

## Hedonism and the Divided Soul in Plato's *Protagoras*

Jessica Moss, May 2013

Penultimate draft. Paper forthcoming in *Archiv fur Geschichte der Philosophie*.

### 1. Introduction

In the *Protagoras*, Plato has Socrates premise his final argument on Hedonism, the doctrine that the pleasant is the good.<sup>1</sup> Does Socrates accept Hedonism? If so, how can we reconcile this with his views about the good elsewhere in the *Protagoras* and in other dialogues?<sup>2</sup> If not, why does he use Hedonism to argue for a conclusion that he evidently does accept, namely that no-one knowingly fails to do the good and thus that virtue is knowledge?

I want to show that to give a satisfactory answer to these questions we need to understand the role Hedonism plays in Socrates' overall argument against Protagoras: precisely which challenges it is introduced to meet, and precisely which claims Socrates takes it to entail. When

---

<sup>1</sup> It is Protagoras who first formulates a version of Hedonism: pleasant and good are the same (τὸ αὐτὸ ἡδὺ τε καὶ ἀγαθόν, 351e5-6). Socrates adopts this equivalence in his substitution argument at 355b ff, where he also treats as synonyms 'pleasure' and 'the good' (ἄλλο γὰρ ὄνομα μετείληφεν ἀντι τῆς ἡδονῆς τὸ ἀγαθόν) (351e5-6). Plato draws no attention to the difference between "the pleasant is the good" and "pleasure is the good", and as I think nothing important for our understanding of the argument rests on that difference I will for the most part ignore it.

<sup>2</sup> Socrates outright rejects Hedonism at *Gorg.* 468c, 474d, 475b, *Rep.* 505c, and throughout the *Philebus*; in the *Phaedo* he explicitly argues against the hedonic calculus recommended in the *Protagoras* (68e-69b). Hedonism also seems to conflict with a whole host of statements that he makes to the effect that virtue is the sole criterion one should look to in choosing how to act: see for example *Apol.* 28b, *Crito* 48c-d and *Gorg.* 479c and 480d. The implication of these passages is clear: what *makes* an action good is that it is just; things are good *because* they are just. Nowhere is there any hint that justice is good because it leads to pleasure; in the *Gorgias* Socrates even argues that justice is beneficial but not pleasant (by analogy with medicine, 478b). If we construe Hedonism as a view about bodily pleasures, as Socrates' examples in the *Protagoras* encourage us to do (see below), the conflict with other dialogues is even starker: avoiding injustice is worth the risk of death (*Apol.* 28b), the sacrifice of wealth (*Crito* 48c), and all manner of physical harm (*Gorg.* 480d and *passim*); moreover, the care of the soul is far more important than the care of the body, as we are told not only in other dialogues (e.g. *Apol.* 29e) but also in the *Protagoras* itself: the soul is "something you value [or *should* value, as is clearly implied] more than your body", and "everything concerning whether you do well or ill in your life depends on whether it [the soul] becomes worthy or worthless" (313a-b).

we do so, I argue, we see that Socrates introduces Hedonism as an attempt to save his views about virtue from a threat which Protagoras has posed to them on behalf of popular morality.

This is a threat implicit in Protagoras' view of moral education, brought to light by his account of courage at 351a, and presented in its starkest form by the Many's account of *akrasia* at 352b-d. If the psychology implied by these accounts is true, then a person's desires for pleasure can conflict with and even overwhelm her motivation to do what she judges best; therefore even if she knows which things are best she may of her own accord do bad; and therefore, contrary to Socrates' claim in this and other dialogues, virtue is not, or not just, a matter of knowledge. Popular psychology thus threatens Socrates' view of virtue by proposing that the human soul has motivations which can conflict with and even overwhelm its desire for the good. We can call this the Threat of the Divided Soul

In introducing Hedonism, I will argue, Plato is considering a response to this threat which centers on what we might call standards of value, or the ends of motivation: things pursued for their own sakes. He recognizes that both the good and the pleasant seem to be ends for the sake of which people act. He furthermore holds that if there is more than one end of human motivation – that is, if the pleasant is desired as an end distinct from the good – then so too must there be more than one *species* of human motivation. If the good is not the only end, then the reason-sensitive, reason-generated kind of desire Socrates thinks we have for the good is not the only kind of desire we have. Hedonism would neutralize this threat: if the pleasant just is the good then (given some assumptions about other putative ends) desires are all for the same end, and therefore are all of the same type. The soul is thus unified, and (given some assumptions about the nature of desire for the good) motivational conflict is impossible: we pursue only what we think best. It is Hedonism's monism about ends, then, which appeals to Socrates, for if Hedonism is true, his psychology and his account of virtue is saved.<sup>3</sup> Indeed, the dialogue implies a stronger claim: arguably Socrates now thinks that *only* if Hedonism is true is his psychology and account of virtue saved.

---

<sup>3</sup> Shields recognizes this connection between Hedonism, end-monism and what he calls "unified agency" (Shields 2007). Nussbaum also holds that Plato is interested in Hedonism as a form of end-monism, although she takes the point to be that only if there is just one end can all possible courses of action be measured against one another, and therefore can virtue be a matter of technical expertise (Nussbaum 1986, 109-117).

Moreover, although Socrates thinks and even hopes that Hedonism may be the true theory of the good, this is not the departure from standard Socratic ethics that it seems to be. On the face-value reading, Hedonism says that what is really good is a certain sensation we get from indulging our bodily appetites. I will argue that a proper understanding of Hedonism's role in the argument shows that Socrates must be implicitly arguing for the converse: that what we seek when we indulge our bodily appetites is the very thing we really get from acting virtuously – i.e. the good.<sup>4</sup> In other words, Socrates' version of Hedonism is a surprising view about pleasure, and not, as most interpreters have assumed, a surprising view about the good.<sup>5</sup> Irwin, for example, argues that in the *Protagoras* Socrates believes that what is really good is the very thing we already know under the name of pleasure.<sup>6</sup> I will argue that the text supports the opposite position: Socrates is considering the view that what is really pleasant is the very thing we already know under the name of good. 'Good' thus continues to apply to exactly the things to which Socrates usually applies it, while 'pleasant' shifts its application from what the Many think of as pleasant to the very things that Socrates usually calls good.

The interpretation that I defend has many rivals: others have argued that Socrates is a straightforward hedonist,<sup>7</sup> that he believes his interlocutors or "the Many" are hedonists and is arguing *ad hominem*,<sup>8</sup> or that pleasure is a place-holder for whatever turns out to be the right single, quantifiable end of the science of living.<sup>9</sup> My own interpretation has two main

---

<sup>4</sup> Rudebusch proposes a similar reading of the *Protagoras*' Hedonism in the last chapter of his 2002, although he gets there by a very different route; see note 39 below. I want to show, as he does not, that something like this reading is entailed by the *Protagoras* itself.

<sup>5</sup> Compare: "The best life turns out to be the life of the sea-tortoise", which is a surprising view about the best life rather than about the life of the sea-tortoise. If we accept this identity we learn lots of new things about the best life, e.g. that it involves swimming around all the time and keeping quiet; meanwhile we learn only one new thing, albeit a very important one, about the life of the sea-tortoise, namely that it is the best life that there is.

<sup>6</sup> Irwin 1995, 87-88.

<sup>7</sup> E.g. Adam and Adam 1905, Irwin 1995, Taylor 1996, and Gosling and Taylor 1982.

<sup>8</sup> E.g. Vlastos 1969, Zeyl 1980, and Kahn 2003.

<sup>9</sup> Nussbaum 1986, 121. A recent article by Dimas argues that Socrates' argument does not in fact rely on Hedonism, but only on the weaker claim with which he begins: that pleasure is good (351c-e) (Dimas 2008). Dimas shows successfully that this is all Socrates explicitly relies on in his argument that courage is knowledge at 358a ff. (see especially 358a6: "We agreed that the pleasant is good"), and raises very important questions about why Plato writes the dialogue this way. Nonetheless, the substitution argument

advantages. First, it shows the *Protagoras*' Hedonism to arise directly from concerns raised in the dialogue itself – and thereby gives us a new understanding of what precisely those concerns are. Second, although I will not treat the point in any detail here, it shows that attention to the *Protagoras*' argument can help us understand the philosophical motivation for the *Republic*'s radical new tri-partite psychology. For if I am right then in the *Protagoras* we see Plato working out the ideas that underlie that psychology: he is formulating the connections between ends of motivation, species of desire, and the nature of virtue. Through the Hedonist hypothesis he proposes that *if* the pleasant is the good, then there is only one species of desire, the soul is unified, and virtue is knowledge; in the *Republic* he rejects the premise, and all that follows from it.

## 2. The Role of Hedonism in the Argument

The *Protagoras* is an inquiry into whether virtue can be taught; Socrates tries to show that it can by arguing that all the virtues are identical, and that they are all knowledge. Toward the end of the dialogue Socrates has made some headway against his interlocutor: Protagoras concedes that most of the virtues form a unity, but insists that courage is distinct from the rest (349d). Socrates wants to convince him that courage is in fact, like the other virtues, knowledge (349e-350c), but Protagoras resists (350c-351a). The remainder of the dialogue is devoted to Socrates arguing that courage is knowledge, with a brief aporetic conclusion.

It is as a prelude to this final argument that Socrates begins questioning Protagoras about the relation between pleasure and the good. He asks if pleasure is good, i.e. if things are good insofar as they are pleasant;<sup>10</sup> Protagoras takes the question to be whether pleasant and good are the same,<sup>11</sup> and this is the question they proceed to investigate. In the argument that ensues, Socrates engages an imaginary interlocutor: “the Many” (*hoi polloi*). He undertakes to show that the Many accept Hedonism, and then to show that this commits them to an important thesis which they initially deny:

---

makes clear that it is the stronger identity thesis, Hedonism, that is crucial in the first stage of the argument (the argument that it is not possible to act against one's knowledge of the good); I will argue below that Hedonism is also required for the second stage of the argument (the argument that courage is knowledge).

<sup>10</sup> More precisely, Socrates first proposes that to live pleasantly is good (351b8); then that in the respect that things pleasant (*καθ' ὃ ἡδέα ἐστίν*), they are good (*κατὰ τοῦτο ἀγαθά*) (351c4-5); then that insofar as (*καθ' ὅσον*) things are pleasant they are good, 351e2-3); and finally that pleasure itself is good (351e3-4).

<sup>11</sup> Τὸ αὐτὸ ἡδὺ τε καὶ ἀγαθόν (351e4-5).

INTELLECT RULES: No appetite or passion can overpower knowledge or belief of what is best.<sup>12</sup>

We can call this argument, from Hedonism to Intellect Rules, Stage 1 of Socrates' final argument (352b-357e). Next, Socrates abandons the Many and directly addresses Protagoras and the other Sophists present. This is Stage 2 (358a-360d): he takes Intellect Rules as established, and argues that it entails his main thesis in the dialogue:

INTELLECTUALISM: Virtue (including courage) is knowledge of what is good and bad.

Notably Socrates does not conclusively endorse any of these theses: the dialogue ends aporetically. Moreover, Socrates most uncharacteristically goes out of his way to ensure that the Many accept Hedonism, and to emphasize that if they reject it his arguments will have no force against them.<sup>13</sup> It is clearly very important to him that he not sneak Hedonism into the

---

<sup>12</sup> Socrates first puts this as a claim only about knowledge (352c3-6), but after arguing for it takes himself to have shown that if Hedonism is true one cannot act against what one "knows or supposes (οἰόμενος)" to be better (358b7; cf. ἡγεῖσθαι, 358e7); thus I will use this broader formulation, with "intellectual judgment" meant to cover both knowledge and mere supposition or belief. This does not mean that the difference is unimportant: if only knowledge is stable, then only knowledge can be virtue, for those with true beliefs about the good may change their minds. For the purposes of this paper, however, we can ignore the difference, for as I will argue in section 5, Socrates' argument for the power of knowledge depends on a prior and more general thesis: that intellectual judgment about the good, where this includes both belief and knowledge, is the only species of motivation that we have.

<sup>13</sup> He asks them repeatedly to confirm not only that they value pleasure, but also that they value nothing else: "Does it not seem to you [...] that these things are bad on account of nothing other than the fact that they result in pain and deprive one of further pleasures?" (353e5-354a1); "Or are you able to mention some other end [...]?" (354b7-c1). The most remarkable point in this imaginary dialogue comes when Socrates is about to demonstrate the absurdity of the Many's account of *akrasia* by using 'pleasant' and 'good' interchangeably. Quite uncharacteristically, before trapping his opponents in a contradiction Socrates here warns them that he is about to do so, and gives them an opportunity to escape: "But even now it is still possible to withdraw, if you are able to say that the good is anything other than pleasure or that the bad is anything other than pain [...] [For if you are not], your position will become ridiculous [...]" (354e8-

conversation unnoticed, nor trick the Many into accepting it, nor in any other way hide the fact that he is attributing it to them. The effect of this tactic is to draw attention to the contingency of the argument on this one thesis: Socrates does not say “Hedonism is true, as I’m sure we all agree; therefore Intellect Rules”, but rather says, very explicitly, “*If* Hedonism is true – and I am working hard to see whether you the Many really do think it is! – then Intellect Rules.” What we have in these pages, although the fact is often ignored, is thus an ostentatiously conditional argument: *if* Hedonism is true, then Socrates’ view of virtue follows. We can outline this section of the dialogue as follows, with the arrows symbolizing indicative conditionals (‘if...then...’):

Hedonist hypothesis

Stage 1: Hedonism -> Intellect Rules

Stage 2: Intellect Rules -> Intellectualism

Aporetic conclusion

The question that has caused so much trouble is: why does Socrates argue for Intellect Rules, and thereby Intellectualism, by way of Hedonism? There are surely other ways to get there, most saliently simply by arguing, as Socrates does in the *Meno*, that all desires are for what one believes good (77b-78a).

A tempting suggestion is that Socrates is arguing *ad hominem*: he suspects Protagoras of Hedonism, and so adopts that view in order to refute him on his own terms (see e.g. Zeyl, cited above). But Protagoras has in no way avowed Hedonism earlier in the dialogue, and when Socrates begins questioning him on the relation of pleasure to goodness he initially denies even the weak thesis that pleasure is a good. Plato has not encouraged us to think Protagoras a hedonist.

I want to show that the argument is indeed tailored to Protagoras – and thereby to popular morality, of which Protagoras is a sophisticated representative – but in quite a different way. Plato presents Protagoras as committed to certain views about human motivation and its objects, views which present formidable threats to Intellectualism and which are implicit in particular in his denial that courage is knowledge. Socrates brings up the topic of how pleasure and goodness

---

355a6). Socrates, who is so often accused by his interlocutors of twisting words around and foisting on them a thesis that they never intended, is scrupulously avoiding that charge.

are related in response to these views. The hypothesis he then proposes, Hedonism, directly counters these views. Moreover, once Stage 1 is complete – that is, once Hedonism has done its argumentative work – Socrates proceeds as if these views have been refuted, helping himself to their contraries. Therefore, we should infer that Hedonism’s role in the argument is precisely to refute these views, and thereby to remove Protagoras’ reasons for resisting the Socratic view of virtue.<sup>14</sup>

In the next sections I identify these views. Socrates suspects that what underlies Protagoras’ resistance to Intellectualism is a belief in a plurality of ends (section 3), and a corresponding belief in a plurality of species of motivation in the human soul (section 4). Furthermore, Socrates thinks that all these views stand and fall together: if Protagoras is right that there is a plurality of ends, and therefore a plurality of motivation-species, then he is right to reject Intellectualism. The point of introducing Hedonism, then, is to hypothesize that all ends reduce to one, and thereby to salvage Intellectualism (section 5). In other words, Hedonism is put forth as a version of:

END MONISM: There is only one kind of thing people pursue as an ultimate end.

This still leaves open the question of why Socrates presents End Monism precisely in the form of Hedonism. In section 6 I argue that this is explained by Socrates’ conviction that Protagoras and the Many really do value pleasure, along with his hope that in doing so they are valuing nothing other than what he is also certain all humans value, the good.

### **3. The pleasant, the fine, and the beneficial**

As we saw above, what directly precedes the introduction of Hedonism into the dialogue is Protagoras’ argument that courage is different from the other virtues. Protagoras claims that courage must be different from wisdom in particular, because it is not a form of knowledge, but rather something analogous to bodily strength (351b). Instead of meeting these arguments head on, Socrates appears to change the topic:

Soc: Would you say, Protagoras, that some people live well and others live badly?

---

<sup>14</sup> For a broadly similar strategy see Klosko 1980, although the conclusions he draws are very different from mine.

Protag: Yes.

Soc: But does it seem to you that a person lives well, if he lives distressed and in pain?

Protag: No indeed.

Soc: Now, if he completed his life having lived pleasantly, does he not seem to you to have lived well?

Protag: It seems that way to me.

Soc: So, then, to live pleasantly is good (τὸ μὲν ἄρα ἡδέως ζῆν ἀγαθόν), and unpleasantly, bad?

Protag: Yes, so long as he lived having taken pleasure in fine things (τοῖς καλοῖς).

Soc: What, Protagoras? Surely you don't, like most people, call some pleasant things bad and some painful things good? (351b3-c2)<sup>15</sup>

Why does Socrates introduce the suggestion that living well is the same as living pleasantly? Why does he bring up the topic of pleasure at all, let alone begin to argue that all pleasant things are good? And why does he do so at this particular point in the dialogue, in response to Protagoras' account of courage?

We find our clue in Protagoras' reply to the suggestion. In adding the caveat that living pleasantly is good only if one takes pleasure in things that are fine (*kalon* – one might also translate the word as 'beautiful,' 'admirable,' or 'noble'), Protagoras is assuming that some pleasant things are *not* fine, and therefore not good. In other words, Protagoras here reveals himself to hold that there is potential divergence between two properties: the pleasant and the fine. Moreover, he assumes that the good (*agathon*) goes together with the fine, and therefore he holds that the pleasant can conflict with the good.

Next comes the introduction of Hedonism as a hypothesis, and then Stage 1, in which it does its argumentative work. Thus Hedonism enters the dialogue as a response to questions about the relation between the pleasant, good, and fine.

---

<sup>15</sup> Translations of the *Protagoras* are based, sometimes loosely, on Lombardo and Bell.

Moreover, Socrates also evidently takes Hedonism to entail some very strong theses about the relations between these three qualities. We see this by looking ahead to what he takes himself to have established by the time he gets to Stage 2. For here, apparently drawing on no further premises beyond those he employed in Stage 1 – that is, Hedonism and its alleged consequence, Intellect Rules – he helps himself to the following:

Are not all actions leading toward living painlessly and pleasantly fine and beneficial (ὠφέλιμοι)?<sup>16</sup> And isn't fine activity good and beneficial? (358b3-6)

If a thing is fine and good, is it also pleasant? (360a2-3)

Socrates cannot be implying that these equivalences are self-evident, for he is addressing Protagoras among others, and Protagoras clearly denied the convergence of the pleasant with the fine and good in the passage we saw above. Moreover, the context is a discussion of courage, and while it is uncontroversial to call courageous actions fine, it is wildly controversial to call them pleasant. Therefore Socrates must take these to be substantive conclusions entailed by what he has shown in Stage 1. In the course of arguing that the Many accept Hedonism and therefore must accept Intellect Rules, he takes himself to have shown what Protagoras initially denied: that the fine, the pleasant, and the good always coincide, and therefore that an action has one of these qualities if and only if it has all three. (In fact we have four terms here: fine, pleasant, good and beneficial. But as 358b suggests, and many other passages confirm, Plato tends to use 'good' and 'beneficial' almost interchangeably.<sup>17</sup> I will return to this point below.)

Thus what directly precedes the introduction of Hedonism in the dialogue is Protagoras' assertion that the pleasant can diverge from the fine and the good, and at least one of Hedonism's functions in the argument is to refute that claim. I want now to show that this is Hedonism's *main* function in the argument: the aim of Stage 1 as a whole is to establish that these three qualities cannot come apart. Socrates introduces the topic of the pleasant because he sees that Protagoras thinks that it diverges from these other two qualities, and suspects that this is the source of his anti-Intellectualism about virtue; Socrates wants to show that the three qualities always coincide because he takes this to be a necessary condition for Intellectualism. (In section

---

<sup>16</sup> Some editors delete "and beneficial" here, as threatening to make the next sentence redundant; whether Socrates makes the point once or twice, however, the passage clearly shows that he thinks all three qualities coincide.

<sup>17</sup> See especially *Gorg.* 499d2-3, and *Rep.* 335d and 379b; cf. Xenophon's *Mem.* IV.6.8.

5 I will argue that he furthermore thinks that the only way to show that they always coincide is to show that they are identical, and that this explains why he relies on Hedonism rather than a weaker thesis about convergence.)

We can begin to see the significance of these three qualities by looking at how they show up in the work of Aristotle, who here as so often takes on ideas from Plato but makes them more explicit. Both in the ethical works and in the *Topics*, an early work with clear Academic influence, he tells us this:

There are three objects of choices (τριῶν γὰρ ὄντων τῶν εἰς τὰς αἰρέσεις) and three of aversions: the fine, the advantageous, and the pleasant (καλοῦ συμφέροντος ἡδέος), and their opposites, the shameful, the harmful, and the painful. (*EN* II.3 1104b30-32)<sup>18</sup>

The fine, the pleasant and the advantageous (σύμφερον, often a synonym for ὠφέλιμον, beneficial) are the three objects of choice: if someone voluntarily undertakes some action, she must do so with a view to obtaining at least one of these three, and if she is averse to some action it must be because she thinks it will yield one of their opposites. These are the three properties which can non-instrumentally motivate us.

Unlike Aristotle Plato does not explicitly identify the pleasant, fine, and beneficial as objects of choice, nor does he explicitly introduce a term like Aristotle's 'objects of choice' (αἰρετά or τὰ εἰς τὰς αἰρέσεις). Nevertheless, in our section of the *Protagoras* we see him developing that notion, arguably for the first time in the history of philosophy. Asking whether the pleasant is the same as the good is a way of getting at the question Aristotle addresses in the above passage: what and how many are the criteria we look toward in choosing what to do? The way Plato articulates the question is as one about *ends*.

This emerges from a passage in Stage 1. Here the Many have claimed to be motivated by two distinct properties, the pleasant and the good; Socrates is trying to get them to acknowledge that they are in fact motivated only by one. This is how he puts it:

---

<sup>18</sup> In the *Topics* he identifies these as the three properties that are αἰρετά, choiceworthy or to-be-chosen (αἰρετόν ἐστι τὸ καλὸν ἢ τὸ ἡδὺ ἢ τὸ συμφέρον, 105a27); cf. 118b27 where these are three things for the sake of which (χάριν) other things are choiceworthy.

Are these things [medical treatment and other painful things the Many call good] good on account of anything else except that they end up (ἀποτελευτᾷ) in pleasures and escapes from and avoidances of pains? Or are you able to state some other end (τέλος), looking toward which you call these things good, other than pleasures and pains? (354b5-c2)

Does it suffice you to live out your life pleasantly without pains? If it suffices, and you are not able to call anything else good or bad, which does not end (τελευτᾷ) in these things, listen to what comes next [...] (355a2-5)

Here ‘end’ clearly has the sense of expected outcome or result: if you call things good because you expect them to end (τελευτᾶν) by producing  $x$ , then  $x$  is the end (τέλος) you look to in calling things good.<sup>19</sup> But Plato is also getting at a related idea: if we call an action good with reference to its expected result, then that expected result is our reason for choosing the action. Valuing an action because you expect it to end in  $x$  is valuing it for the sake of  $x$ .<sup>20</sup> In short, Plato is working toward a notion of the end as the that-for-the-sake-of-which, a notion that becomes explicit in a passage from the *Gorgias* (in which, notably, Socrates is arguing against Hedonism):

For if you recall, it seemed to myself and Polus that all things should be done (πραττέον) for the sake of the good. Do you agree with this, that the good is the end (τέλος) of all actions, and one should do all the other things for its sake (ἐκείνου ἕνεκα) but not it for the sake of other things? (*Gorg.* 499e6-9)

---

<sup>19</sup> Conversely, if you call things bad because you expect them to end by producing  $x$ , then  $x$  is the end you look to in calling things bad: Socrates goes on to claim that pain is the only end looking toward which the Many call bodily indulgences bad (354d1-2). Thus Plato here uses τέλος in the sense that Cicero will use its Latin translation in his *de Finibus Bonorum et Malorum*: “On the ends of” – that is, on the results looking toward which one evaluates actions as – “goods and bads”. There is an excellent discussion of this topic, drawing on *Protagoras* 354a ff., in Allen forthcoming.

<sup>20</sup> See also *Lysis* 220b3, where the good is characterized not only as the that-for-the-sake-of-which of friendship (φιλία, here arguably equated with desire in general), but also that in which all friendships result or end, τελευτῶσιν. See Allen forthcoming for discussion: he argues that there is a natural evolution from the standard sense of result to the novel philosophical sense of goal, for “in paradigm cases of action it will be the fulfillment or consummation of the action that both brings it to a close and was the reason or that for whose sake the agent undertook it in the first place.” Allen also argues persuasively that the sense of “expected result” is in the background of the *Gorgias* passage I quote below, by comparing that passage with *Gorg.* 466c-469e.

In this passage Dodds says we have “Perhaps the earliest clear instance of τέλος in the sense ‘purpose’, ‘end of action’, so common in later Greek from Aristotle onwards.”<sup>21</sup> But we can see now that the same idea is at work in the *Protagoras*: the claim in the *Gorgias* is that the good *should* be the end of all action; the suggestion in the *Protagoras* is that for the Many the pleasant *is* the end of all action.

More precisely, the *Protagoras* suggests that for the Many the pleasant is the end looking toward which they call things good; this turns out, however, to imply that it is also the goal of their actions. Recall that in the preceding lines Socrates has been using ‘good’ in the sense of ‘beneficial’.<sup>22</sup> Thus the claim of 354b-e is this. The Many report that they are moved (to very unequal extents) by two distinct concerns, pleasure and benefit/goodness, and thus that they sometimes value an action *qua* beneficial (i.e. good) even though they are averse to it *qua* painful, or value it *qua* pleasant even though they are averse to it *qua* harmful. Socrates presses them to declare that they call things good solely on the grounds that they expect them to be pleasant in the long term – that pleasure is their only criterion of benefit. If this is right, then even though they claim to evaluate possible actions by considering two distinct factors, these in fact collapse into one. Thus the claim that pleasure is the end looking toward which they call things good entails the claim that pleasure is the end of all their actions – the τέλος in the sense we get in the *Gorgias*. (Or rather, it entails that pleasure is the end of all the actions the Many claim to pursue for the sake of either pleasure or benefit. If we help ourselves to the further assumption that these exhaust their motives – and I will argue below that Socrates has thus helped himself, during Stage 1 – then we get the consequence that it is the sole end of all their actions.)

We can bring out the point more clearly by imagining how the Many might have answered Socrates’ questions at 354b-e in the negative:

*No, pleasure is not the end looking toward which we call actions good. We do of course care about pleasure, but when we consider whether an action is good we are ignoring pleasure and instead looking to something else. We are considering whether it will be beneficial or harmful,*

---

<sup>21</sup> Dodds 1959, commentary *ad loc.*

<sup>22</sup> This is particularly clear from the examples he gives of things the Many call good and bad, for medical treatment and the other painful things are thought good in the sense that they benefit us, and Socrates interchangeably refers to overindulgence in food, drink and sex as ‘bad’ (κακά) and as ‘ruinous’ (πονηρά) – one contrary of beneficial (see 353c7 with 353d5).

*by which we mean e.g. whether it will promote or inhibit the healthy functioning of our bodies; or we are considering whether the action will be fine or shameful, by which we mean whether it will merit honor or disgrace. And considerations of pleasure play no part in our judgments of things as beneficial or fine.*

If this were the Many's view, their psychology would be accurately described by Aristotle in the *Nicomachean Ethics* or *Topics* rather than by Socrates in the *Protagoras*: for them, the pleasant would not be the sole end of action, but only one out of three, alongside the beneficial and the fine.

Thus at 354b-e Plato is developing the notion that some quality of the expected outcome of one's actions serves as one's reason for choosing those actions – as one's end – and mounting a conditional argument that pleasure is the only such end. (I have chosen to adopt 'end' as the most literal translation of Plato's τέλος; 'value' or 'object of choice' would do as well in conveying the sense. 'End' should not therefore be taken to mean 'ultimate end': that  $x$  is one's end, in the relevant sense, means that one chooses at least some of one's actions for the sake of  $x$ , but not necessarily that one chooses all of them for the sake of  $x$ , nor necessarily that one never chooses  $x$  for the sake of anything else. In Aristotle's terminology, I mean to be talking about ends, but not necessarily about final (τέλεια – end-like) ends. On my use, therefore, the claim that the pleasant is an end leaves open the possibility that it may also be a means to some further end: the point is simply that it can motivate us in itself.)

But why does Plato have Socrates bring up this topic in the *Protagoras*, and at the particular point that he does? What in the passage that precedes Socrates' questions about the relation of the pleasant, fine and good – namely Protagoras' anti-Intellectualist account of courage – prompts those questions?

Let us consider why Protagoras should be so adamant that courage is not a form of knowledge. His claim is that courage is analogous to something quite different from knowledge, namely strength. He does not explicitly say how courage is like strength, but readers of Plato's other dialogues will know: courage is generally understood as a kind of perseverance or ability to withstand. What is crucial for our purposes is that Plato, in keeping with popular morality,

consistently identifies what courage enables one to withstand as pleasures and pains.<sup>23</sup> 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” (*Lach.* 191d6-e1); the courageous person is the one whose spirited part “preserves rational commands about what is to be feared or not through both pains and pleasures” (*Rep.* 442b11-c4); courage is “a fight against fears and pains” but also against “longings and pleasures” (*Laws* 633c9-d1). That is, courage is the ability to do the right thing even when one’s desire for pleasure or fear of pain makes this difficult. Moreover, the right thing is the fine thing, as contemporary morality would find obvious, and as Plato makes explicit when he returns to the topic of courage at 359e5. Thus courage will be the ability to choose what is fine, and shun what is shameful and bad, even when the fine is painful and the base pleasant. And thus Protagoras’ account of courage implies what his response to Socrates’ equation of living well and living pleasantly makes explicit: *there is often a conflict between the pleasant and the fine*. Virtue, Protagoras holds, involves the ability to resist the former for the sake of the latter.

Moreover, Protagoras’ statements imply that both the pleasant and the fine are ends. If courage is required to help us resist the pleasant for the sake of the fine (or rather the fine-and-good, since Protagoras assumes that those two go together) then evidently the pleasant is an end in itself. If it were not – if we wanted pleasant things only insofar as they resulted in the fine-and-good – then whenever the two conflicted we would happily abandon the pleasant, and there would be no need for psychological strength to resist it: if pleasure had only instrumental value, there would be no such thing as courage as Protagoras construes it. Furthermore, in arguing that we should reject pleasures insofar as they are not fine, Protagoras evidently also assumes that the fine-and-good is an end in itself: we care about it, and are motivated to pursue it.<sup>24</sup>

Thus implicit in Protagoras’ account of courage are two views about ends:

END PLURALISM: Each person by nature desires [at least two] different ends.

END CONFLICT: The different ends can diverge (some pleasant things are not fine-and-good, and vice versa).

---

<sup>23</sup> I thank John Cooper for helping me to see the relevance of this point.

<sup>24</sup> Compare Taylor: “Protagoras is to be seen as proposing a criterion of value which is [...] independent of pleasantness or unpleasantness” (commentary *ad* 351c1-2).

Notably, these are views that Plato takes to be very widespread: here as throughout the dialogue Protagoras is presented as a proponent of conventional morality (albeit an exceptionally sophisticated and articulate one). For conventional morality sees the lure of pleasure and the fear of pain as obstacles – indeed the chief obstacles – to virtue. As the poets say, the life of virtue is fine but harsh and toilsome, the life of vice base but pleasant and easy, and thus there is a constant temptation to choose the latter (*Rep.* 364a2-4; cf. 358a4-6).

Consider another passage from Aristotle on the three ends, this one from the opening of the *Eudemian Ethics*:

The man who set forth his own opinion in the precinct of the god at Delos composed an inscription for the forecourt of the temple of Leto distinguishing the good, the fine, and the pleasant as not all belonging to the same thing, writing: “Finest is the most just, and best is to be healthy, and most pleasant of all is to get what one loves.” But let us not agree with him. For happiness, being the finest and best of all things, is the pleasantest. (*EE* 1214a1-8; cf. *EN* 1099a24-28)

Popular morality, as represented by the Delian epigram, takes the three ends to diverge; philosophy, as represented by Aristotle, teaches that they in fact converge.<sup>25</sup> We have now seen that a similar debate is carried out in the *Protagoras*: popular morality, as represented by Protagoras, takes the pleasant to diverge from the others, while Socrates argues that they converge.

Thus we can conclude that Socrates finds implicit in Protagoras’ anti-Intellectualism about courage, and so gets him to confirm explicitly through questioning about the relation between the pleasant and the good (351c), two theses that he also attributes to popular morality: End Pluralism and End Conflict. When Socrates goes on to ask if pleasure is good, and then takes up Protagoras’ strong reformulation of the question as Hedonism (351e), we can now see that he is inviting Protagoras to reject precisely these two theses. I will return to this point below (section 5), but first I want to show what it is about End Pluralism and End Conflict that make them so worrisome to Plato.

---

<sup>25</sup> In this passage Aristotle claims only that what is most pleasant converges with what is most fine and most good; other passages make clear however that the three qualities always converge, or more precisely that what is *by nature* or “*without qualification*” (ἀπλῶς) pleasant converges with the other two. See for example *EN* 1099a11-15, and *EE* 1235b32-33 and 1236b27-1237a7.

#### 4. Ends and Motivations

A connection with motivation is built into the very notion of ends: for  $x$  to be an end for humans is for humans to tend to be non-instrumentally motivated by  $x$ . Thus we should not be surprised to see Plato drawing a connection between theses about ends and theses about motivation; he does so in the next phase of the discussion, the transition to Stage 1.

Protagoras, having espoused End Pluralism and End Conflict, is inclined to reject Hedonism, but proposes an inquiry into its truth. Socrates launches the investigation as follows:

All right, will this help to make things clear? When someone evaluates a man's health or other functions of the body through his appearance, having seen the face and extremities he might say: "Show me your chest and back too, so that I can make a better examination." That's the kind of investigation I want to make. Having seen that you have this position on the good and the pleasant, as you say, I need to say something like this to you: Come now, Protagoras, and reveal this about your mind: What's your position about knowledge? (352a1-b2)

Socrates here implies that Protagoras' stance on ends is a symptom of or related to another commitment, as the condition of a man's face and hands is related to his body. In particular, he takes these views about ends as possible symptoms of a view about knowledge, as follows:

Do you go along with the majority or not? Most people think this way about it [knowledge], that it is not a powerful thing, neither a leader nor a ruler. They do not think of it in that way at all, but rather in this way: while knowledge is often present in a man, what rules him is not knowledge but rather something else – sometimes *thumos* [spirit or temper], sometimes pleasure, sometimes pain, at other times *erôs*, often fear; they think of his knowledge as being utterly dragged around by all these other things as if it were a slave. (352b2-c2)

Socrates is questioning Protagoras' stance on what Aristotle will call *akrasia*. He asks if Protagoras endorses the contrary of Intellect Rules, namely:

INTELLECT ENSLAVED: Pleasure, pain, and other passions can overwhelm intellectual judgment (knowledge or belief) about what is good.<sup>26</sup>

Why does Socrates move from questions about ends to a question about the power of knowledge? The natural interpretation is that he takes Protagoras' avowal of End Conflict to imply a commitment to Intellect Enslaved. That Socrates sees the two theses as linked in this way is confirmed by something we saw above: at the start of Stage 2, having refuted Intellect Enslaved, Socrates takes himself to have refuted End Conflict as well, for he now takes as given that actions are pleasant if and only if fine, and fine if and only if good (358b, 360a). Thus Socrates is saying, when he asks Protagoras to "show his chest and back", that if Protagoras wants to maintain his belief in End Conflict he will have to endorse Intellect Enslaved as well.

Obviously it is Intellect Enslaved that Stage 1 is directly concerned to refute: Socrates' aim in this part of the argument is to show that what seem to be cases in which someone's intellect is overcome by pleasure are in fact cases in which she does what she falsely believes best. But Socrates does not attempt to refute this thesis directly. Instead, I will show, he does so by refuting two other theses, presuppositions of Intellect Enslaved. These are theses that are never made explicit in the argument but that Socrates takes Protagoras to have embraced. First, a thesis directly implied by Intellect Enslaved:

MOTIVATIONAL CONFLICT: Pleasure, pain, and other forces can conflict with intellectual judgment about what is good.

Second, a thesis equally obviously implied by this:

MOTIVATIONAL PLURALISM: There are distinct and potentially warring species of motivation in the human soul.

Pleasure, fear and the rest can overwhelm intellectual judgment about the good only if they can conflict with it, and they can conflict with it only if they are distinct from it. Protagoras' account of courage thus presupposes not only a thesis about objects of desire but also a thesis about the

---

<sup>26</sup> As I mentioned in the note to Intellect Rules above, here Socrates mentions only knowledge, but what he says at 358b-d shows that he takes himself to have argued against the broader claim about intellectual judgment, i.e. knowledge or belief.

nature of the soul: that alongside and independent of intellectual judgment about the good it also contains other forces which can determine how we act. This is what I called, in the Introduction, the threat of the divided soul.

In fact, Socrates has reason to believe that Protagoras has assumed this picture of the soul all along. First, it is the view of popular morality: witness the passage from Homer quoted in the *Republic* to illustrate motivational conflict between reason and *thumos*, “He struck his chest and spoke to his heart” (*Od.* xx.17-18, cited at *Rep.* 441b). Second, it evidently formed the background of Protagoras’ major positive contribution to the dialogue, his “Great Speech” on moral education (325d-326e): Protagoras’ program for inculcating virtue has many parallels with the childhood education of the guardians in Books II and III of the *Republic*, and like that program clearly assumes that there are elements in the soul best influenced and shaped by non-intellectual means. But it is not until his speech on courage that Protagoras explicitly appeals to a division in the soul. Here he names *thumos*, *mania* and *phusis* as forces in the soul, contrasting these with knowledge. He argues that we can account neither for courage nor for its lesser counterpart, confidence, only in terms of what the agent knows or believes: there are non-intellectual elements in the soul which contribute to or even constitute courage and other motivational qualities.

Socrates thus takes Protagoras’ avowed positions on ends to entail a psychological theory about motivation, one Socrates also finds in popular morality, and makes explicit in the speech he attributes to Many on the impotence of knowledge. In broad strokes, we can represent the inference Socrates draws as follows:

{End Pluralism & End Conflict} > {Motivational Pluralism & Motivational Conflict & Intellect Enslaved}<sup>27</sup>

Why would Socrates assume that these two sets of views go together? Consider a passage on the *Phaedrus* which puts forth a psychology similar to the one Socrates here attributes

---

<sup>27</sup> I mean to leave open for now the question of how the theses within the brackets are related, and thus of whether the motivational theses are entailed by End Pluralism, or only by End Conflict, and also of whether the theses about ends directly entail all three motivational theses or merely Motivational Pluralism, with additional premises needed to generate Motivational Conflict and Intellect Enslaved; I address all these questions in the next section.

to the Many (this one also presented as a tenet of popular opinion: it comes from Socrates' first, Lysianic speech):

There are two species (ιδέα) that rule and lead in each of us, which we follow wherever they may lead, one being the inborn appetite for pleasures, and the other an acquired belief which aims at the best. (*Phaedrus* 237d6-8)<sup>28</sup>

This psychological theory admits of two ends – the pleasant and the best (or good) – and correspondingly of two distinct species of motivation: inborn appetites, and acquired beliefs.<sup>29</sup> In other words, ends are correlated with species of motivation. This is a pattern we find throughout the corpus. In the *Charmides*, for instance, the desire for pleasure is called appetite again, and the desire for good wish (βούλησις); *erôs* is defined as being for the fine, and fear as being of the fearful (*Charm.* 167e). In the *Republic*, our desires for pleasure and gratification are appetites which arise from “passions and illnesses”, while our motives to pursue what is best have another source altogether, calculation (λογισμοῦ) (*Rep.* 439d). And in the *Protagoras*, in the theory attributed to the Many at 352b-c, the motivation to pursue the good is called “knowledge”, while what drags us toward pleasure is “something else”: *thumos*, *erôs*, or one of the other members of a group which he elsewhere calls the passions (see *Tim.* 69c-d).

The idea behind all these distinctions seems to be that if there is more than one object of human motivation, so too will there be more than one kind or species (ιδέα, as the *Phaedrus* passage puts it) of motivation itself. It is worth noting that the inference is not inevitable: one could in principle hold that there are different ultimate objects of desire (e.g. pleasure and benefit) while also holding that there is only one *type* of desire (e.g. only instinctual drives, or only calculation-sensitive preferences). In all the passages quoted above, however, Plato evidently assumes that the motivation to pursue what is best is different in kind from the motivation to pursue pleasure. In other words, he individuates species of motivation by their objects. Possibly he has in mind the principle he states in distinguishing knowledge, belief and ignorance in *Republic* V: different powers (δυνάμεις) are set over (ἐπί) different things (*Rep.* 477d). (Notably,

---

<sup>28</sup> ἡμῶν ἐν ἑκάστῳ δύο τινέ εἰσιν ἰδέα ἄρχοντε καὶ ἄγοντε [...] ἡ μὲν ἔμφυτος οὕσα ἐπιθυμία ἡδονῶν, ἄλλη δὲ ἐπίκτητος δόξα, ἐφιεμένη τοῦ ἀρίστου.

<sup>29</sup> To call belief (δόξα) a kind of motivation sounds odd to our ears, but, I submit, did not to Plato's; in the lines just quoted from the *Phaedrus* belief “aims at” the best, and a few lines below it “urges” (ὀρμώσης) us toward what is right (*Phaedr.* 238b8).

if this is his view then the converse entailment will also hold: if there are multiple species of motivation there will be multiple ends. I return to this point below.)

If we take this picture of motivation to be at work in the *Protagoras* we can explain Socrates' move from questions about ends to questions about motivations: since the Many avow belief in more than one end, he infers that they must believe in more than one type of motivation as well. Moreover, since they think the different ends can conflict, they must believe the same of the corresponding motivations. Plato takes Protagoras' and the Many's embrace of End Pluralism and End Conflict to entail a commitment to Motivational Pluralism and Motivational Conflict – and that is why, after the discussion of courage and the questions about the relation between the pleasant, fine, and good, he has Socrates broaden the investigation with a discussion of *akrasia*. If the pleasant is distinct and divergent from the fine and good, the desire for the pleasant is distinct from and potentially in conflict with the impulse toward the fine and good.

Now let us look at what has happened by Stage 2 to all the theses Socrates initially attributed to the Many. First, consider the fate of their theses about ends. As we saw above, Socrates now takes End Conflict to have been refuted: the pleasant, fine and good all converge (358b, 360a). As to End Pluralism, this is directly denied by the main premise of the argument, Hedonism. If the pleasant is the good, and the good is the fine (the fine-and-good), there is only one end. To put it another way, if people think of themselves as only wanting either what they think pleasant *or* what they think good, and then it turns out that pleasure is the only end (τέλος) toward which they look in evaluating their actions as good, then pleasure is their only end. Hedonism is a version of End Monism, and thus automatically entails the denial both of End Pluralism and End Conflict.

What about the Many's theses about motivation? Just above we saw evidence that Socrates takes these to be entailed by their theses about ends; now we will see that he also holds the converse (as indeed his apparent individuation elsewhere of species of motivation by their objects implies). For in arguing for the contraries of the Many's theses about ends, it emerges, he takes himself to be establishing the contraries of their motivational theses as well.

Clearly he takes himself to have refuted Intellect Enslaved: that was the main point of Stage 1. Only slightly less obviously he also takes himself to have refuted its direct presupposition, Motivational Conflict: Socrates argues not only that the desire for pleasure cannot

overwhelm knowledge or belief of what is good, but more strongly that the two never come into conflict at all. (We can call this thesis Motivational Harmony.) That is, the Many's alleged experience of being torn between the two motivations is merely a figment of their confused self-analysis. The coward, for example, is averse to standing his ground in battle because he thinks it more painful than running away; he claims to have an intellectual judgment opposing his fear, one that tells him that running away is worse, but in fact he has no such judgment, since in thinking running away more pleasant he also thinks it better. He simply wants to run away, and has no conflicting motivation at all to stand his ground.

What is less obvious, but on close inspection remarkable, is that Socrates evidently also takes himself to have refuted even the weakest of the motivational theses, Motivational Pluralism. This is in fact what we should expect given my argument that Plato consistently correlates objects and species motivation; the expectation is confirmed by Socrates' characterization in Stage 2 of fear:

I call fear a kind of expectation (*προσδοκίαν τινὰ*) of the bad [...] What one fears one supposes (*ἡγεῖσθαι*) to be bad. (358d6-e5)

In stating the Many's belief in Motivational Conflict (352b-c) Socrates contrasted fear, alongside other passions (*thumos*, pleasure, pain and *erôs*), with intellectual judgment about the good. Now he takes it as uncontroversial that fear is a belief or supposition about what is bad. (Note the root *δόξα*, belief, in *προσδοκία*, and the unexplained assumption that fears track suppositions. I argue more fully for this claim in the next section, and also consider a weaker interpretation: that fear merely depends on belief about what is bad.) Thus fear is now shown to be simply one manifestation of the kind of motivation the Many thought opposed to it: intellectual judgment about what is good and bad. If someone is fearful of something she claims to know or believe good, or if she has an appetite for something she claims to know or believe bad, her claims cannot be accepted: in fearing something she is manifesting her belief that it is bad, and in having an appetite for something she is manifesting her belief that it is good. (For compelling defenses of this proto-Stoic interpretation of Socrates' account of fear, see both Segvic and Frede.<sup>30</sup>) Socrates has now rejected Motivational Pluralism and is propounding its contrary:

MOTIVATIONAL MONISM: There is only one species of human motivation.

---

<sup>30</sup> Segvic 2000 and Frede 1992, xxix-xxx.

Thus Socrates' argument in Stage 1, his argument from Hedonism to Intellect Rules, works – again in broad strokes (I consider the details below) – as follows:

{Hedonism & End Monism & End Convergence} > {Motivational Monism & Motivational Harmony & Intellect Rules}

He takes pluralism about motivation to entail pluralism about ends, and pluralism about ends to entail pluralism about motivation; he takes the two sets of views to stand or fall together. He hopes that Motivational Pluralism is false, and so endeavors to show that his interlocutors' apparent commitment to End Pluralism is merely apparent.

### **5. Motivational Monism**

To an extent, then, we have secured an answer to the question with which we began. Plato has Socrates premise his final argument in the *Protagoras* on Hedonism because Hedonism is a form of End Monism, and therefore entails Motivational Monism and thereby Intellect Rules. But this account raises more questions than it answers.

First, the argument looks like overkill. Why bother arguing against Motivational Pluralism and Motivational Conflict if one's target is the weaker thesis Intellect Enslaved? That is, why not concede to common sense that there are multiple and even conflicting species of motivation, while insisting that intellect always wins out? Even granting that this would be a hard case to make, and therefore that Intellect Enslaved may be best undermined via its presuppositions, why not grant common sense at least Motivational Pluralism and rest with an argument against Motivational Conflict? Correspondingly, is it necessary to deny End Pluralism, or would an argument against End Conflict have done just as well? That is, why not allow that there are multiple ends, and therefore multiple species of motivation, while denying that they can conflict?

Second, even allowing that End Monism is, as the strongest thesis in the neighborhood, an expedient premise for Socrates' argument, we are left with no more illumination than we had at the outset as to why he should present this thesis in the form of Hedonism.

I turn to this second question in the next section; let us begin here with the worry about overkill. Socrates could argue directly for Intellect Rules: he could allow Motivational Pluralism and Motivational Conflict but insist that when motivations conflict, knowledge always wins out. Given that the Many think they can refute this claim empirically, however, we can see why he might want to start further upstream: better to show them that they are conceptually confused. In that case, could he merely attack Motivational Conflict, allowing that intellectual judgment is distinct from the desire for pleasure but denying that the two can conflict? Again, there is the worry about empirical counter-arguments. But there is also another factor, which shows that he does need to establish Motivational Monism, and indeed Motivational Monism of a very particular variety.

Consider the argument of Stage 2, which takes us from Intellect Rules to the conclusion that courage is knowledge. Protagoras and the Many think that fear of pain sometimes motivates us to avoid what we judge best (an instance of Motivational Conflict); hence Protagoras' conviction that courage must be something distinct from knowledge of what is best. Socrates refutes this view as follows:

1. No-one willingly goes for what they think or suppose (οἶεται, ἠγεῖται) bad (358c6-d2, 358e5-6).
2. Fear is a kind of expectation (προσδοκίαν) that something is bad (358d6-7).
3. Therefore (from 2), one fears only what one supposes (ἠγεῖσθαι) bad (358e4-5).
4. Therefore (from 1 and 2) courageous people do not differ from cowards in going for what they fear, for no-one does that (358e2-6, 359d1-4).
5. Therefore (from 4) the difference must lie in what they fear: courageous people do not fear noble death (*sc.* for they do not falsely believe it bad) (359d7-e7).
6. Therefore (from 5) cowardice is ignorance, and courage knowledge (360c6-d5).

It is crucial to this argument not merely that fear can never overcome one's intellectual judgment about what is good and bad, but that fear *is* such a judgment: we should read premise 2 as defining fear as a judgment of something as bad. Otherwise Socrates would have provided no argument for premise 3 – no argument against Protagoras' implicit claim that one can fear what one judges good – and therefore no argument against Motivational Conflict. Therefore he would need a direct argument for Intellect Rules: he would need to show that when fear and intellect conflict, intellect is somehow guaranteed to win out. This would

involve psychological factors like non-rational persuasion or force which he nowhere mentions here. For if fear is not itself a belief about the good (or directly dependent on one – see below), it will not be directly affected by beliefs about the good, and so to prevent cowardly behavior some other means must be found of bringing fear into line.

Moreover, even if Socrates were able to give a direct argument for Intellect Rules, in the face of the Many's purported first-person experience, it would not entail Intellectualism. Courage would not be identical with knowledge; instead it would be something much closer to what popular morality thinks it: the ability to withstand fear of things one judges fine and good.

Thus Socrates' argument that courage is knowledge depends not merely on Intellect Rules, but also on its stronger presupposition, Motivational Monism. Furthermore, it depends on a distinctive version of Motivational Monism, which we can call:

INTELLECTUALIST MOTIVATIONAL MONISM: The only species of human motivation is intellectual judgment about the good and bad.

If the thesis were instead that all motivations are, say, unreasoned passions, then Stage 2's argument would fail: without premise 2, premise 1 would be irrelevant and premise 3 unjustified.

It is worth noting an alternate reading on which Socrates holds a slightly modified version of Intellectualist Motivational Monism: the only species of human motivation is motivation that directly depends on and tracks intellectual judgment about the good. (That is, fear is not itself a judgment, but some other state that accompanies and depends upon judgment.) This seems a less charitable reading of the text: what role would premise 2 then play, and what would be the justification for premise 3? Moreover, the main argument for this alternate reading – that judgment and motivation belong to distinct categories – would not have impressed Plato: as we saw from the *Phaedrus* passage quoted above, he is happy to speak of belief (*δόξα*) aiming and urging (*Phdr.* 237d, with 238b, quoted in note 29 above). For Socrates' overall aim of establishing Intellectualism, however, either view would do, and it is probably fair to say that Plato is not here (or anywhere) particularly concerned to distinguish the two. My main point is that Socrates now presents fear as being precisely as closely in line with intellectual judgment as are the motivations that the Many initially recognized themselves as having toward the good. Whether fears *are* intellectual judgments or are merely dependent on them, they track our

judgments about what is good; therefore for our purposes they can be called intellectual motivations, the kind that pose no threat to Intellect Rules or Intellectualism. For simplicity in what follows, I will simply speak of intellectual motivations as identical with judgments.

Given that Stage 2's argument depends on Intellectualist Motivational Monism, and given Plato's assumption that species of motivations are individuated by their objects, we can now see why he has Socrates premise Stage 1 on End Monism rather than merely on End Convergence. If the pleasant is distinct from the good then the desire for the pleasant is independent from intellectual judgment about the good; therefore the question of whether ends can conflict becomes irrelevant to the question of whether motivations can. The trouble is that even if the two properties always converge, someone might fail to recognize that fact; if he values both as independent ends he will therefore have conflicting motivations. To illustrate this point, consider the coward. Let us grant Socrates that standing one's ground is both good and pleasant. The trouble is that someone who values the good and the pleasant as distinct ends might falsely believe standing one's ground painful even while rightly believing it good; therefore he will fear it *qua* painful even while having an intellectual motivation to pursue it *qua* good. If however valuing an action as good is simply identical with valuing it as pleasant, he can never fear what he judges good, never experience that kind of motivational conflict.<sup>31</sup> (In section 6 I will consider how this identity could work.)

Furthermore, if we take Socrates to be assuming Motivational and End Monism in Stage 1's argument we can explain an odd feature of that argument which I have so far ignored. As we saw above, the Many express Motivational Pluralism in an extreme form: intellectual judgments about the good exist alongside not only desires for pleasure but a whole jumble of motivations – fear, *thumos*, *erôs*, and perhaps others. A few lines later Socrates refers back to this whole list: the Many, he says, claim that they act against their knowledge of what is best when they are overcome by “pleasure or pain or are being ruled by one of the things I mentioned just now” (352d). But in the very next sentence, he reduces this whole list to one: he refers to “this experience which they [the Many] call being overcome by pleasure” (352e). And in what follows, he alludes only to the desire for pleasure and the avoidance of pain: *thumos*, *erôs* and fear have dropped out of the picture, evidently subsumed under these two motivations. This is

---

<sup>31</sup> This on the assumption that genuine goods do not conflict, an assumption which Socrates' faith in the measuring art's saving power strongly suggests that he holds in the *Protagoras*. For an argument that Plato's later work questions this assumption see White 1999.

odd: in the *Charmides*, for instance, while appetite was for pleasure, *erôs* was for the fine, and fear of the fearful (*Charm.* 167e). But if we take Socrates' aim to be the securing of Intellectualist Motivational Monism, we see the explanation. He is reducing all the putatively non-intellectual motivations in the soul to intellectual ones, and he does so in two steps: (1) all such motivations are in fact desires for pleasure or aversions from pain; (2) desires for pleasure and aversion from pain are in fact intellectual judgments about the good.<sup>32</sup>

Now we have addressed the overkill worry: Socrates can best defend Intellect Rules (and thereby Intellectualism) on the basis of Motivational Monism, and therefore he needs to premise the whole argument on the radical thesis of End Monism. But the realization that the only form of Motivational Monism that will do is the Intellectualist form – the thesis that we are always motivated by intellectual judgments about the good – only makes more pressing our second worry: why does he choose the particular form of End Monism he does choose, Hedonism?

## 6. Hedonism

The general mystery about Socrates' reliance on Hedonism arises, as we saw in section 1, because the thesis seems so un-Socratic. It seems so un-Socratic because the straightforward way to understand the phrase “the pleasant is the same as the good” – and surely the way the Many are meant to understand it when Socrates urges them to endorse it – is as a reduction of the good to the pleasant. The claim seems to be: nothing is good other than what you think of as pleasant, namely bodily gratification. And the variant Socrates introduces in the substitution passage, that “pleasure and the good are the same”, is most naturally understood as: what it is to be good is to be the tingly feeling or sense of gratification you think of as pleasure. Moreover, these readings seems outright entailed by the examples of pleasures Socrates offers in his dialogue with the Many, for these are all bodily pleasures: food, drink and sex (353c), the paradigm objects of appetites (*ἐπιθυμίας*) in *Republic* IV. There is not even any mention of the higher pleasures we see introduced in *Republic* IX as pleasures of reason (“knowing where the truth lies... while learning” (*Rep.* 581d)) or in the *Philebus* as “true pleasures” (pure aesthetic pleasures, or the

---

<sup>32</sup> In section 3 I noted that we can move from “Pleasure is the end looking toward which the Many call actions good” to “Pleasure is the sole end of all their actions” *if* we establish the further premise that the Many only want things *qua* pleasant or good (or rather, *qua* pleasant, i.e., *qua* good). Here we see Socrates establishing this premise. The Many cannot say “I sometimes want something out of anger, without thinking it pleasant or good”, for anger, along with all other motivations, has here been reduced to desire for pleasure or aversion from pain.

pleasures of learning (*Phil.* 51b-52a)). Thus the obvious face-value reading of the Hedonist hypothesis is this: what most people already think of as pleasant turns out to be the only thing that is good.

What we have seen above about the inferences Socrates draws in Stage 2, however, implies two reasons for doubting that this is what Socrates himself has in mind.

First, this reading makes his denial of End Conflict in Stage 2 (the claims that what is pleasant is fine and good, and vice versa, at 358b and 360a) a complete *non sequitur*. Protagoras and the Many began by contrasting the pleasant on one side with the fine-and-good on the other. Stage 1 makes no mention of the fine at all: it is entirely concerned with the relation between the pleasant and the good. If the hypothesis adopted there is that the good turns out to be nothing other than what they already know as the pleasant, then in the absence of any argument about the fine, the natural inference would be that the good sometimes diverges from the fine. In other words, the Many would be left with the position Polus takes in the *Gorgias*: vicious acts are less fine than their virtuous alternatives, but often more pleasant and therefore better.<sup>33</sup>

Polus holds this view about unjust acts; we might well expect the Many to hold it about cowardly acts like running away from battle. Previously they called that pleasant but bad; they called it bad because they associated the good with the fine and the bad with the shameful, and shared Socrates' assumption – one nowhere questioned in this dialogue – that virtue is fine and vice shameful. If, as Socrates persuades them, they mean by 'bad' nothing other than 'less pleasant overall,' two views are open to them: either running away is overall painful and thus bad, or it is overall pleasant and thus good. Only in the first case will the pleasant and good converge with the fine. But – and this is why cowardly acts are a particularly striking example – *if*, as on the straightforward reading of Hedonism, 'pleasant' retains its familiar meaning, it is very hard to argue that running away is more painful than standing one's ground, and thus very hard to argue that it is worse. Some vicious acts, like intemperate overindulgence in bodily pleasures, are recognizably painful in the long run and thus by the Hedonist hypothesis bad; therefore Hedonism does guarantee that *some* fine acts, like temperate behavior, will also be pleasant and good. But the present case is different: in courageously standing one's ground one risks plenty of bodily

---

<sup>33</sup> *Gorg.* 474c-475e. See Moss 2005 for an argument that Polus thinks committing injustice better just *because* he thinks it more pleasant. In any case he clearly opposes the more beneficial and more pleasant to the finer.

pain, and even death. If ‘pleasant’ means what the Many mean by it, then running away is more pleasant than courageously standing one’s ground.<sup>34</sup> Therefore Hedonism will not preclude End Conflict: running away will be less fine, because cowardly, but pleasanter and thus better. And therefore on the straightforward reading of Hedonism Socrates’ overall argument simply fails: knowledge of the good is no barrier to cowardice, and so courage is not knowledge.

In fact, Socrates’ denial of End Conflict, and specifically his claim that courageous acts are pleasant, seem so obviously counter to common sense, and also so clearly questionable as inferences from Hedonism, that we should take them as signs that Socrates’ characterization of Hedonism may have been deliberately misleading.

This is confirmed by a second oddity in the argument when taken at face value. I argued above that Socrates asserts Hedonism as a form of End Monism, with the aim of establishing Motivational Monism of a special variety: Intellectualist Motivational Monism. Even if we set aside the possibility, raised by my observations about the fine above, that Hedonism on the straightforward interpretation does not entail Motivational Monism at all,<sup>35</sup> there is a major problem: if Hedonism does entail Motivational Monism, it seems to entail Motivational Monism of entirely the wrong sort.

The Many thought they were subject to two species of motivation: in the language of the *Phaedrus* (237d, quoted above), beliefs about the good and appetites for pleasure. If Socrates’ argument in Stage 1 amounts to the claim that all they desire is the very thing they thought of all along as pleasure, then the natural inference would be that they are only motivated by the desires they all along thought of as desires for pleasure: appetites and other passions. As we saw in the previous section, however, Socrates proceeds in Stage 2 as if he has established just the opposite: fear, and by implication all the other motivations the Many opposed to intellectual judgments about the good, are themselves precisely such judgments. Moreover, I will now show, this intellectualist conception of motivation was already at work in Stage 1, and indeed was necessary

---

<sup>34</sup> Would popular belief in after-life rewards solve the problem? Perhaps, but it is notable that Socrates makes no mention of that here: he does not try to convince the Many that courageous acts are more pleasant by appeal to a hedonic calculus, but simply asserts the claim.

<sup>35</sup> If the fine diverges from the pleasant/good then the desire for the fine will be distinct from the desire for the pleasant/good.

for its argument: Appetitive Motivational Monism, as we might call it, would have provided no bulwark against Intellect Enslaved.

In Stage 1 Socrates uses Hedonism to establish Intellect Rules by arguing that if all the Many value is pleasure, they will never knowingly spurn a greater pleasure for a lesser one. (This is the key claim of the substitution argument at 355a-356a.) Notably, this argument treats the desires that motivate the Many as desires for an overall balance of pleasure over pain, or as desires for a life that is more pleasant than painful: only that kind of desire will be sensitive to calculations about long-term consequences in the way required.<sup>36</sup> If, however, in desiring pleasure the Many are led by the kind of desires they think they are led by – impulsive, calculation-insensitive, judgment-insensitive appetites – then the objection Socrates imagines to his argument at 356a will be decisive, and his response irrelevant:

For if someone were to say “But Socrates, the immediate pleasure differs greatly from what is pleasant or painful at a later time”, I would say: “Do they differ in anything other than pleasure or pain? For there is no other way. But like a man good at weighing, having put the pleasures together and the pains together and having weighed both the near and the far on the scale, say which one is greater. For if you weigh pleasures against pleasures, the greater and more must always be taken [...] and if you weigh pleasures against pains [...] that action must be done in which pains are exceeded by pleasures.” (356a5-b8)

Socrates here treats the desire for pleasure as a calculation-sensitive desire for a long-term, overall balance of pleasure over pain. The Many have however already implied the opposite view of this desire, in their claims that they are “overcome by pleasure” to pursue immediate gratifications; moreover, nothing they have said licenses this more sophisticated construal. An imagined objector with more fight in her might thus renew her case as follows: *No Socrates, what I meant was that when I go for pleasure I’m not going for the greatest balance of pleasure over pain in the long run – I’m going for something I can get right now. Just because pleasure is my only end doesn’t mean that I always choose what I take to be the greatest pleasure: my desires are much more haphazard and impulsive than that. Therefore I might know what is best (most pleasant) but still not choose it – so, despite my consent to Hedonism, I am sticking to my guns on Intellect Enslaved.*

---

<sup>36</sup> Compare Taylor: “the previous discussion (e.g. 353e5-354a1) has made it clear that ‘pleasure’ is to be understood as an abbreviation for ‘the predominance of pleasure over pain’ (commentary *ad* 355a2-5).

Does Socrates' argument from Hedonism to Intellect Rules thus rest on an unjustified and indeed question-begging assumption about the rationality of the Many's desires for pleasure? There is one clear way to escape this conclusion: if we take it that Socrates, in equating the pleasant with the good, is implicitly equating desires for pleasure with desires for the good *as he himself ordinarily understands the latter*. Socratic desire for the good, as for example at *Meno* 77b-78b is desire for what benefits one, i.e. for what contributes to one's overall *eudaimonia*. And this is per force a calculation-sensitive desire: it simply makes no sense to say that in desiring my overall or life-long benefit I am insensitive to considerations about overall balance or long-term consequences.

Thus we can save Socrates' argument if we grant him the premise that the desire for pleasure is just like the desire for good as traditionally understood, rather than vice versa. It is very hard, however, to see what would justify this equation on the face-value reading of Hedonism: if we are to understand the good as reducing to what the Many think of as pleasure, why, without a good deal of further argument, should we understand the desire for pleasure as reducing to what Socrates thinks of as the desire for good?

I have raised two problems for Socrates' argument on the face-value reading of Hedonism. One might try to address both by taking Stage 1 as *ad hominem*: Socrates thinks the Many have only end, the good (and thus that they desire pleasure not at all, or only instrumentally), but it would be so hard to convince them of this that he cheats and tells them that in desiring the good they are desiring pleasure. But this subterfuge would undermine Socrates' whole goal, and render the overall argument invalid. The many clearly *do* desire pleasure as an end; there is no point denying this. If their desire for pleasure is distinct from their desire for the good then they have more than one end, and so none of the arguments Socrates gives for Intellect Rules and thereby for Intellectualism succeed.

Nonetheless, the *ad hominem* reading of Stage 1 undeniably has some appeal: clearly Socrates is letting the Many and Protagoras accept something he does not, namely that the good is nothing other than what most people think of as pleasure. Moreover, he is clearly engaging in some kind of subterfuge. I will now propose a reading which allows for this while at the same time rendering the overall argument valid. The basic thought is that while Socrates allows the Many and Protagoras to assume the straightforward interpretation of Hedonism, the thesis he

himself is taking seriously is the converse: that the pleasant turns out to be nothing other than what he and most people already call the good.

Socrates is certain that all people desire the good: that is an axiom of his ethics here and throughout the dialogues. But the Many say that they desire the pleasant, and furthermore that this is something distinct and divergent from the good. If this is true, Socrates' account of virtue is lost, for one might know the good, but have a stronger motivation to pursue the pleasant. To save his account – to show that people are motivated *only* by the desire for the good – he needs the Many to say of themselves, “I really desire only one thing.” If this is true, then the pleasant – what they claim they desire – must be identical to the good, which he knows they desire.<sup>37</sup> Thus his motivation for pressing Hedonism on the Many can be spelled out as follows:

*I believe that there is only one end, the good, i.e. beneficial: all human motivation – yours included – has this one quality as its object. You claim to believe that there are several ends: the pleasant and the fine as well as the good. I know you also claim to equate the fine with the good, so let's just accept that equation (or whatever very close connection it is that you claim to see between the two). What I'm worried about is your attachment to the pleasant. When I began by asking Protagoras if he thinks things are good insofar as they're pleasant, I was trying to show that he thinks pleasure a form of benefit – so that the desire for pleasure is just a desire for good. But he resisted; he seemed to hold that people might want something qua pleasant without in any way finding it beneficial – that they might want pleasure for reasons independent of benefit.<sup>38</sup>*

*So here's the situation: I'm sure that everyone values the beneficial, but it is also clear that you value the pleasant, and this seems to show that you value two different things. If that is right then End Pluralism is true, and so Motivational Pluralism follows, and therefore Motivational Conflict is possible, and Intellectualism is false. But Protagoras said something that gave me an idea. He took my question about whether pleasure is good to be the question of whether the pleasant and the good are one and the same. Now there's a thought: what if the good I'm certain*

---

<sup>37</sup> On this point I closely follow Irwin (1995, 88), although in what ensues my interpretation diverges sharply from his.

<sup>38</sup> Dimas 2008 shows very convincingly that Socrates' initial questions focus only on this claim; the account I offer here explains that, without accepting Dimas' inference that Socrates' argument does not rest on the stronger thesis of Hedonism at all.

*you value and the pleasant you say you value are in fact the same thing? In other words, what if Hedonism is true? I hereby adopt it as a hypothesis.*

*I know how you'll understand Hedonism. You'll take the claim that the pleasant and the good are the same to mean that the very kind of thing you already think of as pleasant is what turns out to be good. And you'll take the claim that pleasure and the good are the same to mean that goodness reduces to what you already think of as pleasure. And for the purposes of argument I'm happy to engage in a little subterfuge and let you understand me that way. But in fact my hypothesis, the one I have in mind when I'm considering that Hedonism may be true, is importantly different. What I mean is: maybe pleasure, the very thing you think you're getting from bodily gratification, turns out to be nothing other than what we both recognize as the good, namely virtuous activity.*

*I'm not in the business at present of giving any theory about how this identification would work: indeed I am notably abstaining from giving any theory of pleasure at all in this dialogue.*

*Perhaps pleasure turns out to be the state one is in when one's body or soul is functioning as it should. I suppose one might put the point by saying "pleasure is unimpeded activity of a natural state" – something like that.<sup>39</sup> Certainly anyone who holds that view will wind up saying things that sound just like Hedonism: the good that we all seek is a pleasure.<sup>40</sup>*

*Mind you, I'm not just making the weak claim that the pleasant and the good always converge (although I do think that true). Nor am I making the claim that the pleasant and the beneficial are two guises of the same thing, the good. That wouldn't do the trick at all: as my 20<sup>th</sup> century successors will explain at length, one can have different attitudes toward the same thing under different guises; therefore if the pleasant is a guise of the good distinct from the beneficial then the desire for the pleasant will be distinct from the desire for the beneficial.<sup>41</sup> No, my hypothesis is a much stronger and stranger one: that these are one and the same guise. Finding something*

---

<sup>39</sup> Aristotle, *Nicomachean Ethics* (EN) 1153a14-15. Rudebusch attributes this theory of pleasure to Socrates on the (somewhat slim) basis of his claim in the *Apology* that dreamless sleep is supremely pleasant, and uses it to argue for an interpretation of Socratic ethics very much like what I'm proposing here: pleasure is happiness (Rudebusch 2002, chapter 10).

<sup>40</sup> Aristotle, EN 1153b12-13.

<sup>41</sup> Plato has been accused of missing this kind of point in the substitution argument: see for example Taylor's commentary on 355c1-8.

*pleasant just is finding it good – and therefore desiring something as pleasant will be desiring it as good.*

*If this is right, then you suffer from some serious conceptual confusions. You think that you judge some things good even though you find them unpleasant, but this simply isn't possible. This is because you mistakenly think that what it is to judge something good is to think of it as an obligation of some kind, the way you now think of medical treatment or exercise. In truth, however, judging something good involves being attracted to it in the way you are toward the things you recognize as pleasant. Therefore, unbeknownst to your own benighted selves, what you are going for when you pursue pleasure is nothing other than the good. You're groping for the good in your own confused way, and you're so lost that you don't even recognize that that is what you're doing.*

*For example, you think you like indulging your appetites for tasty food because of the tingles and thrills and gratifications that provides. What you don't realize is that the positive reaction you have to such indulgences – the pleasure – is in fact a response to them qua good or beneficial. You respond so positively to tasty food because you mistakenly (and unwittingly) find it beneficial; if you didn't, you wouldn't find it pleasant, because what it is to find things pleasant is to find them beneficial.<sup>42</sup>*

*Now I have refuted your argument for End Pluralism: there are not two different kinds of things people go for, the pleasant on the one hand and the good on the other, but just one, most*

---

<sup>42</sup> It may help to consider an aesthetic analogy. Imagine an avowed End Pluralist about music, who thinks that she values the musically pleasant as a distinct end from the musically beautiful. She takes guilty pleasure in soft rock, and is bored by the classical music she calls beautiful. The counterpart of the *Protagoras*' Socrates would tell her: *The trouble with your self-analysis is that you are ignorant of a very important identity: there is just one value in music – the musically pleasant just is the musically beautiful – and therefore what it is to find something musically pleasant is to find it musically beautiful. You think that the feeling of bored obligation you have toward classical music is your finding it beautiful, but in truth finding something beautiful involves being attracted to it in the way you are to the music you call pleasant. Therefore what you are going for when you pursue pleasure in music is in fact nothing other than the beautiful. You respond so positively to Neil Diamond songs because you mistakenly (and unwittingly) find them beautiful; if you didn't, you wouldn't find them pleasant, because what it is to find music pleasant is to find it beautiful.*

*illuminatingly called the good (since that gets at its intrinsic nature), but accurately enough described as the pleasant (since pleasure turns out to be identical with it). Perhaps we can even say something less reductive about pleasure: even though the most informative way into the identity is through 'good,' the ordinary connotations of 'pleasant' capture an important aspect of the good, namely its attractiveness, its subjective, manifest value, and arguably its susceptibility to quantification and measurement.*

*Furthermore, since End Pluralism turns out to be false so too does Motivational Pluralism. There are not two different species of motivation, but just one: intellectual judgment about what is good (or desire directly dependent on it). You think you find within yourself a weak impulse to pursue what you think good, alongside some very different desires: lusts and fears and other so-called unruly passions that drag you toward pleasure regardless of your better judgment. But I've shown you that this putative weak motive is a figment of your confusion, while in pursuing bodily pleasures and fleeing bodily pains you are simply (although with a great deal of ignorance and lack of self-knowledge) going for what you think best. The very same intellectual motivation toward the good that makes virtuous people abstain from bodily indulgences drives you to pursue them.*

*Therefore there is no Motivational Conflict: you are not torn between a strong desire for the pleasant and a weak one for the good, but simply moved by a single species of desire.*

*Therefore Intellectualism turns out to be true: the only difference between you and virtuous people is that you are ignorant of what is truly good while they are knowledgeable. Virtue turns out to be a species of knowledge. This implies that it can be taught, and indeed I believe or at least hope that you can come to learn which things are truly good/pleasant: right now you think good and take pleasure in bodily indulgences, but with the proper education in measuring the value of your actions you will come to think good and take pleasure in virtuous actions.*

*I confess to using some subterfuge in my argument: I know that you will misunderstand my meaning when I suggest Hedonism to you, and I exacerbate that misunderstanding by using examples of bodily gratification to fix the reference of 'pleasure.' But no fear: my subterfuge is harmless. If you really have no other criterion of goodness than pleasure, and if your desire for pleasure is indeed an intellectual one, sensitive to information you get through measurements and calculations, then you will change your behavior if you learn new facts about the pleasant: if you*

*learn that exercise is more pleasant than cake, for example, you'll stop desiring to over-eat and start desiring to exercise. And if the Hedonist hypothesis as I mean it is correct, then it is acting virtuously that provides the greatest amount of the thing you recognize yourself as being after, pleasure. (This is not, of course, because such activity gives you more of those bodily tingles and thrills. It doesn't, but we're working on the hypothesis that those weren't really what you were going for in going for pleasure. Instead, you were going for what pleasure really turns out to be, goodness, which is of course present in fine, virtuous activity to the highest degree. That is why I can so confidently assert to Protagoras and the others that courageous acts are pleasant, and indeed that all fine acts are pleasant and all acts that lead toward livingly pleasantly are fine.) So if you learn how to measure pleasures, soon you'll move from desiring cake to desiring exercise, and onward from there to desiring virtue most of all.*

If the Hedonism Socrates proposes in the *Protagoras* is indeed best understood along these lines, then Socrates can rest his argument on it without contravening his positions in this and other dialogues. He is proposing not popular Hedonism, but a new brand, what we might call Socratic Hedonism.

Is it an objection to this interpretation that it accuses Socrates of subterfuge? It cannot be. I hope to have shown that if the face-value reading of Hedonism is right, he is indulging in much worse subterfuge: he is pretending that the arguments of Stage 1 support the conclusions in Stage 2, when in fact neither they nor anything else he says in the dialogue do that work. Second, the particular form of subterfuge I am charging Socrates with here is not at all out of character. In fact, it is one of his standard tactics, precisely the sort of maneuver which Vlastos famously characterized as Socratic irony: he implicitly gives a radical new meaning to some term – for example, virtue, harm, power, knowledge, or teaching – which he knows his interlocutor to understand in a standard way.<sup>43</sup> (Book I of the *Republic* presents a classic case: Socrates argues against Thrasymachus that justice is to one's advantage, all along using both 'justice' and 'advantage' in ways radically different from his opponent.) My reading of the *Protagoras* thus renders Socrates' reliance on Hedonism not only not uncharacteristic, but in fact just the sort of move we should expect him to make.

---

<sup>43</sup> See Vlastos 1991, chapter 1.

I began by asking why Socrates uses the apparently un-Socratic doctrine of Hedonism as a premise in his argument for the very Socratic thesis of Intellectualism. We now have our full answer. Socrates is not exploiting a view that he rejects but knows his interlocutor embraces, nor is he putting forth a view to which he is unequivocally committed. Instead, he is proposing a hypothesis that he hopes is true, for with it, he now thinks, stand and fall his views about virtue. If Socratic Hedonism is true – if it is an accurate thesis about the ends of human motivation – then Intellectualism stands. If it is false, however – if the pleasant is an end distinct from the good – then Socrates’ view of human psychology, and consequently of human virtue, are cast into doubt.

That the final argument of the *Protagoras* is conditional in this way is confirmed not only by Socrates’ concern to have his interlocutors embrace the Hedonism they initially disavow, and by the dialogue’s aporetic ending, but by a look ahead to the *Republic*. Here Socrates explicitly rejects Hedonism (505c), and gives a detailed account of pleasure that clearly distinguishes it from the good. And along with Hedonism he rejects all the views the *Protagoras* made dependent upon it. End Monism goes, and with it Motivational Monism: the *Protagoras*’ three putatively identical ends – the pleasant, fine, and beneficial – show up as the objects of three distinct species of motivation, the appetitive, the spirited, and the rational.<sup>44</sup> Motivational Harmony and Intellect Rules go too: vicious people have conflicting desires, and are ruled by their passions. Finally, Intellectualism is abandoned: wisdom is a component of virtue, but not the whole thing, for it can neither arise nor rule without the harmonious cooperation of the appetitive and spirited parts of the soul. The convergence of the pleasant with the good, and the resulting lack of conflict within the soul, are no longer starting-points to be used as the basis for moral education, as Socrates proposed in the *Protagoras*. Rather they are the goals of moral education. It is only when we have become virtuous that conflict in the soul disappears, and our appetites and passions reliably desire as pleasant or fine that which we rationally judge best.<sup>45</sup>

---

<sup>44</sup> The rational part is identified as the part that cares about the beneficial (439c, 441c, 442c); the spirited part is called a lover of victory and honor (581b), both of which are naturally taken to be desirable because fine (see Moss 2005 and Lear 2004, chapter 6); the appetitive part is characterized by its desire for “the pleasures of food, drink, sex and others” (436a; cf. *Laws* 689b, where this part is called “the part that feels pain and pleasure”).

<sup>45</sup> I am grateful for enormously helpful discussion and comments on previous drafts from James Allen, John Cooper, Cian Dorr, Jimmy Doyle, Brad Inwood, Terry Irwin and Chris Shields, and to the audience at the Southern Association for Ancient Philosophy in 2011.

## Works Cited

- Adam, J. and A.M. 1905. *Platonis Protagoras*. Cambridge.
- Allen, J.V. forthcoming. "Why there are Ends of Both Goods and Evils in Ancient Philosophy".
- Dimas, P. 2008. "Good and Pleasure in the *Protagoras*". *Ancient Philosophy* 28: 253-284.
- Dodds, E.R. 1959. *Plato: Gorgias*. Oxford.
- Frede, M. 1992. *Plato: Protagoras*. Indianapolis.
- Gosling, J.C.B. and Taylor, C.C.W. 1982. *The Greeks on Pleasure*. Oxford.
- Irwin, T. 1995. *Plato's Ethics*. Oxford.
- Kahn, C. 2003. "Socrates and Hedonism". In *Plato's Protagoras: Proceedings of the Third Symposium Platonicum Pragense*. Ed. A. Havlicek and F. Karfik. Prague.
- Klosko, G. 1980. "On the Analysis of *Protagoras* 351b-360e". *Phoenix* 34: 307-321.
- Lear, G. R. 2004. *Happy Lives and the Highest Good*. Princeton.
- Moss, J. 2005. "Shame, Pleasure and the Divided Soul", *Oxford Studies in Ancient Philosophy* 29: 137-70.
- Nussbaum, M.C. 1986. *The Fragility of Goodness*. Cambridge.
- Rudebusch, G. 2002. *Socrates, Pleasure, and Value*. Oxford.
- Segvic, H. 2000. "No One Errs Willingly: The Meaning of Socratic Intellectualism", *Oxford Studies in Ancient Philosophy* 19: 1-45
- Shields, C. 2007. "Unified Agency and *Akrasia* in Plato's *Republic*". In *Akrasia in Greek Philosophy: From Socrates to Plotinus*. Eds. C. Bobonich and P. Destree. Leiden: 61-86
- Taylor, C.C.W. 1996. *Plato: Protagoras*. Oxford.
- Vlastos, G. 1969. "Socrates on *Acrasia*". *Phoenix* 23: 71-88
- 1991. *Socrates, Ironist and Moral Philosopher*. Cambridge.
- White, N. 1999. "Harmonizing Plato". *Philosophy and Phenomenological Research* 59: 497-512.
- Zeyl, D. 1980. "Socrates and Hedonism: *Protagoras* 351b-358d". *Phronesis* 25: 250-269