbler to those who know as little about it as he does and who judge by colours and shapes' (*ek tôn krômatôn de kai skêmatôn theôrousin*; 601E7–A2). Although the painter will paint only a visual stereotype of what he thinks a cobbler plying his trade looks like, with no knowledge of the technical details of the craft of cobblery, it will be a compelling representation to those who judge on a similarly superficial level, judging only by 'colours and shapes'. [22]

Socrates turns next to an analogous claim about the poet:

> And in the same way, I think, we'll say that a poet uses words and phrases to paint coloured pictures of each of the crafts (*chrômata atta hekastôn tôn tech-*

[22] I take it that the ignorant viewer's error—at least the error of his rational part—is not the unlikely one of mistaking the painting for a real cobbler, but rather mistaking it for a realistic representation of what a cobbler is like and assuming, therefore, that the painter has knowledge of the cobbler's craft. Cf. 598B–D. This is suggested from the analogous mistake attributed to poetry's audience a few lines later, which I consider next. That this is in general the deception perpetrated by painting and poetry is argued extensively by Belfiore (1983) and Moss (2007b). We should note, however, that a person's *non*-rational parts will indeed, as I argued in chapter 4, be unable to distinguish a real and a painted cobbler.

> *nôn tois onomasi kai hrêmasin epichrômatizien*)—not understanding, but imitating them—so that it seems to others as ignorant as he, who judge by words (*ek tôn logôn theôrousi*), that one speaking in meter, rhythm, and harmony speaks extremely well, whether it is about cobblery, generalship, or anything else—so powerful is the natural charms of these things. Yet if you strip a poet's works of their musical colourings (*mousikês krômatôn*) and take them by themselves [they cease to be compelling]. (601A4–B7; cf. *Gorgias* 502C5–D2)

This is a difficult passage. Those who 'judge by words' are clearly meant to be analogous to those who 'judge by colours and shapes', but it is hard to see where this analogy lies. Moreover, while the poet's words may be the wrong ones, judging by words is surely not by itself a bad way of judging—after all, Socrates' own argument must equally be judged by words. However, the passage gives us two clarifications that indicate that Socrates means a particular kind of judging by words. The first is that 'meter, rhythm, and harmony' add a certain natural charm to his words. All Greek poetry was set to metre and much of it, in theatre, was sung. In book 3 we were told that meter, rhythm, and harmony can themselves imitate virtues and vices, such as courage and softness, and their effect can be powerful: they 'permeate the inner part of the soul more than anything else' (401D6–7), and thereby promote these virtues or vices in their listener. Style can also be deceptive: while a correct use of rhythm or harmony is said to fit the true character of the words they accompany, an incorrect use will dress words in a false character (399E8 ff.).[23] Thus, the kind of judging by words that Socrates has in mind is not simply a cool appraisal of the content of the words, but rather involves allowing one's judgement to be affected by the *sound* of the words, the stylistic way in which they are presented. We might imagine, for example, a poet presenting a disorderly, cowardly general in 'musical colourings' that evoke a sense of order and courage.

There is, then, a definite auditory aspect to poetry that can plausibly affect the non-rational parts of the soul. However, this serves only to augment or amplify a poem's representational, imitative content. To understand the latter we

[23] E.g. if 'rhythm and mode should conform to the words and not vice versa, then good rhythm follows fine words and is similar to them, while bad rhythm follows the opposite kind of words' (400D1–4). I take it that the 'vice versa' is bad words fitted to good rhythm.

need to look at a second distinctive feature of the poet's words, namely what he aims to convey with them. For example, the poet uses words and phrases not to give an *account* of each of the crafts, but to 'paint coloured pictures' of them. This is of course a metaphor that facilitates Socrates' comparison with the painter: the poet is like the painter, but he paints with words and phrases.[24] But painting is more than just a metaphor for poetry, as is clear from the fact that Socrates *defines* imitation through the example of painting, defining both the kind of activity it is and the nature of its product, appearances or images. Socrates clearly thinks paintings and other exclusively sensory examples of images suffice to give us an understanding of the images that the poet makes. Indeed, many of Socrates' most important arguments apply only to sensory appearances, while nonetheless clearly being intended to apply to poetry. The most significant examples are the arguments that were, respectively, the subjects of chapters 3 and 4: 602C–603A, which uses the lower part's reliance on sight to establish that it is the part affected by imitation, and 597E–598D, which uses the sensory phenomenon of perspectival variation to establish that the imitator is himself third from the truth. This should encourage us to explore how the images a poet makes can be relevantly similar to the painter's sensory imitations.

On examination, we can see that the metaphor of painting with words captures something unique about the way that the imitative poet uses words, by virtue of being imitative. While the character of Socrates, who is Plato's imitative creation, is equally using words and phrases, he is using them in a different way—not to paint a picture of the imitative poet, but to give an account of him. An important difference between an account and an imitation of, for example, a general is that only the imitation, whether in painting or poetry, tries to create the sense that there is a living general before us: where an account will describe

[24] Compare 377E1–3: poetry that misrepresents the truth 'badly creates images with words (*eikazê tis kakôs tô logô*) of what the gods and hero's are like, like a painter whose painting is not at all like the things he's trying to paint'. Thayer (1975) discusses the comparison between painting and poetry at length in his account of Plato's treatment of the highly visually-descriptive poet Simonides. Simonides himself wrote that 'painting is silently spoken poetry; poetry, painting that speaks'. See also Thayler (1977) for a similar but more general discussion of images in Plato.

the principles behind the craft of generalship, an imitation will conjure Agamemnon, leading the Greeks into Troy.

Recall that in book 3 Socrates distinguished imitative from first-person narrative in the *Iliad* by the fact that when imitating Homer 'speaks as if he were Chryses and tries as far as possible to make us think that the speaker isn't Homer but the priest himself, an old man' (393A8–B2). Homer is not, of course, trying to make us literally believe Chryses is speaking, but to *imagine* that he is. The poetic imitator appeals to our imagination: he uses words and phrases to conjure a compelling image of his characters and stories, inviting us to 'picture' the scenes he creates as if we were really encountering them. The visual metaphor of 'painting a coloured picture' is fitting: while a first-person narrative can describe aspects of the crafts that don't meet the eye—such as their function—imitation aims to get us to imagine that we are seeing and hearing a craftsman and, therefore, will convey only those sensible aspects of the craft that are needed to make a plausible appearance of a craftsman. Importantly, the poet and the painter's subject is exactly the same (within, of course, the limits of each of their mediums). The poet imitates only appearances, so just like the painter, if he represents a bed, he will imitate the appearance rather than the nature of a bed. While he will use the more powerful tool of words, the fact that he is an imitator means that he will use them to describe only how a bed looks. What he imitates, then, are sensory appearances and, consequently, what he presents imitations of are sensory appearances, but a verbal description of them rather than a painting of them. Thus, the kind of judging by words that Socrates has in mind is closely analogous to judging by colours and shapes. Both are ways of judging by sensible appearances, but those who judge the imitative poet by his words are judging the sensible image that his words describe.[25]

There are a number of reasons to think this is a plausible account of poetry's

[25] This might be Shorey's motivation for focusing on the perceptual sense of *theôrein* in his translation of 601A4–B7: '[The imitative poet] lays on with words and phrases the colors of the several arts in such fashion that others equally ignorant, who *see* things only through words (*ek tôn logôn theôrousi*), will deem his words most excellent'; my emphasis.

psychological effect. First and foremost, it clarifies a number of otherwise puzzling aspects of book 10: it avoids the problematic suggestion that the non-rational parts grasp linguistic content; it strengthens the connection between painting and poetry, upon which much of Socrates' argument depends; and it gives us a consistent account, throughout book 10, of what 'appearances' or 'images' are. It also strikes me as true: when we listen to or read a narrative we do conjure a mental image of what it describes. Arguably, picturing a story is an essential part of following it: if an author describes a character's appearance, can we grasp it without picturing it? If we recall a scene from a novel, do we remember the words—a fairly remarkable feat of memory—or imagine the scene the words described?

This solution does require more from the non-rational parts' cognitive abilities than grasping occurrent perceptions, bringing it closer to the territory of *eikasia*, but this is a plausible addition and a more fitting addition than the alternative, which is to grant them linguistic understanding. Specifically, it requires the ability to grasp quasi-perceptual mental images, which are the kind of images that the poet's words conjure. As I suggested in §5.1.1, this and related abilities are arguably a necessary addition. Perception alone would not enable the appetitive or spirited parts to do even some of the simpler tasks expected of them. It would not allow them, for example, to change and develop, as Plato clearly believes they can: they can be habituated through music and gymnastics; they can develop associative attachments like the appetitive part's desire for money (581E5–A1); and as we see in the examination of vicious characters in books 8–9, they can come to acquire entirely new goals (the oligarch's spirited part, for example, comes to value nothing but wealth; 553D1–6). Occurrent perception also fails to explain how the parts of the soul communicate and reach agreement: if they are intransigently responsive to occurrent perceptions, it seems impossible for them to obey the 'declarations of reason' (441C1–3), be 'persuaded … with arguments' (554D1–2), or through 'knowledge and argument' pursue only 'the pleasures that reason approves' (586D5–7). Most fundamentally, since we can only perceive the here and now, it fails to explain how we can form non-rational

passions that relate to the past or future. The most prominent emotion in book 10 is related to a past event: grief that involves 'recollections (*anamnêseis*) of suffering' (604D8). The ability to be moved by both memories and mental images of future events and to form associations between sensory appearances can plausibly fill these gaps. Indeed, since these abilities can be conceived of as extensions of perception—involving sensory appearances that are phenomenologically similar to perceptions—they would seem to be the most parsimonious addition that reaches the minimum requirement of allowing the non-rational parts to function as they are described in the *Republic*.

There is a further reason to be confident about this reading. While these cognitive abilities are implicit in the *Republic*, they are an explicit part of the psychological theory of the *Timaeus* and *Philebus*. In these dialogues non-rational passions are responsive to, and directed by, a range of perceptual experiences: occurrent perceptions, quasi-perceptual mental images and, in the *Philebus*, memory defined as a 'preservation of perception' (*sôtêria aisthêseôs*; 34A10–11). Both dialogues explore how complex, rational information—not dissimilar from the linguistic content of a poem—can be translated into sensory appearances that our non-rational passions can comprehend. The relevant passages are complex and difficult, but they are worth summarising.[26]

In the *Timaeus*' account of the soul, each part is assigned a place in the body: the appetitive part below the midriff; the spirited between midriff and neck; and the rational in the head. The former two, we are told, are 'fused with unreasoning sense-perception' (69D4–5). So placed in the body, Timaeus ascribes a psychophysiological function to the body's organs and the liver, we are told, was made by the gods to solve a problem concerning the appetitive part: 'they knew this part of the soul would not understand *logos* … but would be much more enticed by images and semblances (*eidôlôn kai phantasmatôn*; 71A3–7). The liver was

[26] The summarises I give, condensed as they are, gloss over many interpretive difficulties. For detailed interpretations of the relevant passages with which I am largely sympathetic see Moss (2012) and Lorenz (2006, Ch. 7). Both Moss and Lorenz also argue that these aspects of the *Timaeus* and *Philebus*' psychologies are compatible with, and amend some lacunas in, the psychology of the *Republic*.

made reflective so that the 'the force of thoughts sent down from the intellect might be stamped upon it as upon a mirror' (71B3–5) and these 'painted semblances' (*phatasmata apozôgraphoi*; 71C3–4), accompanied by a pleasant sweetness or painful bitterness, can encourage or discourage our appetitive passions, giving them either 'gentle inspiration' (71C4) or 'threatening them with severe commands' (71B6–7). We are not told what the content of these images are, but a reasonable guess is that they represent pleasant or painful outcomes of available actions. For example, the rational part might threaten or frighten (71B5–7) the appetitive part out if it's desire for a harmful pleasure by vividly picturing the painful illness that it would lead to.[27]

The passage in the *Philebus* uses a metaphor that is even more relevant to the problem posed by poetic appearances. Socrates compares our soul to an illustrated book, made by both a painter and a scribe. Forming a judgement is likened to a scribe writing words in our soul and if these judgements involve sensible information, the painter's task is to add a mental picture of their content:

> [The painter] follows the scribe and paints images (*eikonas*) of the words in the soul … [whenever] a person takes his judgements and assertions from sight or any other perception and then views the images he has formed inside himself, corresponding to those words and judgements. (39A1–C1)

Why do we need a painter? Socrates considers the anticipatory pleasure that directs the passion of hope. When there are 'assertions in each of us that we call hopes' (40A6–7) there are in addition 'painted appearances' (*phantasmata ezôgraphêmena*; 40A9) of these assertions. If someone hopes to have a lot of money, he might imagine himself 'in possession of an enormous amount of gold and a lot of pleasures as a result. And in addition, he also sees himself, in this inner paint-

[27] Bobonich argues that 'what the image signifies is not accessible to the Appetitive part and thus is not, as such, part of the explanation of why it does what it does' (2004, p. 318). Rather, he claims, the explanation is solely the pleasure and pain they cause. On this reading, however, it is very hard to see what purpose the images have: if the appetitive part cannot respond to their content, why would they be needed to produce the pleasure or pain in the liver? It would seem mere epiphenomena of the disturbance in the liver. Bobonich also suggests that perhaps the images are evaluative but not representational—'patches of gloomy colours'—but this fails to capture *eidola* and *apozôgraphoi*, which indicate likeness and representation, or the fact that the images can frighten the appetitive part (71B5).

ing (*enezôgraphêmenon*) of himself, beside himself with delight' (604A10–12). A similar story can be told about our other passions: 'the same account holds in the case of fear, anger, and everything of that sort' (604E2–3). The painter doesn't add any content to the scribes words; he simply illustrates what he says—he 'paints images *of* the words in the soul'—so what he makes is a repetition, in a new medium, of the judgements we make. It seems, then, that it is the medium, not the content, that makes the images necessary: as in the *Timaeus*, passions are sensitive to pleasure- and pain-infused images but insensitive to *logoi*.

Both the *Timaeus* and *Philebus* add the right kind of cognitive abilities and add them for the right kind of reasons: the fact that rational content—explicitly linguistic content in the *Philebus*—is inaccessible to non-rational cognition and must, therefore, be translated into sensory appearances. Their use of the the analogy of painting also has a striking resonance with book 10's claim that 'the poet uses words and phrases to paint coloured pictures' of what he imitates. There are, then a variety of reasons, both from within the *Republic* and from later dialogues, to confine the non-rational parts' comprehension of the world to sensory appearances. To do so in relation to poetry requires us to broaden our conception of sensory appearances so that it include the kind of mental images that the poet evokes but this is, on reflection, the simplest way to explain both what poetic 'appearances' are and how they affect the non-rational parts of the soul. This gives us an answer to the first of the two problems that poetic appearances pose and allows us to refine the second: how can sensory appearances have evaluative content?

5.2.2 *Ethical appearance*

The passage that marks the transition from non-evaluative to evaluative appearances is Socrates' distinction between the user, maker, and imitator of a product (601B–602B), where he introduces both paintings and poems that imitate the evaluative properties of an object. Socrates begins with a succinct statement of the relation of value and function: 'aren't the virtue, fineness, and correctness

(*aretê kai kallos kai orthotês*) of each manufactured object, living creature, and action related to nothing but the use for which each is made or naturally adapted?' (601D4–6). Succinct as it is, this is a very broad account of value, covering a comprehensive range of both value terms and objects to which they can be applied. Socrates focuses here on the fineness of manufactured objects.

Since something's value lies in its function, Socrates tells us that it is the expert who uses it who knows whether something is a good or bad example of its kind: the horseman, for example, has knowledge of how a rein or mouth-bit ought to be and the flute-player knows which flutes are good or bad (which respond well or badly in playing). The maker lacks this use-based knowledge of his products' value but by following the instructions of a horseman or a flute-player he acquires the true beliefs needed to craft his products well. The imitator, in contrast, neither uses nor receives instruction about the objects he paints and so he has neither knowledge nor true belief about what qualities make them good or bad. Nevertheless, an imitator can not only paint them or write about them in poems, but do so in such a way that they *appear* to have the virtue, fineness, or correctness appropriate to them. He can do this because 'what he'll imitate, it seems, is what appears fine to the majority of people who know nothing' (602B2–4). This is, at least on reflection, surprising. It suggests that there is a substantial gap between real and apparent value: what appears fine is not the same as what is fine. Moss gives a good summary of the conclusion we should draw:

> The function of a carpenter's lathe (what it is for) does not meet the eye, but must be understood. Since value is dependent on function (601d), it is also the case that much of what makes for a genuinely *good* lathe does not meet the eye (certainly not the eye of the layman). But if genuine value is nonapparent, there is something related that does meet the eye: apparent excellence or fineness, the quality of appearing, not being, excellent or fine. (2007b, p. 424)

Genuine value is not (at least not clearly or directly) something that can be perceived. A flute that plays badly could look identical to a flute that plays well, so that only the flute-player could tell the difference—not through sight, but through use. Moreover, a grand, ornately embellished flute may appear to be very fine to those who lack knowledge of flute-playing, yet may in fact play very badly.

The same point can be made about human virtue and, given the account of the virtues in the *Republic*, in much the same way as it can be made about a flute. For example, justice is defined as a state of *inner* psychic harmony, which can be expressed in outward behaviour, but is not defined by it. More specifically, justice is a kind of psychic 'doing one's own' of each part of the soul, which is distinguished from doing one's own in relation to others. The latter is only 'a sort of image of justice' (443c4–5); in truth, justice 'isn't concerned with someone's doing his own without, but with what is within, with what is truly himself and his own' (443d9–c1).[28] The justice of outward actions is defined entirely in terms of their relation to this inner state, as those actions that preserve or promote psychic harmony (443e5–6).

Naturally, then, it would be impossible for someone who doesn't know what justice is—who certainly cannot see this inner harmony—to read a person's virtue from their behaviour. Virtue is not a perceptible property. And yet this does not prevent the poet from producing 'images of virtue'. The reason is that, like an ornate flute, there is a second kind of 'virtue' that is expressed in outward appearance: apparent rather than real virtue. Indeed, in book 2 Adeimantus claims that poetry is almost ideologically in favour of merely apparent virtue, encouraging one to value it more highly than real virtue. He imagines a young man indoctrinated by poetry thinking that 'since "seeming overpowers the truth" (*to dokein … kai tan alatheian biatai*) and controls happiness, as wise men [sc. the poets] make clear to me, I must wholly rely on this. As a front and a pretence, I

[28] While certainly interesting, what is meant by 'a sort of image of justice' is not entirely clear, for a number of reasons, and I am not sufficiently confident with any one interpretation of the passage in which it occurs, 443c4–7, to draw a significant conclusion from it. Wilson takes the 'image' to be the outward behaviour of a just man, and he concludes that 'Plato in the *Republic* does speak of outer behaviour as an image or reflection of the inner self' (1976, p. 122). Vlastos agrees (1969, p. 514). However, while it is true that anyone who is psychically just will also do his own in relation to others (441d12-e2), it does not seem that anyone who does his own in relation to others will be psychically just. Socrates' examples of the 'image of justice' is the doing one's own of a carpenter and a cobbler and it is unlikely that he regards them as having met the difficult criteria of psychic justice. So if an interpretation along the lines of Wilson's is correct, we should note that the external image of justice and genuine justice will not always coincide, which accords with the other examples of 'images' that we find in the *Republic*.

should put around me a shadow-painting of virtue (*skiagraphian aretês*)' (365c1–4). Thus, the *Republic* presents a view of apparent value that is similar, and similarly pessimistic, to the view found in the *Gorgias*. Appearance and reality fail to coincide: what *appears* to be virtuous or fine is not what *is* virtuous or fine. As a result, those who lack the relevant expertise—specifically those who will form their judgements on the basis of appearances—are likely to be misled.

The user-maker-imitator heirarchy makes it clear that evaluative appearances present exactly the same kind of deception as non-evaluative appearances: just as a straight stick can falsely appear bent or a bed can falsely appear to vary in shape, an object, person, or action can falsely appear to be virtuous or fine.[29] However, a substantial question still remains. While it is clear how a submerged stick merely appears bent, it is not yet entirely clear how a person or action merely appears virtuous or fine. It is especially unclear how they appear so in a way that can move the non-rational parts of our soul to desire and action, as Socrates undoubtedly thinks poetry does. Why would the appetitive part of the soul in particular care whether a character or action is virtuous or vicious? To answer this, we need to look at the qualities that give rise to false evaluative appearances.

While it has not always been noticed, if we look carefully at the discussion of poetry's effect on the soul (603c–607a) we will see that Socrates makes it quite

[29] In contrast to Moss, I take the the non-evaluative case to be fully analogous (in structure at least) to the evaluative case. Moss sees 'a very marked *dis*analogy between the two cases. The appearance of a bed supervenes on and is caused by a particular actual bed; an apparently good lathe, however, floats quite freely of any genuinely good lathe' (2007b, pp. 425–426). I think this does not correctly align the features of the two cases. An appearance of a lathe—whether apparently good or apparently bad—will indeed supervene on an actual lathe, just as an appearance of a bed will supervene on an actual bed. What fails to supervene in both cases, however, is the properties the appearance attributes to the actual bed or lathe. As I argued in §4.3, the apparent bed not only fails to be a real bed, but also fails to be a veridical representation of a real bed: its shape appears to vary; a real bed's shape remains the same. Thus, just as the apparently good lathe supervenes on a poor lathe, not a good lathe, the apparently variably-shaped bed supervenes on a consistently-shaped bed, not a variably-shaped bed. This is clearer in the less contentious case of the submerged stick that falsely appears bent: the apparently bent stick floats freely of a genuinely bent stick, supervening instead on a straight stick. In general terms, whether in an evaluative case or not, the appearances Plato refers to are always removed from what they are an appearances of in two ways: first, they fail to *be* what they are an appearance of—the apparent bed is not really a bed—and, second, they fail to *represent* (i.e. differ in some relevant property from) what they are an appearance of—the bed appears to vary shape, while in fact remaining the same.

explicit that what adds evaluative content to a poem is pleasure and pain.[30] He opens with a careful, exhaustive definition of what poetry imitates:

> S: We say that imitative poetry imitates human beings acting voluntarily or under compulsion, who believe that, as a result of these actions, they are doing either well or badly and who in all this experience pleasure or pain. Does it imitate anything apart from this? G: Nothing. (603C4–9)

It is here that the poet finds 'images of virtue'. Socrates is not explicit about the role pleasure and pain play. A natural assumption would be that someone's belief that they are doing well aligns with an experience of pleasure and their belief that they are doing badly aligns with an experience of pain. As we will see, the connection is more complicated than this. In Greek poetry, and in tragedy as a matter of definition, the hero is often doing very badly. This might even suggest a *prima facie* objection: why would an audience admire or wish to emulate a suffering, tragic character, even if they do appear virtuous? Socrates pre-empts this objection with an analysis of grief. Grieving is introduced as an example of behaviour that meets the definition of what poetry imitates.[31] It can be seen as an example of the genesis of an evaluative appearance in poetry, beginning with the kinds of real grieving that the poet imitates (603E–604D) and ending in his representation of this grief in his poetry and the effect it has on his audience (604D–606B).

Socrates introduces two ways to grieve—two candidates for how one *ought* to grieve—only one of which is an appropriate object of poetic imitation. The first is the response to a loss that 'reason and law' (*logos kai nomos*; 604A10) prescribe, which is 'to keep as quiet (or 'still'; *hêsuchian*) as possible in misfortunes and not get excited about them' (604B9–10). Opposing reason and law is pleasure and pain: 'isn't it reason and law that tells him to resist his pain, while his experience of it tells him to give in to it?' (604A10–B1). At the centre of grief is the pain of suffering a loss, which motivates expressions of one's suffering: remem-

[30] An exception is Belfiore (1983). For example: the imitative poet 'leads us to judge good and evil by the false standards of pleasure and pain, to take images of virtue, *eidôla aretês*, for the real thing' (p. 44). This is precisely the claim I aim to establish.

[31] Immediately following the definition, Socrates asks if 'a person of one mind in all *these* circumstances?' (603C10–D1; my emphasis) and grieving is introduced as one such circumstance—one circumstance poetry can imitate—in which a person is not of one mind.

bering one's loss, weeping, and lamentation. But these expressions of suffering are also kinds of pleasures, affording some release from one's suffering. Thus, the 'lamenting part' of the soul—which for the present we can tentatively identify with the appetitive part (but see n. 33)—'hungers for the satisfaction of weeping and wailing' (606A4–5).[32] The second kind of response to a loss, then, is to follow the direction pleasure and pain lead us, expressing one's grief rather than standing firm against the pain and being 'quiet'.

As pleasure and pain are experienced in one's own grief, they are likewise experienced when we engage in poetic imitations of grief. The part of our soul that 'hungers for the satisfaction of weeping … is the very part that receives satisfaction and enjoyment (*pimplamenon kai chairon*) from poets' (606A6–7), presumably by vicariously experiencing the satisfaction the poem's character finds in weeping, as a vicarious relief from the character's pain. Socrates takes such pleasure to be the *aim* of poetry: the decent man allows himself to indulge in poetry solely because 'it has this advantage: pleasure' (606B3–4) and imitative poetry is described as 'the art of poetry related to pleasure' (*hê pros hêdonên poiêtikê*; 607C4–5). Finally, and most decisively, in the closing words of his argument Socrates makes it clear that the role of pleasure and pain that is seen in imitations of grief can be applied generally to poetic imitations:

> If you admit the pleasure-giving Muse, whether in lyric or epic poetry, pleasure and pain will reign (*basileuseton*) in your city instead of law or what is always commonly believed to be best, reason. (607A5–8)

This contrast between pleasure and pain, on one side, and law and reason, on the other, is precisely what was illustrated in the analysis of grief, suggesting that Socrates introduces grief as a model for all evaluative appearances in poetry. The reason it causes pleasure and pain to 'reign' in a city is presumably that it encourages people to take this as the measure of how one ought to behave, rather than the measure of reason and law.[33]

[32] Compare the *Philebus*, where lamentation is described as one of the 'pleasures mixed in with pain' (48A1–2).

[33] If poetry has evaluative content by virtue of representations of pleasure and pain, how does

Poetry encourages this and not the reign of reason and law because behaviour governed by reason and law is outside of imitation's remit (recall the definition of its subject: people acting 'who in all this experience pleasure and pain'). Of the two responses to grief, Socrates tells us why only one of them can be effectively imitated:

> Now, this excitable character admits of many multicoloured imitations. But a rational and quiet character, which always remains pretty much the same, is neither easy to imitate nor easy to understand when imitated, especially not by a crowd consisting of all sorts of people gathered together at a theatre festival, for the experience being imitated is alien to them.... So clearly an imitative poet isn't by nature related to the part of the soul that rules in such a character, and, if he's to attain a good reputation with the majority of people, his cleverness isn't directed to pleasing it. Rather, he's related to the excitable and multicoloured character, since it is easy to imitate. (604E1–605A6)

What motivates the rational, 'quiet' character is the reasons for avoiding excessive grieving that Socrates lists (604B9–D4). He is not easy to imitate because his motivations have no highly visible external manifestation (one cannot *see* his reasons). Rather, he 'remains pretty much the same', since reason dictates that his loss should not cause any great change in his behaviour. If he experiences pain at his loss, it will only be very subtly displayed in his actions. In contrast, what motivates the character who follows what pleasure and pain dictate is writ large in his behaviour. His tears and lamenting are visible suffering, direct expressions of the pain and pleasure that leads him to act as he does. Thus, an audience can read the excitable character's motivations directly from his behaviour and easily understand and identify with them (and receive pleasure in the process). In short, then, the motivation for acting on the pain of grief and pleasure of lamentation is highly apparent, while the reasons to remain 'quiet' or 'still' in grief are not.

At this point it is useful to recall a point I made in §5.2.1, which answers a

---

it appeal to the honour-loving, spirited part of our soul (as implied by 606D1–2)? While it requires greater defence that I can offer here, I believe the answer is that the spirited part is moved by the pleasure of honour and pain of shame. Book 9 not only tells us that the spirited part has it's own pleasures and pains (e.g. 586C–587A), it also clearly implies that when it is misled, it is misled by the *pleasure* of anger, honour, and victory (586C7–D2). We might also note that in the *Philebus* all passions, including typically spirited passions (47E1–3), are identified with pleasures and pains (see Moss (2012)).

possible objection to Socrates' claim that the rational character isn't easily imitated. The objection is that the poet could, in theory at least, do an excellent job of portraying the rational character in a poem. To do so he just needs to make his character state and defend the reasons for his 'quiet' behaviour, making him explain that 'it won't make the future better to take my misfortunes hard; in fact it would make it worse, since excessive grieving would cloud the very thing I most need now, deliberation' (loosely paraphrasing 604B9–D4). This is how Plato represents characters in his dialogues—think of Socrates in the *Crito* explaining his reasons for submitting to the death penalty rather than escaping. It is important to understand why Socrates thinks this kind of representation is not open to the imitative poet. Only the least important part of it can be called an imitation. It is an imitation of a rational character making an argument insofar as 'making an argument' is a kind of behaviour, but the content of the argument is not itself an imitation. To imitate is to make images at a third remove from the truth, but the reasons the rational character will give are not mere images or appearances of reasons and they are not third from the truth. So the skill one needs to reproduce an argument in a poem is not the imitator's skill, the ability to reproduce appearances, but the philosopher's skill, the ability to understand and make rational arguments. Thus, the rational character is difficult to imitate, and make comprehensible if imitated, because his motivations—his reasons for acting as he does—are simply opaque to imitation.

Furthermore, if the poet were to venture beyond imitation and present the case for being measured in one's grief, this itself will not be easy to understand for a theatre audience who are judging by appearances, in contrast to the immediately apparent motivations of the excitable character. The difference between the rational and excitable characters is very similar to the difference between the doctor and pastry chef of the *Gorgias*: in both cases the contrast is between a rationally defensible case for what one ought to do, on the one hand, and what merely appears to be what one ought to do, on the other, where the latter is what the inexpert majority find most persuasive. The *Republic*, however, has a more

sophisticated psychological explanation for why the truth can fail to persuade, but also an explanation that, like the *Gorgias*, takes pleasure and pain to be central. This will be explored in detail in the next section, but it can be summarised as follows:

The part of the soul that the poet is 'directed to pleasing' is the lamenting part, a part that lacks the cognitive resources needed to reason or to grasp the truth. It is also a part of the soul that is by nature susceptible to imitative poetry, that is, to appearances evaluatively tinted by pleasure or pain. It is no mystery why the excitable character appeals to this part of the soul and the rational character doesn't: not only are the rational character's motivations beyond its limited cognitive grasp, but even if it did understand the reasons for remaining quiet in grief it is not clear that it would be *moved* by them—it is obviously less likely to respond to the idea that, for example, 'human affairs aren't worth taking seriously' (604B12–C1) than it is to respond to the pain suffered in grief. So the lamenting part of the soul is easily convinced by the poet. The real danger of poetry, however, comes from the fact that an effect on one part of the soul is not easily isolated from the rest of the soul. By corrupting our lamenting part, poetry threatens to corrupt the whole person: a sufficiently strong lamenting part of the soul can not only overcome the rational part, causing a person to akratically grieve excessively, but can also cause the rational part to endorse its strong desires and so come to believe that grieving excessively is how one ought to grieve. Thus, imitative poetry uses pleasure and pain to persuade us indirectly, changing the intellectually stronger element in us by preying on the intellectually weaker element.

5.2.3 *An appetite for virtue?*

Before concluding the discussion of ethical appearances, I wish to briefly answer a possible objection to my reading, one that draws out one of its fundamental assumptions. The objection is that by reducing poetry's appeal to pleasure and pain Socrates has come too far from images of *virtue*.

To begin with, a few clarifications about what images of virtue are. First and fairly obviously, it is not that the heroes of a poem falsely appear to act in accordance with what really is virtue—e.g. appearing to act like they have well-ordered souls—but that the way they *do* act falsely appears to be what virtuous behaviour is. To take Socrates' example: it is not that the hero appears, but is not, measured in his grief, as the genuinely good man would be, but that he gives into his grief and laments in a dramatic and moving way that makes this appear like commendable behaviour. Second, it would seem that Socrates is using 'virtue' in its most basic normative sense. That is, it doesn't indicate specifically moral behaviour or even behaviour that accords with the canonical Greek virtues, but the more general notion of how one ought to behave. This seems to fit Socrates' account of poetry's subject as well as his examples. There is no suggestion that the quiet character appears less *temperate* in his grief than the dramatically lamenting character. The false appearance is that dramatically lamenting, whether intemperate or not, appears to be the better way to behave in the circumstances. So while the lamenting part of the soul may lack high-minded concepts like temperance or justice, it is moved by images of virtue in this less sophisticated sense, specifically by pleasure- and pain-tinted images of how one ought to behave and of what sort of person one ought to be. It is only once a whole person is persuaded by such appearances that this will inform a persons view of either the value or nature of such virtues as temperance or justice (e.g. believing, like Polus, that justice is secondary to the pursuit of pleasure or, like Callicles, that justice is the pursuit of pleasure).

This picture needs to be complicated still further. The position I've been defending implicitly sides with one of two accounts of how each part of the soul 'chooses' its object of desire. The first view is what I will call the 'hard-wired' account. On this view, each part of the soul has desires that are intrinsically responsive to their characteristic object: the appetitive part's desires are hard-wired to the pleasant; the spirited part's to the fine; and the rational part's to the good. Thus, the appetitive part's desire for pleasure is axiomatic; it requires no explanation

beyond the fact that this just happens to be how its desires function. Recently, an alternative to this view has gained popularity, which I will call the 'intellectualist' account.[34] On this reading, in the *Republic* Plato still adheres to the Socratic claim that we desire something if and only if we find it good and desire it because we find it good. Rather than being an axiomatic feature of their desires, the characteristic object of each part of the soul is what it typically believes (or otherwise cognises) to be good. What each typically believes to be good is, in turn, explained by the cognitive resources available to it. The appetitive part, for example, lacks the ability to calculate, but has the ability to apprehend appearances and, consequently, it believes that what is good is what it appears to be, pleasure. In the present context, what is at stake between these two views is the idea that the non-rational parts are capable of having a conception of how one ought to behave. On the hard-wired reading, it seems that we should explain at least the lowest part of the soul in descriptive rather than normative language: the appetitive part just reacts to pleasures and pains, being capable only of instinctive, automatic responses. On the intellectualist reading, on the other hand, it makes more sense to say that each part of the soul is motivated by a conception of what it ought to do or what it is best for it to do.[35]

The account of non-rational cognition I have defended strongly favours the intellectualist reading. In book 4 Socrates distinguishes between parts of the soul that have distinct objects of desire; in book 10 Socrates distinguishes between parts of the soul that have distinct cognitive abilities.[36] According to the account I've been developing, the former is explained by the latter. The model for this explanation is 602C–603A: the lower parts of the soul accept the mere appearance

[34] Variants of this view are defended by Lesses (1987); Carone (2001); Bobonich (2004); Morris (2006); Moss (2006) and (2008); and Singpurwalla (2011). See also Price (1995, p. 48 ff.).

[35] The difference can be illustrated by the following counterfactual: if it were possible to make the appetitive part fully understand that a pleasure is harmful to it (some harm not measured by pain), according to the intellectualist reading this would be sufficient for it to cease desiring the pleasure but according to the hard-wired reading it would, irrationally, continue pursuing the pleasure.

[36] In books 8–9, it could be argued that we find an intermediary position: the parts of the soul are distinguished by their distinct beliefs and values. Some examples of this are found in the next section.

that a stick is bent because they lack calculation. Similarly, poetry appeals to the lower parts of the soul—to the passions of the lower parts of the soul—because it presents appearances that these parts lack the calculative abilities to see through. And the lamenting part of us favours lamenting and weeping because it is incapable of appreciating the reasons for grieving quietly, yet is fully cognisant of the apparent value of grieving like the excitable character does. (The hard-wired reading must reject the 'because' in the last two sentences: the appetitive part of the soul would pursue pleasures *even if* it could use calculation to figure out that they are bad pleasures—its problem is not that it fails to appreciate that lamenting dramatically is bad or shameful, but that it is just not moved by this kind of consideration.) More importantly for the present context, it also explains how Socrates can consistently argue both that poetry imitates apparent value (recall especially 601B–602B: he imitates apparent 'virtue, fineness, and correctness') and that these imitations appeal to the non-rational parts of our soul. As we will see in the following section, this claim is essential to the 'greatest charge' against poetry.[37]

[37] Detailed defences of this reading within the context of book 10 can be found in Moss (2006) and Singpurwalla (2011). It is often thought that the most obvious textual evidence against the intellectualist reading is 437D8–438A5. In this passage Socrates argues that thirst is simply a desire for drink, not for drink of a qualified kind, and he rejects the following objection: 'let no one catch us unprepared or disturb us by claiming that [A] no one has a desire for drink but rather for good drink, or food but good food, on the grounds that [B] everyone after all desires good things'. This is often taken to be a rejection of B and, therefore, a rejection of Socratic intellectualism. I find Lesses (1987) and Morris's (2006) response persuasive: Socrates rejects not B, but the claim that B implies A. That is, he is pointing out that it is a mistake to infer from the fact that I desire drink *qua* good—i.e. believing that drink itself is good—that what I desire is drink of only a qualified sort, namely good rather than bad drink. (Compare the difference between saying one desires 'good pleasures', as a virtuous person undoubtedly does, and saying one desires 'pleasure *qua* good', which is true of a vicious person.) That this is how the passage should be read is perhaps clearest when one realises that Socrates could have made precisely the same point about pleasure, that is, he could have said 'let no one disturb us by claiming that no one has a desire for drink but rather for pleasant drink, or food but pleasant food, on the grounds that the appetitive part of the soul after all desires pleasant things.' In fact, we should find it interesting that this isn't the objection Socrates considers: if the hard-wired reading is right that the appetitive part of the soul, the part that thirsts, desires pleasant things intrinsically—rather than pleasure *qua* apparent good—then the more relevant objection would indeed be that thirst is not for drink of an unqualified kind, but for 'pleasant drink'. As for some more explicit (albeit contentious) evidence that Socrates does accept B in the *Republic*, see 505D11–506A2: e.g. 'every soul pursues the good and does whatever it does for its sake'.

5.3 THE PSYCHOLOGICAL EFFECT OF POETRY

We are now in a position to consider the culmination of the *Republic*'s account of appearances. In §5.1.2 we saw that poetry, being a serious source of ethical miseducation, is prominently represented in the Cave. Book 10 explains why and how poetry miseducates. Its stated aim is to show that all imitative poetry 'is likely to distort the thought (*dianoias*) of anyone who hears it, unless he has the knowledge of what it is really like, as a drug to counteract it' (595B5–7). This is a topic that deserves the quite considerable attention that Plato gives it in part because of the danger of poetry itself, as the most influential and appealing of the image-making crafts, but also in part because it is an opportunity to explore the psychological effect of appearances *per se* and, therefore, to explain why the majority have such a low level of ethical understanding—why most people, like the prisoners in the Cave, are convinced by mere shadows of justice. This is a topic that Plato has explored before, in the *Gorgias*, and as in the *Gorgias* the most important claim is that the majority of people find appearances even more persuasive than the truth. What Plato adds in the *Republic* is a sophisticated—and, I will argue, quite convincing—account of how this persuasion works.

Poetry's danger lies not simply in the fact that it presents misleading appearances but in the fact that these appearances bypass the rational part of a person's soul, which might be more guarded against its influence. This leads to what Socrates calls the 'the greatest charge' against poetry (605C6): even the 'decent man' (*ho epieikês*)—someone who has a reasonably sound grasp of ethical arguments—can be corrupted by the compelling appearances that poetry presents. The decent man's error is one that we can easily be sympathetic to: he believes that poetry is harmless entertainment and so can be safely enjoyed. What he has failed to appreciate is the arguments of book 10—the 'knowledge of what it is really like' that counteracts poetry's effect—which show that this seemingly innocuous enjoyment of poetry has a powerful effect on the non-rational parts of his soul, shaping their aspirations to the shameful behaviour of the poem's heroes and

strengthening their position in the soul. Once they are powerful enough, they can lead him to emulate this behaviour in his own actions and, ultimately, to mistake it for virtue. It goes without saying that if it has this effect on the 'best among us' (605C10), its effect on less morally aware citizens will be even more acute.

I will begin by looking at what it means to say that poetry 'is likely to distort the thought of anyone who hears it', specifically at what change it has the potential to cause in the soul. I will then look in more detail at how it does this. I will argue that Socrates presents an account of what we might loosely call 'suspension of disbelief': in order to enjoy a poem, we set aside any beliefs that conflict with a poem's evaluative content. I will argue that this is achieved by suspending rational cognition in favour of non-rational cognition. We derive pleasure from a poem not through the rational part of our soul, but through the non-rational parts of our soul and, accordingly, the evaluative beliefs that are active as we enjoy a poem are non-rational beliefs. Strange as this account might sound, I will argue that it is in fact defensible.

5.3.1 *How poetry corrupts*

The poet is said to put 'a bad constitution in the soul' because he 'arouses, nourishes, and strengthens [the inferior] part of the soul and so destroys the rational part, in just the way that someone destroys the better sort of citizens when he strengthens the vicious ones and surrenders the city to them' (605B3–6). A little later we are told that: 'in the case of lusts and anger and all that is appetitive and painful and pleasurable in the soul … [the poet] nurtures and waters them and establishes them as rulers in us when they ought to wither and be ruled' (606D1–5). It is not simply that poetry can strengthen non-rational passions sufficiently to cause *akrasia*—although this is likely to be a intermediary step—but that it causes them to *rule* the soul, setting up a certain bad psychic 'constitution' or 'government' (*politeia*), and in this process 'destroys the rational part'. The political analogy refers us back to the discussion of the four vicious political and

psychic constitutions described in books 8 and 9. Here we find more detailed (though still not fully satisfactory) descriptions of what happens when one part of the soul 'rules' the other parts. There are three passages that are particularly helpful.

Passage 1 describes the oligarch who is controlled by necessary appetites (the appetites needed to sustain normal bodily well-being). These necessary appetites are called 'money-making' (*chrêmatistikos*; 553C2) since they are satisfied by means of money yet are sufficiently restrained to allow an accumulation of wealth. Thus, the oligarch's soul becomes entirely devoted to wealth:

> The rational and spirited parts are made to sit on the ground beneath it [sc. the money-making part], one on either side, reducing them to slaves. He won't allow the first to reason about or examine anything except how a little money can be made into great wealth. And he won't allow the second to admire and value (*thaumazein kai timan*) anything but wealth and wealthy people or to have any ambition (*philotimeisthai*) other than the acquisition of wealth. (553D1–6)

Passage 2 describes the democratic man who is controlled by unnecessary appetites, appetites that seek pleasures beyond what is necessary for health and, in contrast to money-making appetites, are called 'spend-thrift' (*analôtikos*; 599C3). Socrates uses a political analogy similar to the one in book 10:

> S: [The democrat's appetites] occupy the citadel of the young man's soul, finding it empty of knowledge, fine ways of living, and words of truth, which are the best watchmen and guardians of the thoughts of those men whom the gods love. A: They certainty are. S: But now false and boastful words and beliefs rush up and occupy this part of him. (560B7–C3)

The result for the democratic man is an inversion of values: he'll call 'respect foolishness and moderation cowardice' (560D2–3) and 'insolence good breeding, anarchy freedom, extravagance magnificence, and shamelessness courage' (560E4–561A1).

Finally, passage 3 describes a similar change, somewhat more literally, in the tyrannical man. In most of us our worst desires are only 'awakened in sleep, when the rest of the soul slumbers—the rational, gentle, ruling part' (571C3–5). During this time this part finds satisfaction in dreams, engaging in murder or 'trying to have sex with a mother, as it supposes (*hôs oietai*), or with anyone else, whether

man, god, or beast' (571C9–D2). In the tyrannical man, in contrast, this lower opining part overcomes his rational part even while he is awake:

> The old traditional opinions that he held from childhood about what is fine and shameful—opinions that are accounted just—are overcome by the opinions, newly released from slavery, that are the bodyguard of *eros* … [These] opinions used only be freed in sleep, but under the tyranny of *eros*, he has permanently become while awake what he used to become occasionaly while asleep. (574D5–E3)

These three passages tell us that being 'ruled' by a part of the soul—or, as it seems in at least the oligarch's case, a set of the desires within this part—effects a change in our whole psychology. Passage 1 describes how the goals of the non-ruling parts are subordinated to the goal of the ruling part, the goal of money-making. All of the oligarch's cognitive and conative abilities become devoted to this goal: the rational part calculates the means to wealth; the spirited part admires wealth and the wealthy; and the appetitive part, the controlling part, sets all its appetites (even its unnecessary ones: 554C11–D3) on wealth. While there is some language of force—the rational and spirited part's are controlled like 'slaves'—the ruling part does not, for example, force the spirited part to act contrary to its natural love of honour, acting in ways that it finds shameful, but rather causes it to find wealth itself honourable. Likewise, the rational part of the soul comes to equate the best life with the wealthiest life and to dedicate its love of wisdom to financial truths.

Passages 2 and 3 describe the effect being ruled has on the rational part in particular. The defining feature in both accounts is that, once subjugated, the rational part comes to agree with the beliefs of the newly ruling part. Socrates' two accounts of how this comes about are brief and opaquely metaphorical: the opinions of the part fighting for control of the soul 'rush up and occupy' the intellectually weakened rational part (passage 2) or 'overcome' the traditional opinions that it held (passage 3). This makes it difficult to figure out the process by which this doxastic conversion takes place, but for our purposes the process is less important than its conclusion, which Socrates is far more explicit about: the

rational part comes to value the same things as its ruling part.[38] The specific beliefs that it acquires are those evaluative beliefs that were previously held only by the ruling part: the democratic man takes on a new, inverted view of traditional virtues and the tyrannical man acquires his ruling part's opinions 'about what is fine and shameful', opinions that used to only be released in deviant dreams. We should note that these passages are also consistent with Socrates' description of what happens when the ruling part is the rational part of the soul: for example, each part will 'follow knowledge and argument and pursue with their help those pleasures that reason approves' (586D6–7) and will 'believe in common that the rational part should rule' (442C11–D1).

These passages tell us what to expect if poetry establishes a non-rational part as a ruler in one's soul and makes it clear how it 'distorts the thought' and 'destroys the rational part' of its audience. The crucial consequence is that the evaluative beliefs of the parts being ruled are caused to agree with the beliefs of the ruling part.[39] Books 8 and 9 describe the successful usurpation of a whole soul, but we should expect the overthrow to occur in degrees and, as book 10 suggests, to occur more locally, from a specific passion to a specific supporting belief.

In Socrates' 'most serious charge' against poetry we see an account of this process—or at least the threat of this process—in a decent man who allows himself to set aside his moral incredulity in order to enjoy and sympathise with a poem's questionable hero. I aim to show that the conflict that the decent man

[38] I take it that it does not in fact involve a literal transference of beliefs between parts of the soul. To speculate on less direct processes we might appeal, for example, to book 6 where we hear that 'when someone's desires incline strongly for one thing, they are thereby weakened for others, just like a stream that has been partly diverted into a channel' (485D6–8), so someone with strengthened non-rational passions will become less interested in other objects of desire. We are told in book 10 that being overcome by the passion of grief prevents one from deliberating, thereby weakening one's defence against, for example, being taken in by the false appearances of a poet's excitable character. More generally we might appeal to the plausible suggestion that we sometimes rationalise our stronger desires, acquiring beliefs that allow us to satisfy them without guilt.

[39] See Scott (1999, pp. 34–36) for a similar comparison: in both the corruption of the democratic man in book 8 and the corruption caused by poetry in book 10 'the conflict between rational and non-rational can result in the expulsion of certain beliefs from the rational part and, in their place, the adoption of new beliefs in accordance of the desires of the non-rational part.'

suffers in this passage, 605C–606B, is the same in kind as the conflict caused by painting that was examined in chapter 3, 602C–603A. As we have seen, poetry and optical illusions are closely related: insofar as both are appearances, they affect the soul in the same way. That said, there is one crucial difference, arising from a difference in their content: the false appearances that poetry presents are not benign, motivationally-inert illusions, but evaluative appearances that stir our non-rational passions and can, therefore, 'destroy' the rational part of our soul.

In 605C–606B Socrates draws attention to the discordant attitudes that the decent man shows in response to, on the one hand, a poem's hero's grief and, on the other, his own grief:

> (A) When even the best among us hears Homer or some other tragedian imitating one of the heroes sorrowing and making a long lamenting speech or singing and beating his breast, you know that we enjoy it, give ourselves up to following it, are affected along with (*sumpaschontes*) the hero, taking his sufferings seriously, and praise (*epainoumen*) as a good poet the one who affects us most in this way. (605D1–5)
>
> (B) But when one of us suffers a private loss, you realise that the opposite happens. We pride ourselves if we are able to keep quiet and master our grief, for we think that this is the manly thing to do and that the behaviour we praised before is womanish. (605D7–E2)

At first sight, this does not seem to be the kind of opposition to which the Principle of Opposites (PO) would apply. There are two features of the opposition that make PO inapplicable: first, they are not strictly speaking about the same thing, since A is about someone else's suffering and B is about one's own suffering and, second, the opposing attitudes occur at different times. With some reflection, however, we will see that it does indeed reveal a *bona fide* instance of psychological conflict.

With respect to the first feature, it is unclear how one can rationally hold opposing views about the same kind of behaviour—to the extent of censuring one and praising the other—merely on the grounds that it is performed by different subjects. To make such a distinction, the decent man would need to believe something like the following: to grieve dramatically is shameful for *me*, but it is worthy

of approval in *others*. Such a belief is possible only if one thinks that one's moral responsibilities are unique in some way. Consider, for example, tolerating a child who cries easily, yet disapproving of the same behaviour in an adult, or disapproving of a distinguished general who fails to be exceptionally courageous, yet excusing the same behaviour among the rank and file. Unlike these examples, however, it is hard to see what difference between the decent man and the tragic hero, whether real or imagined, would account for such a disparity between their responsibilities. The decent man does not just excuse the hero's behaviour, putting aside his moral censure out of compassion for his suffering, but rather 'approves of' or 'praises' it (*epainein*; 605E4–6, 606B3).[40] Nor does the decent man ever appeal to any unique, extenuating circumstances that warrant his reversal from blame to praise. It seems, then, that he holds more seriously conflicting beliefs, beliefs that fail to find resolution in an implicit distinction between subjects: in A, the belief that dramatic grieving—whether one's own or another's—is appropriate and praiseworthy and, in B, the belief that dramatic grieving—whether one's own or another's—is 'womanly' and shameful.

This brings us to the second putative exception to PO: that the opposing beliefs are held at different times and so never, synchronically, conflict. It is likely, however, that the conflict between A and B indicates a similar conflict within each of A and B. We already know that in one's own grief there can be conflict between opposing parts of the soul. Showing this was the aim of Socrates' earlier analysis of grief, discussed in §5.2.2. It seems reasonable, then, to assume that we should apply the conflict considered in the analysis of grief to the personal grief described in B. One apparent difference is that in the analysis of grief the conflict was desiderative but the proposed conflict in B would involve beliefs: approving and disapproving of a certain behaviour. As we've seen, however, insofar as the grieving part is guided by evaluative appearances, its passions are guided by beliefs and, therefore, it enters into cognitive conflict with the rational part of the

[40] At 605E4–6 Socrates summarises the opposing attitudes as 'approval' and enjoyment, on the one hand, and finding something 'unworthy' (*mê axia*) and shameful, on the other.

soul, which believes one should grieve quietly. At least in B, then, is seems safe to assume that the decent man suffers a synchronic conflict that appears very close to the diachronic conflict between A and B: a conflict between the lamenting part of us that is led by the merely apparent value of dramatically grieving, and so desires this, and the rational part that believes that this is shameful, and so desires to grieve quietly.

Does the same conflict occur when, as in A, we enjoy a grieving hero in poetry? There is one significant difference between A and B: there is no struggle in A. The decent man simply enjoys the hero's performance. Socrates explains why: unlike in his own grief, the decent man believes there is no harm or shame in releasing his 'lamenting part' and allowing it to be emotionally affected by, enjoy, and approve of the hero's actions. He mistakenly believes that this will not affect *his* actions and, moreover, he knows that if he didn't release his lamenting part, he would not be able to enjoy the poem, since his reaction would be dominated by rational condemnation of the hero's shameful actions. It is not that for the duration of the poem he—or specifically his rational part—ceases to think that the hero's actions are shameful, but that he 'suspends' his moral judgements so that he can enter into the ambient morality of the poem and enjoy it as the poet intended. I believe this is a plausible, defensible account of a specifically moral sort of suspension of disbelief, and of its danger. Socrates' first insight into this common experience is that to 'suspend' our judgements is to allow a less critical, more appearance-sensitive part of us dominate for the duration of the show. His second insight is that in doing so, we strengthen this part and put ourselves in danger of being dominated by it in real life, both by its desires and beliefs.

Let us first look at the passage from which I drew the above explanation. Glaucon says that our attitude in A, given our different attitude in B, seems to be without 'good reason' (*eulogô*). Socrates agrees and tells him that the following thought will explain why:

> [You will understand why] if you reflect that the part restrained by force (*to bia katechomenon*) in our own misfortune, and that hungered for the satisfaction of weeping and lamenting, because it desires these things by nature, is the very

> part that the poets satisfy and please. But the best part of our nature, since it has not been sufficiently educated in reason or habit, relaxes its guard over this sorrowful part, since this is watching another man's sufferings and he thinks it is no shame to himself to approve of and pity (*epainein kai eleein*) someone else who, claiming to be a good man, grieves excessively, but rather that it has this advantage, pleasure, and he would not choose to be deprived of this by despising the whole poem. For few, I think, are among those who can conclude that the pleasure we take from another's suffering necessarily transfers to our own. For once the pitying part (*to eleinon*) is nursed to strength in another's suffering, it is not easy to restrain it in our own suffering. (606A3–B8)[41]

The decent man, then, willingly relaxes his guard over his 'sorrowful part' because he has made a mistake. He thinks that there is no harm or shame *for him* if he approves of *another* man's suffering. The fact that he needs to draw this distinction in order to feel justified in taking pleasure in the poem implies that he does in fact know that the other man's lamenting is shameful—otherwise he would approve of his lamenting simply because he approves of it, not because he believes he is somehow isolated from something shameful. This is as we would expect if, as is surely the case, he has not forgotten or failed to apply the beliefs he expressed in

[41] My translation. There is a significant disagreement about the translation of this passage, specifically of 606B1–5: *'…, hate allotria pathê theôroun kai heautô ouden aischron on ei allos anêr agathos phaskôn einai akairôs penthei, touton epainein kai eleein, all' ekeino kerdainein hêgeitai, tên hêdonên, kai ouk an dexaito autês sterêthênai kataphronêsas holou tou poiêmatos.'* The masculine *kataphronêsas* near the end of this sentence tells us that at some point the subject has moved from the neuter *to beltiston*, the best part of a person, back to the whole person, who is the natural contrast to *allos anêr*, 'another man'. The disagreement concerns where this first happens. I side with Jowett & Campbell (1894) and Adam (1902) who take this point to be the *heautô* at B1. Mastrangelo & Harris (1997) argue that the change of subject occurs further down, beginning at the *all'* at B3 and Grube & Reeve take the subject to be *to beltiston* throughout. A very strong reason to favour Jowett & Campbell and Adam's reading is that if the *heautô* were neuter, taking the subject *to beltiston*, then 'the best part' is the same as 'the pitying part' that is said to be *restrained* in our own grief at the end of the passage. The best part, then, would be both the restrainer and the restrained, a conclusion that is both odd in itself and inconsistent with Socrates' discussion of grief. A second reason is that it makes better sense to assume that, as Grube & Reeve translate, the *hêgeitai* at B4—what is merely *thought* to be the case—should encompass the whole correlative conjunction 'it is no shame … but rather …,' taking it as expressing a single thought that the spectator gives to justify relaxing his guard over his 'sorrowful part'. All commentators agree that the *hêgeitai*, given its proximity to the *kataphronêsas*, must take a masculine subject, making it unambiguous that the whole man's thoughts are in question. But in Mastrangelo & Harris's reading the *hêgeitai* applies only from the *all'* at B3, which would suggest that Socrates is first *stating* that it is no shame to praise and pity another man's excessive grieving (despite clearly disagreeing: 605E4–7) and only then introducing the subject's own thoughts on the matter—this is surely not Socrates' intention.

B. However, in addition he *does* 'approve of and pity' the excessively lamenting character, as if it were not shameful. He is, then, in a state of cognitive conflict: approving of and disapproving of the same thing at the same time.

This is a purely cognitive, non-motivational conflict—just like the conflict described in 602C–603A—because it is something the decent man willingly permits: since he believes it is harmless and necessary to enjoy the poem, he refrains from exerting any countervailing motivational force to hold back his approval and pity. This cognitive conflict is possible because, again like the conflict in 602C–603A, each side of the conflict is attributed to a different part of the soul. His rational part knows that the grieving character is acting shamefully because it can use reason and so can appreciate the arguments for grieving quietly. Another part of him, however, is by nature taken in by the hero's apparently honourable grieving and so pities the poet's grieving character and believes that he is, as he claims, an example of a good man. Socrates tells us what part of the soul this is. It is 'the pitying part' (*to eleinon*),[42] which, once strengthened, 'is not easy to restrain' in our own grief. This is of course the same part described as 'the part restrained by force in our own misfortunes' at the start of the passage, which is the part that was described as the 'lamenting part' in the analysis of grief, which is in turn a non-rational part of the soul.

We can conclude that it is a non-rational part of the soul that is moved to approval and pity by the poem. The rational part of our soul relaxes its guard over this part, but it does not change its opinion of how one ought to grieve. It continues to believe that dramatically lamenting is shameful, but for the sake of enjoying the poem it doesn't restrain the lamenting part, as it would in its own suffering. Consequently, there is cognitive conflict between the lamenting and rational part but no motivational conflict. To this extent, the conflict caused by poetry is the same as the conflict caused by paintings and optical illusions. However, with poetry the lower part of us does not stop at simply apprehending an ap-

[42] Or simply 'the pity', i.e. the emotion of pity. On either translation the result is the same: the part that feels the emotion of pity is restrained.

pearance. The appearance stirs its pity and its natural desire to weep and lament, strengthening these passions and making them more difficult to restrain in our own suffering. Socrates goes on to list other passions that can be equally stirred. Comedy appeals to us with jokes that 'you yourself would be ashamed to tell but that you enjoy hearing and don't detest as something bad in comic imitations', relaxing your guard on 'the part of you that wanted to tell the jokes that was held back by reason' (606C2–6). Similarly 'in the case of lusts and anger and all that is appetitive and painful and pleasurable in the soul … [the poet] nurtures and waters them and establishes them as rulers in us' (606D1–5). Without the decent man's noticing, all these passions become strengthened by poetry and threaten to guide his behaviour.

There are degrees of danger involved in poetry. Socrates says the decent man, having strengthened his pitying part on poetry, will not find it 'easy' to restrain this part in his own grief. This falls short of allowing this part to rule us, as 'ruling' was characterised in books 8 and 9. Even if he fails, the decent man will *try* to restrain his pity, which shows that his rational part is functioning reasonably well, even if it now has to deal with a more troublesome non-rational part. This falls short, but is a step in the direction of, poetry's most pernicious effect, its potential to 'distort the thought' or 'destroy the rational part' of its audience—not just making a lower part of us unruly but setting it up to rule in the soul. As we learn from books 8 and 9, once the rational part is ruled by another part of the soul, both its cognitive and conative abilities are subordinated to the goals of the ruling part. Consequently, not just the passions, but the entire attitude released during a poem will spill over into our own affairs. With respect to grieving, this will mean that our rational part ceases to attend to or be persuaded by the arguments in favour of grieving quietly and is instead convinced, in agreement with the lamenting part, by what merely appears to be the best way to grieve.

We can describe poetry's effect in a way that recalls the Allegory of the Cave: it encourages the decent man to wilfully engage in the kind of ekastic cognition, the judging by appearances, that keeps the prisoners ethically shackled. Just as

Socrates' account of the dangers of poetry explains the prominent role of poetry in the shadow-show of the Cave, the Cave should remind us that Socrates takes the danger of appearances to be broader than poetry alone. Poetic appearances have the psychological effect that they do by virtue of being evaluative appearances that appeal to the non-rational parts of our soul. Any similarly compelling evaluative appearance is likely to have the same effect. Those who are less morally sophisticated than the decent man will, it seems likely, have long since been corrupted by poetry, oratory, sophistry, and any other source of misleading appearances, and will therefore consistently approach ethical questions through eikastic cognition that ought to be confined to our non-rational parts. Their comprehension of the evaluative world, then, will extend only as far as mere appearances of what is best, just as the prisoners' awareness never goes beyond mere shadows. Plato's insight, which is an advance on the account we find in the *Gorgias*, is that the danger of such appearances is not simply that they are both false and persuasive in the absence of the reasoning that would reveal them as mere appearances, but that they exert a power over us even, as in the decent man's case, in the presence of such reasoning. In the next, final section I consider the merits of this theory. I argue that his central insight is correct: we cannot comprehensively explain human behaviour and cognition without positing a desire- and action-guiding source of belief, such as appearances, that affects us independently of our rational beliefs.

5.3.2 *The paradox of fiction and other cases of belief-discordant behaviour*

I concluded chapter 3 by arguing that the *Republic*'s characterisation of sensory appearances as a form of non-rational belief is defensible. In the argument from visual illusions in 602C–603A Socrates shows that a consequence of this view is that we have two sources of belief that have the potential to conflict: a set of beliefs that fall into a rational structure—that tend towards mutual implication, abhor contradiction, and are responsive to evidence and reasoning—and that we feel ownership over and a second, independent set of perceptual beliefs that can

lie outside this rational structure. The account of our engagement in poetry in 605C–606B extends that account by illustrating how such conflict can explain belief-discordant *behaviour*, that is, behaviour that is at variance with the beliefs that we would explicitly attribute to ourselves.

Consider first a putative paradox, the so-called 'paradox of fiction' that was first raised by Radford (1975) and can be illustrated with the following inconsistent triad:

(1) To feel a desire or emotion in response to a person or situation, one must have certain relevant beliefs: e.g. to fear, we must believe that we are under threat; to pity, we must believe that someone is suffering.

(2) With respect to fictional people and situations, we lack the relevant beliefs: e.g. we do not believe we can be threatened by a fictional monster; when an actor weeps on stage, we do not believe a real person is suffering.

(3) Fictional people and situations can elicit desiderative or emotional responses: e.g. we can fear a fictional monster; we can pity a fictional character's plight.

I have referred rather vaguely in (1) to 'certain relevant beliefs'. It is not easy to define a set of the beliefs that are required by all desires and emotions, although example by example they are quite clear. Discussion of the paradox usually focuses on 'existence beliefs', the belief—or failure to believe—that the threatening monster or the suffering person exists. Many other kinds of beliefs could be adduced, depending on the passion in question, but one is particularly relevant to the *Republic*: to pity, we must not only believe that someone is suffering but that their suffering is pitiable, that it is *worthy* of pity. Put more generally, desires and emotions rely on certain *evaluative* beliefs, beliefs that can—arguably—be lacking in our emotional responses to fiction. The account of our engagement in poetry in book 10 furnishes us with a response that is equally relevant to existence and evaluative beliefs, but Socrates is concerned with the latter, since it is with respect to these beliefs that poetry is dangerous.

Being a cognitivist about desires and emotions, the Socrates of the *Republic* would agree with (1). As Radford puts the thought behind it: 'it would seem then that I can only be moved by someone's plight if I believe that something terrible has happened to him. If I … believe that he has not and is not suffering … I cannot grieve or be moved to tears' (p. 68). He also clearly agrees with (3), that fiction does move us. What Socrates disagrees with is (2), that the passions that poetry elicits are not supported by the appropriate beliefs.

Socrates doesn't mention 'existence beliefs', but we can be fairly sure of what he'd say. Confined to appearances, the part of the soul that poetry appeals to will be unable to distinguish between what is merely acted and what is real—the aim of acting, after all, is to appear real—so it will indeed mistake a fictional monster for a real threat or stage tears for real tears. Thus, Socrates would deny (2) in such cases because he believes a part of us is naturally responsive to certain kinds of representations (representations that constitute passion-guiding beliefs) whether or not our rational part believes they represent something real.[43] The claim Socrates actually makes, however, is more controversial, but not implausible: we can 'suspend' not only existence, but also evaluative beliefs. Even though we believe a character's behaviour is reprehensible and unworthy of pity, we can admire and pity him.

Consider film again. It is easy to fail to fully appreciate just how morally dubious the generic 'heroes' of many enjoyable films are—think of *The Godfather* or *Pulp Fiction*. Behaviour well into the criminal is often glorified: debonair gangsters outwit dull, fumbling police; soldiers make the Geneva Convention look like liberal naïvety; hit-men kill without compunction, but with style. This is not to say that we wouldn't, if asked, agree that these actions are unambiguously wrong. But to enjoy a film, one can't be dominated by outrage at the actions of

[43] Note that if we exclude the parenthesis, his solution bears similarities to those that deny (1) by arguing that not just beliefs but other representational states can gives rise to desires and emotions. While the difference between denying (1) and (2) is more than merely linguistic for Plato—it is not for arbitrary reasons that he calls perceptual representations 'beliefs' in the *Republic*—the difference could lie solely in one's view of what counts as a belief, rather than in one's view of perceptual representations and their role in cognition.

its characters, as we might be in reality. Rather, we allow ourselves to be won over by the features of a story that the writer has made evaluatively salient and, since it's *just* a film, excuse it from any more objective, 'real life' moral appraisal. We sympathetically enter into, say, the Mafia moral code. As the narrative bids us to, we take the characters who stick by the code to be 'good' and those who break it to be 'bad' and, therefore, treat the criminality of the heroes lightly. Of course, even if we agree with Plato this far, not many people would agree that the consequences are as severely morally corrupting as he thinks and still less would agree that the solution is outright censorship. Nonetheless, while the *Republic*'s view of the arts is Draconian in degree, and perhaps reprehensible for this reason, this should not prevent us from entertaining the idea that its analysis of fiction could serve as a reasonable justification for the level of censorship, such as the rating system, that we do have.[44]

Fiction provides an example of one form of belief-discordant behaviour, but what Plato offers us is a theory of belief-discordant behaviour as such. Consider the following cases of belief-discordant behaviour described by Tamar Gendler:

> A sports fan watching a televised rerun of a baseball game loudly encourages her favourite player to remain on second base. A cinema-goer watching a horror film 'emits a shriek and clutches desperately at his chair'. A man suspended safely in an iron cage above a cliff 'trembl[es] when he surveys the precipice below him'. An avowed anti-racist exhibits differential startle responses when Caucasian and African faces are flashed before her eyes. In each of these cases, an experience or behavior is elicited that violates an apparent norm. At least in some sense of 'should' … the movie-goer should exhibit fear only in the face of actual danger, the anti-racist should react similarly to members of different racial groups. (2008b, pp. 552-553)

Like Smith's case of a phobic response to a hallucinated spider that I discussed at the end of chapter 3, these are all cases where our behaviour seems to be at variance with our beliefs, or at least with the beliefs that, as Smith put it, we would *say* we have or that we would rely on if we were to make a bet. Since actions typically become intelligible when we assign appropriate beliefs and de-

[44] For a detailed comparison between modern attitudes to television and Plato's view of poetry see Nehamas (1988).

sires to the agent (or, more strongly, become *bona fide* actions, rather than mere bodily movements), if we deny the possibility of conflicting occurrent beliefs or insist that there are only rational beliefs—beliefs that accord with our betting behaviour—then our behaviour in these cases is very difficult to explain. Yet in each of them it is very easy to point to the beliefs that *would* make our responses intelligible and rational: it would be perfectly rational to encourage a player to stay on second base if we believed that he was in front of us with his choice still to be determined; to tremble if we believed we there was a danger of falling to our death; or to react differently to Caucasian and African faces if we believed a relevant difference existed. A reasonable response, then, is to deny that all our beliefs accord with our linguistic and betting behaviour: in some situations our actions and emotions are guided by lower, non-rational beliefs that are insensitive to the kinds of evidence and reasoning that inform our higher, rational beliefs.

Gendler rejects this response on the grounds that 'whatever belief is … it is normatively governed by the following constraint: belief aims to "track truth" in the sense that belief is subject to immediate revision in the face of changes in our all-things-considered evidence' (p. 565). While no one would deny that beliefs typically do obey this norm, this is not by itself sufficient to show that they must do so universally or that to do so is part of what it is to be a belief (we can at least conceive of beliefs that are flagrantly recalcitrant in the face of contrary evidence: imagine a theist who agrees that the all-things-considered evidence favours atheism, yet 'just believes'). If instead of evidence-responsiveness, we place more weight on another, no less characteristic and near-universal feature of beliefs, namely the fact that they guide and render intelligible our actions and emotions, then we can reasonably take the cases Gendler lists as exceptions to the constraint she imposes.[45]

[45] It should be noted that while Gendler rejects the idea of conflicting beliefs in cases of belief-discordant behaviour, her own response is reasonably close in spirit to Plato's. She argues that these cases indicate that the traditional psychological repertoire—belief, desire, perception, imagination, and so forth—is too narrow and that a new mental state, which she calls 'aliefs', is required. While a component of this state bears similarities to belief, she characterises it as a 'lower' state that involves aspects of a number of different 'higher' mental states, making it an 'associative

Someone might reasonably respond by pointing out that while beliefs might be needed to make behaviour intelligible, evidence-responsiveness is needed to make *beliefs* intelligible: if we 'solve' Gendler's cases by positing a belief that is simply irrational, with no systematic regard for evidence whatsoever, then we make the behaviour intelligible at the expense of introducing a belief that is itself unintelligible. However, Plato's claim is not that some beliefs are insensitive to evidence *as such*, but that different kinds of beliefs are sensitive to different kinds of evidence. He denies that all our beliefs, desires, and cognitive abilities fall into one consistent, mutually supportive set. On the one hand, we have higher, reflective beliefs that are sensitive to reasoning and that guide our rational desires. On the other hand, we have rich perceptual representations—normal perceptions, sensory appearances, and perhaps also quasi-perceptual mental images—that are insensitive to reasoning and that guide our non-rational passions. These two kinds of belief usually do agree, since usually perception is accurate, but cases of belief-discordant behaviour reveal that this agreement is contingent. Nonetheless, when they do come into conflict, the beliefs (and desires or actions) on each side of the conflict, as well as the reason for the conflict, can be fully explained with reference to the different all-things-considered evidence available to each. Neither side of the conflict, then, is brutely irrational or inexplicable.

What emerges is a theory of irrationality. As Davidson would say, irrationality is always attributed against a given background of rationality. A belief or desire of the appetitive part may be seen as irrational if this background is our set of rational beliefs, yet against the background of the less complete evidence that is in fact available to it, it is perfectly rational (i.e. 'rational' in the sense of being

repertoire, one that paradigmatically includes not only representational (or 'registered') content, but also affective states, behavioral propensities, patterns of attentiveness, and the like.' (2008b, p. 559). She posits this 'associative repertoire' as something more primitive than, and conceptually antecedent to, more recent mental states like belief or desire (p. 558). A number of recent authors have argued persuasively that aliefs are an unnecessary innovation and that Gendler's project is better served if aliefs are broken into more familiar states like belief, desire, and perception. For example, Muller & Bashour argue that 'Gendler gets unwarranted conclusions about the existence of aliefs from belief discordant cases. She should, instead, have concluded that our beliefs do not necessarily form a consistent set' (2011, p. 386). This is precisely the solution I attribute to Plato. For a similar response to Gendler, see Currie & Ichino (Forthcoming, *Analysis Reviews*).

the appropriate response to the available evidence). Fearing a fictional monster seems irrational and unmotivated given our rational belief that the monster is a mere appearance, but fear *is* rational as a response to our non-rational belief that it is indeed, as it appears, a threat. Each part of us, then, is internally intelligible (at least to a reasonable degree: the non-rational parts of our soul will only be as consistent as sensory appearances are) with the result that seemingly inexplicable desires or behaviour can be rendered intelligible once we identify the appropriate part of the soul to attribute them to.

The claim that each part of the soul is internally intelligible can be put another way: to make a whole person intelligible in cases of conflict, we need to interpret each part of his soul as an agent in its own right. That is, certain instances of a person's behaviour, especially discordant behaviour, can be explained by a psychological story that is entirely specific to just one part of the soul, the kind of psychological story that we would usually apply to a whole person: from the information available, to a corresponding belief, to a desire this generates, and finally to an action this motivates.[46] Plato understands the parts of the soul in this way at least in part because he believes that all practical errors are at root intellectual errors, intellectual errors that I have characterised as the uncritical acceptance of appearances in a world in which appearance and reality fail to coincide. Plato does not believe, then, that any passions, however unruly, are brutely irrational, that is, passions that are not appropriate motivational responses to some belief the agent holds. Passions can be misguided only to the extent that our beliefs are misguided. This much is common to Plato's position in *Protagoras*, *Gorgias*, and *Republic*. The *Republic* differs in positing two new bearers of beliefs, an appetitive and a spirited part that are responsive only to sensory appearances.

[46] This is simply to say that we should identify as an agent—the source of an action—any subject that has sufficient cognitive and motivational resources to give rise to an action. For a detailed defence of the claim that this is how Plato sees the appetitive and spirited parts of the soul in the *Republic*, see Lorenz (2006, part 1). Note that, contra Santas (2010, pp. 81–89), this does not require one to attribute to each part of the soul the full spectrum of motivational and cognitive resources that we typically apply to a whole person. Crucially, it consistent with the appetitive or spirited part being entirely incapable of reasoning.

Nonetheless, the most interesting and original elements of Plato's account of how humans err, and why they err so much, remains essentially the same as what he described as falling prey to the 'power of appearance' in the *Protagoras* or relying on a knack of appearance-guessing in the *Gorgias*.

Bibliography

Adam, J., ed. (1902), *The Republic of Plato*, vol. 2 (Cambridge: Cambridge University Press) (cit. on pp. 108, 155, 197).

Annas, J. (1997), 'Understanding and the Good: Sun, Line, and Cave' in *Plato's Republic: Critical Essays*, Kraut, R. (ed.) (Lanham, Md: Rowman & Littlefield), pp. 143–168 (cit. on p. 150).

Armstrong, D. M. (1968), *A Materialist Theory of the Mind* (New York: Humanities Press) (cit. on pp. 127, 128).

Barney, R. (1992), 'Appearances and Impressions', *Phronesis* 37.3, pp. 283–318 (cit. on pp. 108, 114, 136).

Belfiore, E. (1983), 'Plato's Greatest Accusation against Poetry', *Canadian Journal of Philosophy* suppl. vol. 9, pp. 39–62 (cit. on pp. 170, 181).

Beresford, A. (Unpublished), 'The Strange Hedonism of Plato's *Protagoras*', url: `http://www.faculty.umb.edu/adam_beresford/research/hedonism-protagoras.pdf`.

Berman, S. (1991), 'Socrates and Callicles on Pleasure', *Phronesis* 36.2, pp. 117–140 (cit. on pp. 42, 75).

Bobonich, C. (2004), *Plato's Utopia Recast* (Oxford: Oxford University Press) (cit. on pp. 101, 176, 187).

— (2007), 'Plato on *Akrasia* and Knowing Your Own Mind' in *Akrasia in Greek Philosophy*, Bobonich, C. and P. Destrée (eds.) (Leiden–Boston: Brill), pp. 41–61 (cit. on p. 19).

Brickhouse, T. C. and N. D. Smith (2006), 'The Socratic Paradoxes' in *A Companion to Plato*, Benson, H. H. (ed.) (Oxford: Blackwell Publishing), pp. 263–277 (cit. on p. 5).

— (2007), 'Socrates On *Akrasia*, Knowledge, And The Power Of Appearance' in *Akrasia in Greek Philosophy: From Socrates to Plotinus*, Bobonich, C. and P. Destrée (eds.) (Boston: Brill), pp. 1–17 (cit. on pp. 5, 21).

Burnyeat, M. F. (1976), 'Plato on the Grammar of Perceiving', *The Classical Quarterly* 26.1, pp. 29–51 (cit. on pp. 101, 120, 123).

— (1999), 'Culture and Society in Plato's *Republic*' in *The Tanner Lectures on Human Values*, Peterson, G. (ed.), vol. 20 (Salt Lake City: University of Utah Press), pp. 215–354 (cit. on pp. 99, 104, 105, 119, 131, 160).

Carone, G. R. (2001), 'Akrasia in the *Republic*: Does Plato Change his Mind?', *Oxford Studies in Ancient Philosophy* 20, pp. 107–148 (cit. on pp. 15, 187).
— (2004), 'Calculating Machines or Leaky Jars? The Moral Psychology of Plato's *Gorgias*', *Oxford Studies in Ancient Philosophy* 26, pp. 55–96 (cit. on pp. 73, 84).
Cooper, J. M. (1999), 'Socrates and Plato in Plato's *Gorgias*' in *Reason and Emotion* (West Sussex: Princeton University Press) (cit. on pp. 70, 79).
Currie, G. and A. Ichino (Forthcoming), 'Aliefs don't exist, though some of their relatives do', *Analysis Reviews* (cit. on p. 205).
Devereux, D. T. (1995), 'Socrates' Kantian Conception of Virtue', *Journal of the Philosophy of History* 33.3, pp. 381–408 (cit. on pp. 5, 8).
Dodds, E. R. (1959), *Plato, Gorgias* (Oxford: Oxford University Press) (cit. on p. 56).
Dominick, Y. H. (2010), 'Seeing Through Images: The Bottom of Plato's Divided Line', *Journal of the History of Philosophy* 48.1, pp. 1–13 (cit. on pp. 150, 154).
Ferrari, J. (1990), '*Akrasia* as neurosis in Plato's *Protagoras*', *Proceedings of the Boston Area Colloquium in Ancient Philosophy* 6, pp. 115–139 (cit. on p. 8).
Fine, G. (1988), 'Plato on Perception: A Reply to Professor Turnbull, 'Becoming and Intelligibility'', *Oxford Studies in Ancient Philosophy* Suppl. Vol. Pp. 15–28 (cit. on p. 119).
— (1993), *On Ideas: Aristotle's Criticism of Plato's Theory of Forms* (Oxford: Oxford University Press) (cit. on p. 131).
— (2003), 'Knowledge and Belief in *Republic* V–VII' in *Plato on Knowledge and Forms* (Oxford: Oxford University Press), pp. 85–116 (cit. on pp. 150, 154).
Frede, M. (1987), 'Observations on Perception in Plato's Later Dialogues' in *Essays in Ancient Philosophy* (Oxford: Oxford University Press), pp. 3–8 (cit. on pp. 123, 125).
Ganson, T. S. (2009), 'The Rational/Non-Rational Distinction in Plato's *Republic*', *Oxford Studies in Ancient Philosophy* 36, pp. 179–197 (cit. on p. 117).
Gendler, T. S. (2008a), 'Alief and Belief', *Journal of Philosophy* 105.10, pp. 634–663 (cit. on p. 128).
— (2008b), 'Alief in Action (and Reaction)', *Mind & Language* 23.5, pp. 552–585 (cit. on pp. 203, 205).
Glüer, K. (2005), 'Perception and Justification (Unpublishd)' (cit. on pp. 120, 129).
Gosling, J. C. B. and C. C. W. Taylor (1982), *The Greeks on Pleasure* (Oxford: Oxford University Press) (cit. on pp. 29, 42).
Gulley, N. (1965), 'The Interpretation of "No One Does Wrong Willingly" in Plato's Dialogues', *Phronesis* 10, pp. 82–96 (cit. on p. 50).
Halliwell, S. (2002), *The Aesthetics of Mimesis* (Princeton: Princeton University Press) (cit. on pp. 131, 148).

Hamlyn, D. W. (1958), '*Eikasia* in Plato's *Republic*', *The Philosophical Quarterly* 8.30, pp. 12–23 (cit. on pp. 137, 141, 150).
Heck, R. G. (2000), 'Nonconceptual Content and the 'Space of Reasons'', *The Philosophical Review* 109.4, pp. 483–523 (cit. on p. 120).
Hemmenway, S. R. (1996), 'Sophistry Exposed: Socrates on the Unity of Virtue in the *Protagoras*', *Ancient Philosophy* 16.1, pp. 1–23 (cit. on p. 27).
Holland, A. J. (1973), 'An Argument in Plato's *Theaetetus*: 184–6', *The Philosophical Quarterly* 23, pp. 97–116 (cit. on p. 123).
Irwin, T. (1979), *Plato: Gorgias* (Oxford: Clarendon Press) (cit. on pp. 56, 89).
— (1995), *Plato's Ethics* (Oxford: Oxford University Press) (cit. on pp. 26, 151).
Jowett, B. and L. Campbell, eds. (1894), *Plato's Republic: The Greek Text*, vol. 3 (Oxford: Clarendon Press) (cit. on p. 197).
Kahn, C. (1998), *Plato and the Socratic Dialogue: The Philosophical Use of a Literary Form* (Cambridge: Cambridge University Press) (cit. on p. 7).
Klosko, G. (1983), 'The Insufficiency of Reason in Plato's *Gorgias*', *The Western Political Quarterly* 36.4, pp. 579–595 (cit. on p. 62).
— (1984), 'The Refutation of Callicles in Plato's *Gorgias*', *Greece & Rome* 31.2, pp. 126–139 (cit. on p. 42).
Lemmon, E. J. (1962), 'Moral Dilemmas', *The Philosophical Review* 71.2, pp. 139–158 (cit. on p. 8).
Lesses, G. (1987), 'Weakness, Reason, and the Divided Soul in Plato's *Republic*', *History of Philosophy Quarterly* 4.2, pp. 147–161 (cit. on pp. 187, 188).
Lorenz, H. (2006), *The Brute Within: Appetitive Desire in Plato and Aristotle* (Oxford: Oxford University Press) (cit. on pp. 101, 102, 108, 111, 146, 175, 206).
MacKenzie, M. M. (1981), *Plato on Punishment* (London: University of California Press) (cit. on p. 88).
Malcolm, J. (1962), 'The Line and the Cave', *Phronesis* 7.1, pp. 38–45 (cit. on pp. 150, 158, 160).
— (1981), 'The Cave Revisited', *The Classical Quarterly* 31.1, pp. 60–68 (cit. on p. 150).
Martin, M. G. F. (2002), 'The Transparency of Experience', *Mind & Language* 17.4, pp. 376–425 (cit. on p. 120).
Mastrangelo, M. and J. Harris (1997), 'The Meaning of Republic 606a3–b5', *The Classical Quarterly* 47.1, pp. 301–305 (cit. on p. 197).
McCoy, M. B. (1998), 'Protagoras on Human Nature, Wisdom, and the Good: The Great Speech and the Hedonism of Plato's *Protagoras*', *Ancient Philosophy* 18.1, pp. 21–39 (cit. on p. 27).
McTighe, K. (1984), 'Socrates on Desire for the Good and the Involuntariness of Wrongdoing: *Gorgias* 466a–468e', *Phronesis* 29, pp. 193–236 (cit. on p. 50).

Morris, M. (2006), 'Akrasia in the *Protagoras* and the *Republic*', *Phronesis* 51.3, pp. 185–229 (cit. on pp. 187, 188).

Moss, J. (2006), 'Pleasure and Illusion in Plato', *Philosophy and Phenomenological Research* 72.3, pp. 503–535 (cit. on pp. 187, 188).

— (2007a), 'The Doctor and the Pastry Chef: Pleasure and Persuasion in Plato's *Gorgias*', *Ancient Philosophy* 27, pp. 229–249 (cit. on pp. 47, 58, 62, 84, 89, 144).

— (2007b), 'What is Imitative Poetry and Why is it Bad?' in *The Cambridge Companion to Plato's Republic*, Ferrari, G. R. F. (ed.) (Cambridge: Cambridge University Press), pp. 415–444 (cit. on pp. 137, 170, 178, 180).

— (2008), 'Appearances and Calculation: Plato's Division of the Soul', *Oxford Studies in Ancient Philosophy* 34, pp. 35–68 (cit. on pp. 102, 144, 187).

— (2012), 'Pictures and Passions in Plato' in *Plato and the Divided Self*, Barney, R., T. Brennan and C. Brittain (eds.) (Cambridge: Cambridge University Press), pp. 259–280 (cit. on pp. 175, 183).

Muller, H. and B. Bashour (2011), 'Why Alief is Not a Ligitimate Psychological Category', *Journal of Philosophical Research* 36, pp. 371–389 (cit. on p. 205).

Nehamas, A. (1982), 'Plato on Imitation and Poetry' in *Plato on Beauty, Wisdom and the Arts*, Moravcsik, J. and P. Temko (eds.) (Totowa, NJ: Rowman and Littlefield), pp. 47–78 (cit. on p. 137).

— (1988), 'Plato and the Mass Media', *The Monist* 71, pp. 214–234 (cit. on pp. 137, 203).

— (1999), 'Plato on Imitation and Poetry in *Republic* X' in *Virtues of Authenticity* (Princeton: Princeton University Press), pp. 252–278 (cit. on pp. 104, 105).

Paton, H. J. (1922), 'Plato's Theory of EIKASIA', *Proceedings of the Aristotelian Society* 22, pp. 69–104 (cit. on pp. 137, 141, 150).

Penner, T. (1990), 'Plato and Davidson: Parts of the Soul and Weakness of Will', *Canadian Journal of Philosophy, Suppl. Vol.* 16, pp. 35–74 (cit. on pp. 8, 15).

— (1991), 'Desire and Power in Socrates: The Argument of *Gorgias* 466A–468E that Orators and Tyrants have No Power in the City', *Apeiron* 24.3, pp. 147–202 (cit. on p. 51).

— (1996), 'Knowledge vs. True Belief in the Socratic Psychology of Action', *Apeiron* 29.3, pp. 199–230 (cit. on pp. 8, 15).

— (1997), 'Socrates on the Strength of Knowledge: *Protagoras* 351B–357E', *Archiv für Geschichte der Philosophie* 79.2, pp. 117–149 (cit. on pp. 8, 15).

Price, A. W. (1995), *Mental Conflict* (London: Routledge) (cit. on p. 187).

Principe, M. A. (1982), 'Restraint of Desire in the Gorgias', *The Southern Journal of Philosophy* 20.1, pp. 121–132 (cit. on p. 87).

Radford, C. (1975), 'How Can We Be Moved by the Fate of Anna Karenina?', *Proceedings of the Aristotelian Society* 49, pp. 67–80 (cit. on p. 201).

Reshotko, N. (2006), *Socratic Virtue* (Cambridge: Cambridge University Press) (cit. on pp. 5, 15).

Rowe, C. (2007), 'A Problem in the *Gorgias*: How is Punishment Supposed to Help with Intellectual Error?' in *Akrasia in Greek Philosophy*, Bobonich, C. and P. Destrée (eds.) (Boston: Brill), pp. 19–40 (cit. on pp. 84, 85).

Rudebusch, G. (1999), *Socrates, Pleasure, and Value* (Oxford: Oxford University Press) (cit. on pp. 30, 42).

Russell, D. C. (2000), 'Protagoras and Socrates on Courage and Pleasure: *Protagoras* 349d ad finem', *Ancient Philosophy* 20.2, pp. 311–338 (cit. on p. 26).

Santas, G. (1966), 'Plato's *Protagoras* and Explanations of Weakness', *The Philosophical Review,* 75.1, pp. 3–33 (cit. on p. 5).

— (2010), *Understanding Plato's Republic* (Sussex: Wiley-Blackwell) (cit. on p. 206).

Scott, D. (1999), 'Platonic Pessimism and Moral Education', *Oxford Studies in Anciet Philosophy* 17, pp. 15–36 (cit. on pp. 62, 193).

Sedley, D. (2007), 'Philosophy, Forms, and the Art of Ruling' in *The Cambridge Companion to Plato's Republic*, Ferrari, G. R. F. (ed.) (Cambridge: Cambridge University Press), pp. 256–283 (cit. on p. 150).

Segvic, H. (2000), 'No One Errs Willingly: The Meaning of Socratic Intellectualism', *Oxford Studies in Ancient Philosophy* 19, pp. 1–47 (cit. on p. 7).

Silverman, A. (1991), 'Plato on *Phantasia*', *Classical Antiquity* 10.1, pp. 123–147 (cit. on p. 125).

Singpurwalla, R. (2006), 'Reasoning with the Irrational: Moral Psychology in the *Protagoras*', *Ancient Philosophy* 26.2, pp. 243–258 (cit. on p. 5).

— (2011), 'Soul Division and Mimesis in *Republic* X' in *Plato and the Poets*, Destrée, P. and F. Herrmann (eds.) (Leiden-Boston: Brill), pp. 283–298 (cit. on pp. 102, 187, 188).

Smith, A. D. (2001), 'Perception and Belief', *Philosophy and Phenomenological Research* 63.2, pp. 283–309 (cit. on p. 128).

Thayer, H. S. (1975), 'Plato's Quarrel with Poetry: Simonides', *Journal of the History of Ideas* 36.1, pp. 3–26 (cit. on p. 172).

— (1977), 'Plato on the Morality of Imagination', *The Review of Metaphysics* 30.4, pp. 594–618 (cit. on p. 172).

Vlastos, G. (1956), 'Introduction' in *Protagoras* (Indianapolis: Library of Liberal Arts) (cit. on p. 8).

— (1965), 'Degrees of Reality in Plato' in *New Essays in Plato and Aristotle*, Bambrough (ed.) (Routledge & Kegan Paul), pp. 1–19 (cit. on p. 133).

— (1969), 'Justice and Psychic Harmony in the *Republic*', *The Journal of Philosophy* 66.16, pp. 505–521 (cit. on p. 179).

Weiss, R. (1990), 'Hedonism in the *Protagoras* and the Sophist's Guarantee', *Ancient Philosophy* 10.1, pp. 17–39 (cit. on p. 27).

Weiss, R. (2006), *The Socratic Paradox And Its Enemies* (Chicago: University of Chicago Press) (cit. on p. 28).

Wilberding, J. (2004), 'Prisoners and Puppeteers in the Cave', *Oxford Studies in Ancient Philosophy* 27, pp. 117–39 (cit. on pp. 150, 151, 155, 160).

Wilson, J. R. S. (1976), 'The Contents of the Cave', *Canadian Journal of Philosophy* Supp. vol. II, pp. 117–127 (cit. on p. 179).

Woolf, R. (2000), 'Callicles and Socrates: Psychic (Dis)harmony in the *Gorgias*', *Oxford Studies in Ancient Philosophy* 18, pp. 1–40 (cit. on pp. 62, 73).

— (2002), 'Consistency and Akrasia in Plato's *Protagoras*', *Phronesis* 47.3, pp. 224–252 (cit. on p. 12).

Zeyl, D. J. (1980), 'Socrates and Hedonism: *Protagoras* 351B-358D', *Phronesis* 25.3, pp. 250–269 (cit. on p. 25).