The Socratic Paradoxes and Plato's Epistemology

Natalie Hannan

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2021

Abstract

The Socratic Paradoxes and Plato's Epistemology

Natalie Hannan

Plato's "Socratic paradoxes" state that no one does wrong voluntarily and that virtue is knowledge. Outside of moral psychology, the importance of the Socratic paradoxes has been neglected. My dissertation defends two related proposals that showcase their importance in ancient epistemology. The first proposal is that they are a major motivation for Plato to develop a unique view of *epistēmē* (knowledge or understanding) as an infallible and robust cognitive power that is set over a special class of objects. The second proposal is that understanding the influence of the Socratic paradoxes can help us see how *epistēmē* improves our *doxai* (beliefs or opinions) about the world around us, solving a long-standing problem in Plato's epistemology. I will start by examining the *Hippias Minor*, in which we see Plato seeking to embrace the Socratic paradoxes (rather than already assuming them) and looking to develop his notion of *epistēmē* as a result. I will then move to the *Protagoras*, in order to show Plato proceeding with this project by embracing *epistēmē* as something that produces good action and involves measurement. I will show the *Protagoras'* picture to be fully developed in the *Republic*, in which *epistēmē* emerges as something that measures the truth of our *doxai* and has clear practical benefits as a result. Finally, I will compare this account to Aristotle's treatment of virtue and *epistēmē* in the *Eudemian Ethics*, in order to consider the legacy of the Socratic paradoxes after Plato.

Table of Contents

Acknowledgements

Many wonderful people were instrumental in bringing this dissertation to completion, and though these brief acknowledgements can hardly do them justice, it is a joy to be able to thank them here.

It is difficult to overstate the advice and insights I have gained from Katja Vogt, both regarding this dissertation and throughout my time at Columbia. I had the good fortune of meeting her as a curious undergraduate and I have benefited from her mentorship ever since. Her enthusiasm and diligence towards her students are unparalleled in the field. I am exceedingly grateful for the opportunity to have grown philosophically and personally as a result of her guidance and for the thriving and collegial community of ancient philosophy graduate students that she has created at Columbia.

I am thankful to Dhananjay Jagannathan for being a tireless force for ancient philosophy at Columbia, and I have benefited greatly from his numerous reading groups, exceptional teaching, and generous mentorship. My chapter on the *Eudemian Ethics* was made possible through many meetings in the summer of 2018, in which we attempted to untangle the philological and philosophical puzzles of the text.

John Morrison has been a cherished mentor and teacher throughout my time at Columbia, and I am grateful to him for agreeing to serve on my committee and offering the insights that come with a non-ancient perspective. Thanks as well to Iakovos Vasiliou and Whitney Schwab, who so kindly served on my examining committee and offered many helpful comments, not only on the dissertation itself but also through conversations about these issues for many years.

I have been lucky enough to have presented parts of the first three chapters to helpful audiences, whose questions and feedback have greatly improved each one. For Chapter 1, I am

iv

Introduction: The Socratic Paradoxes and Plato's Epistemology

Overview

Though Plato is no stranger to surprising doctrines, two features of his philosophy number among the most peculiar. The first is what are typically called the "Socratic paradoxes." The Socratic paradoxes are a cluster of theses that make unintuitive claims about virtue, knowledge, and action. A number of theses have been designated Socratic paradoxes in the literature, but I will focus on two of the most prominent ones:

> Socratic Paradox 1 (SP1): No one does wrong voluntarily.
> Socratic Paradox 2 (SP2): Virtue is knowledge.

The second distinctive feature is Plato's theory of *epistēmē*. *Epistēmē* is often translated as knowledge or understanding, but it is unlike anything in modern epistemology: it is a highly specialized cognitive power set over the highest class of objects, which are completely separate from the things that we encounter in the world. *Epistēmē* is highly important not only to Plato's epistemology but also to his ethics, taking on a crucial role in his description of the best possible city and soul. Though Plato clearly places high value on *epistēmē* in his philosophy, it, like the Socratic paradoxes, has usually been found unintuitive and thought to raise more problems than it solves. Without clear answers to these problems, one is tempted to conclude that Plato's conception of *epistēmē* may be more trouble than it is worth.

In my dissertation, I aim to do two things. The first is to show that these two features are closely connected. In particular, I will argue that Plato's commitment to the Socratic paradoxes is a primary motivation for him to develop his unique view of *epistēmē*. My second aim is to demonstrate that we will gain greater insight into *epistēmē* itself through understanding its relationship to the Socratic paradoxes. In this regard, I will be focusing on one of the most

notorious problem with *epistēmē*—namely, what benefit it can provide given its highly specialized status. I maintain that we will be able to solve this problem when we understand how *epistēmē* relates to and is affected by the Socratic paradoxes.

I will provide an overview of how this project will proceed. First, however, I will say a bit more about the Socratic paradoxes and about Plato's treatment of *epistēmē*.

The Socratic Paradoxes

The first thing to note about the "Socratic paradoxes" is that they are neither paradoxical nor Socratic. With regard to their status as paradoxes, we can see pretty clearly that there is nothing logically contradictory about SP1 or SP2. What makes them *feel* paradoxical, rather, is that they strike most readers as terribly misguided, perhaps to the point of incoherence. Surely virtue is not knowledge; Aristotle dismissed that idea rather quickly, and it seems much too strong and simple for the complex pictures of virtue that have developed since.[1] The claim that no one does wrong voluntarily smacks of a desperate attempt to let humanity off the hook. Only a few modern attempts have been made in this vein, by raising the standards for moral responsibility especially high.[2] Even these cases, however, stop short of concluding outright that no one is morally responsible for wrongdoing, let alone that no one does wrong voluntarily. It just seems obvious that voluntary wrongdoing—whether through succumbing to temptation, acting with clear-eyed evil, or something else—is not only possible but prevalent.

---

[1] See *Nicomachean Ethics* II.4: "Where the virtues are concerned, however, knowing has little or no strength, whereas the other factors have not just a little but, rather, *all* the significance, and these are the very ones that come about from frequently doing just and temperate actions" (1105b1-4). Even the promisingly-named field of virtue epistemology is not really in the business of showing that virtue is to be identified with knowledge so much as the other way around. Roberts and Wood (2007) provide a bit of an exception by considering several virtues on epistemological terms, and there is a blossoming field of study on epistemic injustice stemming from Fricker (2007), but even these, I take it, do not go quite so far as to say that virtue *just is* knowledge.

[2] See Rosen (2003, 2004), who probably goes the furthest down this path in expanding the kinds of ignorance that would release an agent from moral responsibility. Even he, however, adopts more of a skeptical attitude towards the endeavor, much different from what we see in Plato.

Before we declare Plato completely off the mark, however, we should consider whether we are interpreting SP1 and SP2 as he would have intended. This is not to say that they would not have been unintuitive, even in Plato's time, but rather that their meaning might be less than obvious. In particular, we might wish to get clear on two terms: "voluntarily" in SP1 and "knowledge" in SP2. Plato never gives us an explicit and in-depth explanation of voluntary action, and scholars have not made much of an attempt to draw one out of his writings.[3] In order to get clear on whether voluntary action is possible, however, we must first determine what voluntary action involves; as we begin to explore SP1, we should make sure to find out how to interpret it on Plato's terms. What Plato means by "knowledge" in SP2 will, perhaps unsurprisingly, come to light as we examine his epistemology more closely. For now, we should suspend judgment on the plausibility of SP2 until, like SP1, it is further clarified.

So much for SP1 and SP2 as paradoxes. We should also hesitate, I contend, to treat them as Socratic. The question of what Socrates the historical figure thought and taught is an incredibly fraught one, since Socrates famously did not write anything.[4] The task then falls to trying to discern which of Plato's dialogues held the "real" views of Socrates, aided imperfectly by the testimony of Xenophon and (even more imperfectly) Aristophanes. If we accept the standard chronology of the dialogues (more on that in a moment), then we might expect the

---

[3] Wolt (2019) does some work in this vein, constructing a conception of the voluntary from the *Timaeus* that involves a connection to the agent's intellect. Though a full examination of how Plato uses the voluntary throughout his corpus is best saved for further research, we shall see that Plato puts forward a similar account in the *Hippias Minor*, in which voluntary action involves having power and know-how for each domain of action. This is probably a good place to mention my choice to render ἑκών as "voluntarily," rather than "willingly" or "deliberately" as it has also been translated. The main reason for choosing "voluntarily" is that I deem it to be the most neutral term of the lot. "Willingly" inclines more strongly to the notion of the will, which I do not take to be prominent, if existent, in Plato (see Frede (2011), but see Segvic (2009) for arguments in favor of the will in Plato). "Deliberately" builds in too much of a focus on intention for my taste. "Voluntarily" is, of course, not perfect, but I believe that it can be read with as little baggage as possible and would encourage the reader to do so.

[4] The one exception may be some poetry and a hymn mentioned at *Phaedo* 60d, but they do not survive, may have been part of Plato's fictionalization, and would likely not be a clear guide to Socrates' philosophy in general.

"earlier" dialogues to contain more of Socrates' thought, before Plato develops his philosophy and branches out on his own. Chronologically, we would thus expect the Socratic paradoxes to be primarily a feature of Plato's "Socratic" dialogues. Philosophically, we would not expect them to feature very prominently. If they were passed down as important theses of Socrates', Plato would likely consider them axiomatic, not worth spending much energy on but important to his teacher's thought and his own—until he develops his own ideas. As Plato progresses, it would not be surprising to see the Socratic paradoxes quickly abandoned as interesting but defunct philosophical leftovers.

The Socratic paradoxes, however, appear all over the canon. SP1, for instance, pops up in the *Laws*, known to be one of Plato's latest dialogues.[5] If the paradoxes are original to Socrates, they are enduring parts of Plato's thought throughout his whole career. In future research, I intend to argue that SP1 is a purely Platonic, rather than Socratic, commitment.[6] For now, I propose that we suspend judgment about both the Socratic and the paradoxical natures of the Socratic paradoxes.[7] Instead, when we see Plato advocate for them, we should take them on their own terms, not as obscure and confusing pre-commitments from Plato's philosophical education. With this attitude, we may begin to see their value.

The existing literature on the Socratic paradoxes is quite bare. The only recent systematic study focused on the paradoxes themselves comes from Roslyn Weiss, in *The Socratic Paradox and Its Enemies* (2006). Weiss presents an extremely *ad hominem* reading of

---

[5] See *Laws* 860d-e.
[6] My current thinking is that there may be a difference between SP1 and SP2 on this count. There does seem to be some external confirmation that Socrates thought something like virtue is knowledge. There is no such confirmation, however, about the impossibility of voluntary wrongdoing, and as we shall see in Chapter 1, we should doubt that this thesis was securely in place in the early dialogues. Apart from these considerations, the Socrates I refer to is Plato's Socrates, and I will not attempt to give further color here on the historical figure.
[7] I will, however, continue to use the term "Socratic paradox" in order to avoid confusion and inconsistency with the existing literature. A colleague once suggested to me that "Surprising Principles" would be a more apt name than "Socratic Paradoxes;" one can read SP1 and SP2 with this gloss in mind.

the Socratic paradoxes; her primary contention is that SP1 and SP2 (in addition to a third

paradox, "all the virtues are one") are used as tools of attack against the sophists and their

ignoble views. We should, in her view, remain agnostic about whether Socrates actually holds

these beliefs, since their use is merely fitted to the dialectical occasion. I believe that we shall

see that picture begin to unravel when we turn to the *Hippias Minor* in Chapter 1. For now, let

us simply note that the primary (by default) view on the Socratic paradoxes does not assign them

any importance in Plato's own philosophical theories.

Earlier scholarship that does take the Socratic paradoxes seriously tends to focus on their

import for Plato's moral psychology—their bearing on desire, wish, or related ideas.[8] Though

knowledge is recognized to be important to SP2 in particular, the knowledge in question is

usually taken to be fairly easily explainable: that a certain action is unjust, that unjust actions

harm the agent.[9] What is not considered is the *effect* the Socratic paradoxes may have on

knowledge—if Plato is seriously committed to them, how it might change his epistemology, not

just his moral psychology, as a result. The Socratic paradoxes, after all, are not the only

surprising or strange pieces of Plato's philosophy; as I said at the start, his theory of knowledge

also makes the list.

Epistēmē

My investigation of Plato's epistemology will focus primarily on his treatment of

epistēmē. As I said at the start, *epistēmē* is usually rendered as some kind of knowledge or

understanding. Much of my project involves trying to determine what Plato has in mind when he

---

[8] See Bambrough (1960), Santas (1964), O'Brien (1967), McTighe (1984), Weiss (1985b), Segvic (2009).
[9] Allen (1960), and to some extent O'Brien (1967), consider whether there is some relation between knowledge of the Forms and the Socratic paradoxes, which we will see is on the right track. No one, as far as I can tell, takes the Socratic paradoxes to *inform* or *motivate* Plato's epistemology in any way, rather than simply being in accord with parts of it.

discusses *epistēmē*, so in order to avoid making any unwarranted assumptions about it, I will

leave it transliterated throughout.[10] I will at times draw conclusions based on Plato's use of

epistēmē instead of a similar term and vice versa, but it is important to flag straightaway that I

am not interested in simply tracking the uses of the term "*epistēmē*." Plato is not always

consistent in his employment of epistemological terms, a fact to which he calls explicit

attention.[11] There will be occasions where Plato speaks of *epistēmē* but is not picking out what is

of interest to us, and there will be times that Plato speaks of something else and has *epistēmē* as I

understand it in mind. Though *epistēmē* is his best and most typical term for our topic of

interest, I will not be tracking "*epistēmē*" unfailingly.

What I *am* interested in tracking is the notion that *epistēmē* often picks out: an infallible

and robust cognitive power that is set over a special class of objects.[12] In this description of

epistēmē, albeit an imperfect one, we can see its similarities to and differences from

contemporary views of knowledge and understanding. Like *epistēmē*, knowledge and

understanding are factive epistemological terms: one has knowledge or understanding only of

what is true. Most would agree that knowledge and understanding are more robust than states of

mind like opinion or conjecture; there is some non-accidental connection with the truth, and

---

[10] See Burnyeat (2011) for some of the different "kinds" of knowledge in contemporary discourse and their relation to *epistēmē* and related terms. For a defense of *epistēmē* as understanding (in the *Meno*, but with broader applications) and further discussion of recent debates regarding *epistēmē* and contemporary epistemology, see Schwab (2015).

[11] E.g. *Republic* VII, in which Socrates flags the relative unimportance and occasionally sloppiness of epistemological terms: "What we often used to call *epistēmai* from habit, we need to find another name for [. . .] but, it seems to me, it is not worth disputing about a name, since the investigation of many greater things than these lies before us" (533d4-9). Translations of the dialogue of focus for Chapters 1 through 3 are my own unless otherwise noted, with consultation of the translations in Cooper, ed. (1997). Translations of dialogues other than the one of focus are also taken from that edition. I have sometimes sacrificed a translation's readability for a closer adherence to the text, though I hope that the former has never been sacrificed completely.

[12] I will use "notion" and "conception" often when talking about Plato's view of *epistēmē*, but I do not mean anything technical by them. These terms are only meant to flag that Plato has *some* view in mind of what *epistēmē* is, and whatever conditions or attributes fit that view is what we see in his notion or conception of *epistēmē*. I do not mean to imply that *epistēmē* is merely a mental concept or something along those lines.

perhaps with stability as a result.[13] On these counts, *epistēmē* would fit with some ease into the

modern epistemological conversation.

Where the similarities end, however, is at the second part of my description, that *epistēmē*

is set over a special class of objects. To specify a bit more, *epistēmē* is set over a class of objects

that are fully (and only) intelligible, completely unchanging, and in no way deficient regarding

their natures. These are usually taken to be Plato's famous Forms: Goodness, Beauty, Justice,

and whatever else might be included among their number.[14] In contrast to modern views of

knowledge and understanding, nothing on Earth as we know it is a candidate object of *epistēmē*.

Rather, everything in the sense-perceptible world is cognized by the power of *doxa*, usually

translated as belief or opinion.[15] Plato posits two independent cognitive powers that deal with

two completely separate kinds of objects: *epistēmē* is not a subcategory of *doxa*, in the way that

knowledge is usually taken to be a subcategory of belief.[16] Rather, *epistēmē* and *doxa*, like their

objects, do not overlap.[17] This understanding of Plato's epistemology is commonly called the

---

[13] Hence the enduring force of Gettier (1963), whose counterexamples to knowledge as justified belief turn on showing that such an account of knowledge would lead to knowledge being acquired only accidentally in some cases. I do not take robustness (an intentionally vague term) to be equivalent to non-accidentality but rather to be an umbrella term for *something* special about *epistēmē* that has to do with its strong connection to the truth and its stability.

[14] I will not be taking up the question of how much further, if at all, the realm of the Form extends; taking the above three as Forms, a fairly uncontroversial assumption, will be sufficient for our purposes.

[15] Of these two translations, opinion seems more accurate than belief, but I will also leave *doxa* transliterated. See Moss and Schwab (2019).

[16] As with the English knowledge and belief, *epistēmē* and *doxa* can be used to designate both the cognitive powers and the content with which these powers deal. We can thus, for instance, speak of both my *doxa* that cognizes certain things and the corresponding *doxa* or *doxai* (the plural form, as with *epistēmai*) in my soul. I will use both terms in both ways; see Crombie (1963) for more discussion of these different components. As we shall see in Chapter 3, the question of these powers' content bears on how separate one takes these powers to be, but I do not think we should rule on these issues just yet.

[17] Williamson (2000) is notable for taking knowledge to be a separate state of mind from belief, rather than a species thereof. His account, however, diverges from Plato's in many of the same ways as contemporary epistemology does; the objects of knowledge are much more in line with other contemporary views of the ordinary things that one can know.

Two Worlds view and comes about most clearly, according to its supporters, in "doctrinal" works like the *Republic*.[18]

Plato outlines the Two Worlds view of *epistēmē* and *doxa* most clearly in the *Republic*, in which he spends some time on the differences between them. Acquiring *epistēmē* is a privilege reserved for a select few: the philosophers who undergo a long and rigorous process of education beforehand. *Epistēmē* is important to forming the best city because it makes these philosophers uniquely situated to rule; through having *epistēmē* of the Forms, they enact laws and customs that model Goodness, Justice, and so on and develop a citizenry of upright character. With *epistēmē*, Plato is not just outlining an epistemological curiosity; he is describing an essential part of political and ethical life.

In contrast to the Socratic paradoxes, Plato's treatment of *epistēmē* has received considerable attention in the scholarly literature. Much of this attention has been devoted to trying to solve a fairly obvious problem with the Two Worlds picture as it has been laid out so far. I will call this problem the Problem of the Gap. It can be stated simply: if *epistēmē* and *doxa* are completely different powers, with completely different objects, how can one influence the other? For Plato does posit such an influence, in both directions. We are raised in the sense-perceptible world with only *doxa*, but somehow the philosophers are able to grasp objects to

---

[18] There has been a lot of wiggle room in the literature as to how strict of a Two Worlds-er one must be. A weaker reading would say that *epistēmē* and *doxa* are naturally set over, though not exclusively about, their objects in both directions (e.g. Vogt (2012)). A stronger reading might say that one of these directions is exclusive, though not necessarily the other (e.g. Schwab (2016)); the strongest would say that there is absolutely no *epistēmē* of perceptibles or *doxa* of intelligibles (e.g. Gerson (2003); Vlastos (1965) also seems likely to embrace this but faults Plato on account of it, and most interpreters shy away from the strongest reading). I take all of these readings to count as sufficient to consider one a Two Worlds-er. At its core, I take the Two Worlds theory to adopt this basic commitment: the objects that *epistēmē* is set over and that *doxa* is set over do not ever overlap. This commitment leaves room for some kind of cognition of the other set of objects, even though those objects are not what the power is "set over." But despite the wiggle room, I take this commitment to be sufficient for raising the Problem of the Gap that I discuss presently; the existing literature has not yet done a sufficient job of solving this problem on any Two Worlds reading, even the weaker ones.

which our world gives us no access. We do not encounter Forms in our day-to-day lives in any

way, so how do philosophers escape the "cave" that figures in Socrates' famous analogy?[19]

Similarly, once the philosophers acquire *epistēmē* of the Forms, they are supposed to be able to

make a better city and better citizens. But city laws and the thoughts and actions of citizens all

happen on the level of *doxa*. If the objects of *epistēmē* are in no way like the objects of *doxa*,

how will *epistēmē* bring about these benefits? How can it be of any value at all?[20]

Proposed solutions to the Problem of the Gap, I submit, have not so far been satisfactory.

One famous response, which has generated more discussion than converts, is to deny Two

Worlds entirely and maintain that *epistēmē* and *doxa* are not completely separate but overlap

much like knowledge and belief.[21] But many scholars have argued, convincingly in my opinion,

that the text shows pretty clearly that we must adopt some version of Two Worlds.[22] Pro-Two

Worlds solutions have either left the gap something of a mystery or given general descriptions of

how *doxa* and *epistēmē* can inform each other despite it. What I intend to do is to show how we

can get a more concrete picture of *epistēmē*'s influence on *doxa*, through understanding how

epistēmē has been developed as a response to the Socratic paradoxes.[23]

---

[19] See *Republic* VII, 514a-517c.
[20] Fine (1990) puts the problem succinctly: "If [the philosophers'] knowledge is only of the Forms—if, like the rest of us, they have only belief about the sensible world—it is unclear why they are specially fitted to rule in this world. They don't know, any more than the rest of us do, which laws to enact" (86).
[21] Fine (1978, 1990) is the notable defender of this position. Gosling (1968), Annas (1981), and Szaif (2007) also incline this way to some extent.
[22] See Gonzalez (1996), Baltzly (1997), Vogt (2012), Schwab (2016).
[23] My focus in closing the Problem of the Gap will be on one direction of the *epistēmē-doxa* divide: how *epistēmē* can influence and improve one's *doxai*. I focus on this direction because I see it as more relevant to Socrates' project with the Socratic paradoxes, in which he seeks to find a notion of *epistēmē* that explains virtuous action. I believe that the relationship between *doxa* and *epistēmē* that I lay out in Chapter 3 will also make some progress towards explaining how the world of *doxa* can lead to *epistēmē*, but I shall not try to give a full solution to this problem here.

Chronology and Philosophical Development

My dissertation has four chapters. The first three chapters are each devoted to one of Plato's dialogues: in order, the *Hippias Minor*, *Protagoras*, and *Republic*. The fourth chapter focuses primarily on Aristotle's *Eudemian Ethics*. Before I give an overview of how the argument will proceed, let me first say something briefly about the chronology of Plato's dialogues.

The only dialogue for which we have strong confirming chronological evidence is the *Laws*, which came very late in Plato's career. Some other points of evidence point to certain temporal placings—the *Theaetetus* after Theaetetus' death, say, or the *Apology* fairly close to Socrates' actual court defense. Stylistic and philosophical similarities between these dialogues and others have provided a rough sketch of Plato's philosophical development regarding his corpus. The standard picture tends to go like this: Plato started with a number of shorter, "Socratic" dialogues that mainly captured the historical Socrates and his method and did not put forward much doctrine.[24] As Plato developed some of his own ideas, the dialogues lengthened and more "Platonic" doctrines appeared; Plato's latest period is a bit of a mix, with some dialogues abandoning the character Socrates altogether, some views from Plato's middle period called into question, and some works in more of the original Socratic spirit.

According to this picture, the dialogues of focus in my dissertation proceed roughly in order of composition: the *Hippias Minor* as an early dialogue, the *Protagoras* probably also early but more transitional, and the *Republic* solidly in Plato's middle, doctrinal period. The story I am telling will be one that speaks of a development of Plato's epistemology, which starts with an uncertain attitude towards *epistēmē* in the *Hippias Minor* and progresses by forming

---

[24] For several recent perspectives on chronology and developmentalism in Plato, see the contributions to Annas and Rowe, eds. (2002).

some solutions in the *Protagoras* and fleshing out those solutions fully in the *Republic*. This picture lines up nicely with the standard chronological development of Plato's thought. If one does not accept the standard chronology, however, one can still make good sense of my account. At the most fundamental level, I wish to argue that Plato raises questions and kicks off philosophical projects in some dialogues to which he gives solutions in others. Embracing the standard chronology is beneficial, but not necessary, for explaining this picture. Accordingly, when I speak of Plato's development of *epistēmē*, I wish to refer mainly to philosophical, rather than chronological, development.[25] With this background in mind, let us turn to the arc of the dissertation.

Summary of Dissertation Chapters

My investigation will begin in Chapter 1 with the *Hippias Minor*, a short, tricky dialogue that has been the recipient of (relatively) little in-depth study and much suspicion. The *Hippias Minor* is best known for proposing its own paradoxical thesis: those who do wrong voluntarily are better than those who do wrong involuntarily. This thesis has caused the *Hippias Minor* to be poorly regarded overall, occasionally designated as not even Platonic but more often as fallacious or primarily an exercise in *ad hominem* sophistry. I believe that a closer look, however, will reveal the dialogue to be grappling seriously with issues of virtue and wrongdoing and, in doing so, kick off our primary philosophical endeavor. Socrates feels acutely the pull of the thesis and is seeking to find a way to square it with other intuitions about virtue. If voluntary wrongdoers are indeed considered better in many domains, then why do we show more lenience

---

[25] Aristotle's chronology will be a relatively important part of Chapter 4, so I will save further discussion for that time.

towards those who do *injustice* involuntarily? We see Plato trying to figure out how to explain justice's distinctive status in this regard.

What the *Hippias Minor* shows us, I will argue, is that Plato has made progress towards solving this puzzle, which lies in committing to SP1. Only when we accept that no one does wrong voluntarily (with regard to justice, at least) can we explain the superiority of so many voluntary wrongdoers as well as our attitude towards those whom we consider involuntarily unjust. The issue is that this solution is neither clear nor satisfying—recall that there is something unsettlingly unintuitive about the Socratic "paradoxes"—and we need a reason to think that SP1 is at all plausible. This, I argue, is where Plato's project turns epistemological. He characterizes justice as an *epistēmē*, hence bringing in SP2, and realizes that something about its status as *epistēmē* must explain SP1. Where the *Hippias Minor* leaves us is with the task of searching for this "something" about *epistēmē*. A run-of-the-mill conception of knowledge, even craft-knowledge, will not do the necessary work; we must find some way to understand *epistēmē* that will reveal voluntary injustice to be impossible.

With this task in place, Chapter 2 turns to the final sections of the *Protagoras*. One of the *Protagoras'* primary themes is exploring SP2 in more detail, by arguing for the unity of the virtues. In setting up an argument for identifying courage and wisdom, Socrates discusses the strength of *epistēmē*, and in the so-called "denial of *akrasia*," Socrates concludes two important claims: first, that someone with *epistēmē* will not be overcome by any other motivation, and second, that this *epistēmē* involves some kind of measurement essential for living well. This picture of *epistēmē* is one that will answer the questions posed by the *Hippias Minor*: someone with justice, understood as this kind of *epistēmē*, will never do injustice voluntarily. As Socrates

has here expanded his project to include all of the virtues, we can conclude that the *Protagoras'* treatment of *epistēmē* shows us how it is involved in living virtuously and avoiding wrongdoing.

Before we can conclude that Plato is offering us this conception of *epistēmē*, however, we must face a serious interpretive issue. Plato's arguments for *epistēmē* as strong and pertaining to measurement appear to depend on an argument for hedonism. Socrates endorses hedonism nowhere else in the corpus and often argues against it, so this leaves our conclusions in jeopardy: either we get this picture of *epistēmē* only when Plato (temporarily) embraced hedonism, or the arguments Socrates presents are little more than unserious sophistry against an opponent. In order to rescue the *Protagoras'* conception of *epistēmē*, I will present a new interpretation of the text that shows Socrates to make two parallel arguments about *epistēmē*, both hinging on the role of the good life. One of these arguments takes the common conception of the good life— revealed to be a hedonistic one—but the other applies to the good life in general, whatever it may turn out to be. As a result, we can receive Socrates' conclusions about the latter argument without committing to hedonism, leaving the picture of *epistēmē* unscathed.

We have, then, a notion of *epistēmē* as the strongest cognitive power, and one that exhibits this strength in our lives through measurement. This account is quite general, however, and we might hope that Plato will give us a more satisfying explanation. The most important question is what, exactly, *epistēmē* would measure, a question that we will answer in Chapter 3, when we turn to the *Republic*. *Epistēmē* produces virtuous action, I will argue, by measuring truth. In order to understand this claim, we will need to look closely at both the *Republic*'s treatment of truth and its epistemology. We first see that the treatment of truth is important because of the new role of the Socratic paradoxes; SP1 now appears not about voluntary injustice or vice but about voluntary falsehoods. Truth itself is very different from our typical binary

account. Instead, Plato proposes two types of truth. One of these truths involves how much something turns our soul towards or away from the highest realities. The other type of truth involves how much something corresponds to the world as we perceive it. We shall see that both of these types of truth are measurable, in themselves and compared to the other.

The first thing to note about this picture is that the "things" we are measuring are *doxai*. *Doxai* are never fully true or false, and we must be able to measure their overall truth and falsity in order to decide whether or not to accept them into our souls. Our remaining task is to show that *epistēmē* is in the business of doing this measuring. By looking at Plato's comparisons of *doxa* and *epistēmē* in the middle books of the *Republic*, we will see how *epistēmē* can play this role: both *epistēmē* and *doxa* share in truth and can be related to each other along that scale. This insight is our tool to solve the Problem of the Gap: even though the objects of *doxa* and *epistēmē* are completely separate, one can allow for measurement of the other with regard to what they both share. This conception of *epistēmē* allows us to make sense of some further features of the *Republic*, such as the importance of calculation for *epistēmē*'s acquisition and use and its role as a model for the philosopher-ruler. It also provides a picture of *epistēmē* necessarily accompanied by virtue, which cements our understanding of how it satisfies the Socratic paradoxes.

Chapter 4 will examine these themes in Aristotle. My primary aim will be to understand a notoriously difficult passage in Aristotle's ethics, *Eudemian Ethics* VIII.1. VIII.1 is noteworthy for arguing against the picture we saw in Plato; Aristotle begins by outlining how *epistēmē* can be voluntarily misused and concludes that the virtues are not *epistēmai*. I will show that we can provide a coherent understanding of his argument by understanding *epistēmē* as Aristotle conceives of it in a similar discussion in *Metaphysics* Θ. In *Metaphysics* Θ, Aristotle

treats *epistēmē* as a contrary-use capacity, which is a capacity that can be used to produce opposite states of affairs. We shall see that the argument in VIII.1 proceeds by first establishing that *epistēmē* is a contrary-use capacity and then showing that the virtues and *phronēsis*—practical wisdom—in particular are not.

When we have understood this argument, we are faced with a puzzle: Aristotle has already argued that *phronēsis* is not *epistēmē* earlier in the *Eudemian Ethics*, and in a very different way. I believe that this puzzle will lead us to two conclusions. First, VIII.1 supports the theory that the "earlier" discussion of *epistēmē* and *phronēsis* was not originally part of the *Eudemian Ethics* but was developed as part of Aristotle's later ethical thought. Second, being able to trace this development shows us that Aristotle still feels the pull of the Socratic paradoxes, even when he tries to argue explicitly against them. We may conclude, then, that the Socratic paradoxes were not some obscure and uninteresting part of Plato's thought; their legacy, like *epistēmē* itself, was surprisingly robust.

Chapter 1: Knowledge and Voluntary Injustice in the *Hippias Minor*

Introduction

The *Hippias Minor* depicts a conversation between Socrates and the sophist Hippias after a supposedly great speech on Homer by the latter. Socrates and Hippias start by discussing the relative merits of Homeric heroes and quickly turn to the topic of voluntary wrongdoing. In this context, Socrates proposes and defends one of his most notorious theses:

> Superiority of the Voluntary Wrongdoer (SVW): The person doing something
> wrong voluntarily is better than the person doing it wrong involuntarily.

Hippias resists SVW, declaring at the end of the dialogue that, for all of Socrates' arguments, he does not know how he can agree with him. Interpreters have tended to follow suit: SVW is one of the most surprising and potentially troubling claims in Plato's corpus. A few scholars in the past couple centuries have gone so far as to demote the dialogue as not Platonic on that basis. Unfortunately for these scholars, a citation from Aristotle provides overwhelming evidence that the *Hippias Minor* is genuine, raising the need to explain why such a pernicious thesis would have been proposed.[26]

Those who take the dialogue as genuine have tended to adopt a shared strategy for dealing with SVW: appeal to SP1—or a version thereof, that no one does injustice voluntarily— as a pre-commitment that informs Socrates' arguments. One most views, this leads us to see Socrates' arguments as unserious in some way; SP1 shows that he is only promoting SVW to prove some larger point. When we understand that SP1 is in the background, it is argued, we can understand how Socrates treats SVW. Even if we think that Socrates is, in fact, serious about

---

[26] *Metaphysics* Δ 1025a6-9. See Hoerber (1962) on attempts to excise the *Hippias Minor* from the canon.

SVW the assumption is still that SP1 is important to coloring Socrates' arguments. We should

understand SVW and the *Hippias Minor* in light of Socrates' commitment to SP1.

In this chapter, I aim to present a more compelling approach. As we shall see, Socrates

himself points towards a different understanding of SP1 in the *Hippias Minor*, raising the need to

reexamine its role and, by extension, the *Hippias Minor* as a whole. To that end, we will look

closely at how each interlocutor proceeds and interacts with the other. Socrates and Hippias both

put forward claims about virtue and wrongdoing; viewing these claims together causes Socrates

to recognize his ignorance and desire to investigate the issues further. Contrary to widespread

assumptions, Socrates' aim is not to pick one interlocutor's position over the other; we see

instead that he and Hippias do not contradict each other and seek a larger explanation to

reconcile their claims. This, I shall argue, is the true role of SP1: Socrates is not already

committed to it but instead recognizes it as a potential solution to his puzzle about wrongdoing.

As SP1 is not already assumed, however, we need something to provide further explanation of its

plausibility. To provide this explanation, Plato commences an epistemological project that I will

be tracing throughout this dissertation: finding a notion of *epistēmē* that can show why no one

does wrong voluntarily.

I will start by providing an overview of the *Hippias Minor* and then will discuss how the

scholarship has treated the problems associated with SVW. I divide the recent literature into two

camps: the traditional view that Socrates cannot be committed to SVW, and a more recent view

that Socrates is, in fact, committed to SVW and ultimately rejects Hippias' resistance to it. I

argue that neither of these positions fully grasps Socrates' attitude towards his and Hippias'

arguments, and I provide a preliminary reason as to why: Socrates discusses how he experiences

wavering (πλάνη) throughout the *Hippias Minor*, and this wavering indicates that he wants to

17

accept what seems true to him about both his and Hippias' arguments. Given his uncertainty, we must determine what exactly Socrates' and Hippias' positions are and what parts of each one are compelling. I look closely at the text to highlight some key theses proposed by Socrates or Hippias, which both accept as shared commitments. A surprising observation is that Socrates and Hippias do not disagree much. Socrates argues with SVW that voluntary wrongdoers are better, while Hippias contends that we are more lenient towards those who are voluntarily unjust. All they lack is a way to understand how both of these commitments can be true together.

As I mentioned, the standard view is that when confronted with what SVW would require for justice, we should reject SVW because Socrates hints that he is already committed to SP1. I argue that this position gets things backwards. Socrates is trying to find a larger theory for his and Hippias' views, and one theory that can accommodate both of them involves committing to SP1 as it applies to justice. It is not, as has been assumed, an axiom that colors the discussion; on my view, it is itself a thesis that the *Hippias Minor*'s arguments would lead Plato to accept. Accordingly, we cannot just assume that SP1 will be in place; we need the *Hippias Minor* to indicate to us why it could plausibly hold. I argue that the dialogue's arguments provide a framework for doing so, by highlighting SP2 as important for figuring out justice's special status. In order to understand how voluntary injustice is impossible, we must understand what it means for it to be an *epistēmē*, which involves getting clear on *epistēmē* itself. Though the *Hippias Minor* does not fully determine what is required for *epistēmē* in this context, we see Plato begin to consider it separately from craft knowledge and seek to develop it into something that is powerful enough to support SP1.

18

The Superiority of the Voluntary Wrongdoer

The *Hippias Minor* does not exactly number among Plato's most well-known dialogues; even its name, though a reference to its relative length, imparts a sense of obscurity. Given its somewhat unfamiliar status, I will here outline the main points of the dialogue and then discuss how some of the recent scholarship has tackled the problems it raises.

The *Hippias Minor* can be divided roughly into two parts.[27] At the start of the dialogue, Hippias has just finished a rousing speech on Homer, and Socrates asks Hippias whether he thinks Achilles or Odysseus is the better man. In response, Hippias asserts that Homer made Achilles the best, Nestor the wisest, and Odysseus the wiliest (364c).[28] Socrates presses Hippias to give him a straight answer, and Hippias responds that Achilles is better because he is simple and tells the truth, while Odysseus is wily (πολύτροπος) and a liar (364e-365b). Socrates spends the rest of the first half of the dialogue arguing that the liar and the truth-teller are the same person. A skilled mathematician (like Hippias, apparently), when asked a question of arithmetic, could always tell lies, if he wanted to lie (366e). An amateur mathematician, on the other hand, would often involuntarily tell the truth, even if he wanted to lie (367a). Therefore, the liar and the truth-teller are the same (367c).

Despite these arguments, Hippias still wants to hold onto his initial claim. He admits that both Achilles and Odysseus speak falsely, but Achilles, unlike Odysseus, does so involuntarily. At this point, the second half of the dialogue begins, with Socrates stating a version of SVW:

---

[27] In this chapter, my main focus is on the arguments found in the second half of the dialogue, but my aim is in part to show how the whole dialogue affects the arguments around SVW. To this end, the first half will be relevant especially concerning its discussion of power, to which I will return later on.

[28] Nestor is not a main hero in Homer, and certainly not of such a kind as Achilles or Odysseus. Thus, his inclusion at the start of the dialogue is striking and, I think, not insignificant. Though I do not elaborate further here, I would maintain that the inclusion of Nestor is an early sign of the importance of knowledge in the dialogue in determining who is the best person.

"Didn't those lying voluntarily appear just now to be better than those doing so involuntarily?" (371e7-8). Hippias is incredulous and Socrates agrees that he is also puzzled. Then, however, Socrates raises a number of examples for which SVW would hold, and Hippias accepts that SVW applies to running (373d), wrestling (374a), other physical activities (374b-c), and the functioning of body parts (374c-e), tools (374e), souls of animals (375a), souls of craftsmen (375b-c), and souls of slaves (375c). He follows Socrates relatively easily until they come to the case of their own souls, at which point Hippias complains that the voluntarily unjust cannot be better. Socrates presents an argument for why SVW would also hold for justice, and Hippias is left befuddled rather than convinced; the dialogue ends with Socrates lamenting Hippias' failure, as a wise man, to help him out with this puzzle.

The arguments in the *Hippias Minor*, especially those concerning SVW, have rendered many suspicious. As mentioned earlier, a citation from Aristotle blocks the most extreme manifestations of this suspicion by confirming the *Hippias Minor* as canonical. We should, for our part, try to banish this suspicion as much as possible. Surprising claims from Socrates are surely not an anomaly, and a work focused on "Socratic paradoxes" should be especially loath to shy away from them. In addition, although the *Hippias Minor* is chief among Plato's corpus for advocating for SVW, we find traces of this line of thought elsewhere as well. Near the beginning of the *Crito*, for example, Crito laments that "the majority can inflict not the least but pretty well the greatest evils" (44d3-4). Socrates provides the following response: "Would that the majority could inflict the greatest evils, for they would then be capable of the greatest good, and that would be fine, but now they cannot do either. They cannot make a man either wise or foolish, but they inflict things haphazardly" (44d6-10). The majority is unable to do great harm but acts haphazardly instead. Socrates says tellingly that it would be better if they could commit

the greatest evils. Unlike SVW in the *Hippias Minor*, this *Crito* passage does not include the words "voluntary" or "involuntary." Nevertheless, the talk of not being able (οἷοί τ᾽ εἶναι) to inflict the greatest goods or evils and doing things haphazardly (ποιοῦσι δὲ τοῦτο ὅτι ἂν τύχωσι) indicates that the majority is unable to do bad things consistently, just like the amateur mathematician in the *Hippias Minor*. In some way (which we have not quite yet figured out), being skilled at doing bad things makes one better than a haphazard actor.[29] Our goal, then, is to try to make sense of Socrates' surprising claim within the *Hippias Minor* and where it might appear elsewhere in Plato.

The literature on the *Hippias Minor* has tended to take one of two primary positions regarding SVW. The first and most prominent is that Socrates must reject SVW; it is simply too pernicious a thesis for us to think he would take it seriously. A traditional line of thought claims that the arguments for SVW are undermined by fallacies of equivocation.[30] More recently, scholars have demonstrated that the arguments in the *Hippias Minor* are valid, not fallacious.[31] The arguments' validity, however, does not entail Socrates' commitment to them. Most authors who accept the arguments for SVW as valid consider them instead to be a *reductio* that Socrates has constructed in response to Hippias. Socrates' reasons for forming the argument in this way

---

[29] Though we can already begin to see its importance here, I will return to the relation of power/ability to voluntary wrongdoing in later sections. Thanks to Paul Woodruff for highlighting to me the relationship of the *Crito* passage to the *Hippias Minor*.

[30] See Hoerber (1962), Sprague (1963), Mulhern (1968), and Haden (1997). One of the most cited cases of equivocation has been Socrates' use of "good," which could mean either good at some domain (for which SVW may hold) or morally good (for which it would not). Such fallacious reasoning could not have escaped Plato's notice, so, it is argued, he must have had some purpose in making a faulty argument, such as to re-examine our use of "good" and other key terms.

[31] The first and most notable advocate for this view has been Weiss (1981, 2006), who thinks that the *Hippias Minor* demonstrates an *ad hominem* attack on Hippias and his commitments. For other perspectives on why Socrates' arguments are valid but not endorsed, see Lampert (2002), Balaban (2008), Stefou (2012), Carelli (2016), and Naddaff (2017). More generally, there has been skepticism in the last few decades about morality and a separate category of "moral" goodness operating in ancient philosophy, stemming from Anscombe (1958).

21

vary on each interpretation, though they all share the position that the argument is meant to lead to a false, unacceptable conclusion.

Against these views, the second primary position regarding SVW has recently been advanced by Jones and Sharma (2017). They argue that Socrates is, in fact, committed to SVW, and he seeks to accept it over Hippias' objections. To show this, they pay careful attention to how Socrates interacts with Hippias, arguing that Socrates leads Hippias into uncertainty about Hippias' own commitments. We should conclude that Socrates still wants to endorse SVW; what he needs is a theory that can show why SVW is preferable to Hippias' position. This theory, they argue, comes about through a further thesis that surfaces in other dialogues: one cannot unjustly promote one's own interests. Such a thesis allows Socrates to continue to endorse SVW without its unpalatable implications. If one cannot unjustly promote one's interests, then there is no worry about the person who is an "expert" at justice but does injustice voluntarily. This person would, instead, recognize that injustice will not be to her benefit, so she will only act justly.[32]

One point that all of the above views share is that they recognize the importance of Socrates' final conclusion in support of SVW. Socrates states that someone voluntarily committing injustice would abide by SVW, "that is, if there is such a person" (εἴπερ τίς ἐστιν οὗτος, 376b5-6). However, for SVW to apply, there must be someone who voluntarily does injustice, which, it is argued, Socrates is already committed to denying as a result of holding SP1. By framing SVW in a conditional, we can sidestep its force, and most scholars take this

---

[32] I find Jones and Sharma's account to be more compelling and closer to my own view than the previous interpretations of the *Hippias Minor*, and I follow them on certain points, such as the importance of πλάνη and Socrates' desire for a larger theory. Nevertheless, I find that their position falls short in a number of important ways, which I will explain in more detail in the following sections, starting with what we should make of Socrates' and Hippias' πλάνη.

conditional as a sign that Plato is not really committed to SVW. We will return to this

conditional later on. For now, we have two primary positions about SVW: either it must be

rejected, or Socrates' arguments show it to be superior to Hippias' intuitions. We will find

reason to doubt that either of these proposals is sufficient, however, when we examine how

Socrates is thinking about SVW and where he thinks he and Hippias have ended up in their

debate.

Socrates' and Hippias' Πλάνη

The *Hippias Minor* exhibits signs of being a typical aporetic dialogue: Hippias ends up

at a loss for how to proceed, and the dialogue ends without the interlocutors agreeing on any

clear conclusions. Jones and Sharma, in arguing that Hippias is led into uncertainty, point out

that Hippias' final line displays questioning unease rather than outright disagreement with

Socrates' arguments; they translate 376b7 as "Socrates, I don't know how I'm to agree with you

on these matters."[33] Socrates indicates that he is in a similar confusion, but he has more to say

about what it involves. It is worth showing the close of the dialogue in full:

> Nor do I know how to agree with myself, Hippias; but it is
> necessary that it appears so to us now from the argument. But still,
> as I was saying earlier, I waver back and forth (ἄνω καὶ κάτω
> πλανῶμαι) concerning these things and it never seems the same to
> me. And it is no wonder that I or any other layman wavers; but if
> you wise men waver too, this is certainly terrible for us, if when
> we come into your company we still will not stop from wavering.
> (376b8-c6)

Socrates notes that he and Hippias have both been rendered unable to come to an agreement, but

not in the sense that each has failed to convince the other the truth of his position. Hippias has

been led to see that Socrates' arguments have force, hence that things "appear so to us now from

---

[33] Οὐκ ἔχω ὅπως σοι συγχωρήσω, ὦ Σώκρατες, ταῦτα. By comparison, consider Nicholas Smith's more decisive rendering: "I can't agree with you in that, Socrates."

the argument." Socrates has not won the day, however. He emphasizes that they are both in a state of πλάνη, or wavering—Socrates constantly wavers about these things, and now Hippias is wavering too.[34] Πλάνη itself is important in this picture: it is the last word of the dialogue and it can clarify Socrates' attitude towards what has been put forward and how to respond.[35]

Socrates first mentions his πλάνη earlier in the dialogue, after Hippias raises initial objections to SVW. He characterizes his wavering by stating that Hippias' objections do not seem right to him, but that this is not always the case: "it appears to me, Hippias, the complete opposite of what you're saying [. . .] but sometimes too the opposite of these seems true to me and I waver about these things, clearly on account of not knowing" (372d3-e1). Socrates' claim and his ending speech above highlight three key features of wavering. First, wavering is a sign of a lack of knowledge; although Socrates is not so blunt at the end of the dialogue, Hippias must be ignorant if he also ends up wavering. Second, wavering results when opposite things both seem true to a person, as Socrates says often happens to him about these matters. Third, the ending lines show wavering to involve not being able to come into agreement, not so much with someone else but concerning the various points one would want to accept within oneself. Put together, Socrates and Hippias waver because they feel compelled by the truth of seemingly opposite things and cannot come to an internal agreement about how to proceed. In order to stop wavering in this case, two strategies present themselves: either determine which of the opposites is not true, or determine how the supposed opposites do not, in fact, contradict. Having reached the end of the dialogue, the interlocutors still waver and have not been able to set aside what has

---

[34] Πλάνη is sometimes translated "wandering" in the *Hippias Minor* and elsewhere. For reasons that will become clear presently, I follow Smith's translation and prefer "wavering," emphasizing the two destinations that draw Socrates.

[35] I am grateful to an anonymous reviewer at *Apeiron* for helping me to clarify the implications of the last lines of the dialogue, as well as many other points about wavering that are discussed in this section.

seemed true to them. The second strategy would thus appear more attractive: examine whether the opposing claims could end up being reconciled.

We can see more clearly the desire to accommodate claims from both interlocutors when we recognize how Socrates and Hippias have caused each other to waver. Hippias starts the discussion of SVW with strong convictions, but by the end he displays uncertainty. Something about Socrates' arguments has driven him into wavering.[36] Socrates also indicates that Hippias has brought him to wavering. Hippias' strong objections to Socrates' first proposal of SVW send Socrates into his primary speech on wavering. He remarks at the start that "whenever I converse with someone among those highly reputed in wisdom and to whose wisdom all the Greeks bear witness, I appear to know nothing; for none of the same things seem true to me and you, so to speak" (372b4-c1). This statement may be suffused with at least a little irony, but the main point is serious: Socrates does not have a ready reply to Hippias, and this displays his ignorance. Notably, Socrates wavers because the opposite of what he was saying also seems true to him. Socrates sometimes feels that Hippias is wrong, but he comes to waver when he is faced with Hippias' objections and considers that they do seem true.[37] The end of the dialogue has not relieved Socrates; he still wavers and hence must still feel that Hippias is stating something true.

---

[36] Jones and Sharma argue that Socrates' final argument about justice is what specifically draws Hippias into uncertainty. I think that this is largely right, although the previous discussion does also display signs of Hippias' discomfort, such as his specification that SVW applies only to certain domains as Socrates advances his arguments. He anticipates where the discussion will lead, and it makes him nervous.

[37] Socrates' use of "sometimes" and present-tense verbs throughout the discussion of wavering may imply that he has considered Hippias' arguments before; Hippias would thus not be the first person to cause Socrates to waver about the matters at hand. Whether Hippias himself is unique in causing Socrates' wavering about SVW does not matter much for our purposes. The same conclusions still result if Hippias is merely reminding Socrates of certain intuitions that have seemed true to him; in fact, it may better explain why Hippias need only register his objections as brief outbursts, rather than prolonged arguments, for Socrates to waver.

The *Hippias Minor* is the only dialogue, besides one mention in the *Hippias Major*, where Socrates describes himself specifically as wavering in this way.[38] It stands out as unique, and we should consider Plato's purpose in constructing the *Hippias Minor* in this way. We can contrast it to the many dialogues that show Socrates engaging in elenchus and refutation. In these elenctic dialogues, an interlocutor puts forward a number of proposals for consideration, and when Socrates' examination raises difficulties for each one, the interlocutor must start over and try again.[39] The *Hippias Minor*, on the other hand, already has everything on the table, so to speak. Socrates has managed to convince Hippias that his arguments should be taken seriously, but neither interlocutor has the final say, and Socrates takes Hippias' points seriously as well. Both waver back and forth at the end of the dialogue, and we do not have a proposal that we can simply throw out and start over. We still have instead a set of points that all seem both true and contradictory, and there is a need for further investigation.

---

[38] Πλανῶμαι and πλάνη are not completely absent from other dialogues, however. The terms are often used with a negative connotation, to describe the state of souls who do not have knowledge and are confused by believing obscure things. See, for example, the extensive discussion in *Phaedo* 79c-81d, where being dragged down by the body causes impure souls to waver while completely separating from the body puts a stop to wavering and other ills. Similarly, as we shall see in the next chapter, *Protagoras* 356d claims that the power of appearance causes us to waver rather than to know the truth. The most extensive discussion of wavering comes in *Alcibiades I* (112d-118a), where Socrates shows Alcibiades that he wavers because he is ignorant of the fact that he does not know. If someone knows that she does not know something, she can go to the expert who does know, but if she does not know that she does not know, then she remains confused and wavers about those things (in Alcibiades' case, the just, the beautiful, and the good). I think this discussion in *Alcibiades I* helps illuminate two more features of wavering in the *Hippias Minor*. First, Socrates cannot end his wavering because Hippias is not an expert: as Socrates says at the end of the dialogue, it is concerning if a wise man such as Hippias cannot cure him of his wavering. Second, Socrates does not know what he does not know because, in fact, both his and Hippias' proposals seem compelling. Escaping ignorance does not involve realizing his or Hippias' proposals are wrong so much as determining what about each of them is right and how to fit them into a larger theory.

[39] See, for example, the end of the *Euthyphro*, where Socrates exhorts his eponymous interlocutor that "we must investigate again from the beginning what piety is," not having come much closer to the account they are seeking (15c13-14). Vlastos in particular has contributed greatly to understanding these elenctic dialogues, including (1994) raising the "problem of the Socratic elenchus." Briefly, Socrates shows that an interlocutor's proposal is inconsistent with other premises that the interlocutor holds. All that this implies, however, is that the conjunction of premises and proposal is false, not that the proposal in particular must be. I deem it a strength of the *Hippias Minor* that it avoids this problem by making it explicit: we cannot hold conflicting beliefs, so we must carefully examine our commitments in order to deliver a final verdict. Plato will use this strategy at other points of interest to us, such as the *Protagoras'* debate about virtue, which we will examine in the next chapter. These "debater" dialogues, in contrast to the elenctic dialogues, help us out by already providing (seemingly) opposing arguments and inviting us to consider each one further.

Jones and Sharma are the first interpreters to emphasize the importance of πλάνη, recognizing that both Socrates and Hippias find themselves in uncertainty as a result of their arguments. Nevertheless, they draw the wrong conclusion from Socrates' wavering. They write that "the point is not that Socrates feels compelled to 'save the phenomena' of ordinary usage by somehow accommodating them into his theory [. . .] lacking a broader theory of voluntary action, Socrates is not prepared to say without reservation that common practices are misguided" (121-2). What this indicates is that the common practices *are* misguided; Socrates will cease his wavering once he realizes how he can reject Hippias' objections. But if Socrates is truly wavering back and forth, sometimes viewing Hippias' position as correct, then it seems that Socrates does not wish to reject Hippias' arguments after all. Rather, he does find himself drawn to "save the phenomena;" ceasing his wavering must involve considering what about them seems true and how that might be included in a larger theory. We should hesitate, therefore, to say that Socrates' wavering shows him to prefer his position over Hippias'.

One way in which confusion can arise about Socrates' wavering is through a lack of clarity about the positions that seem true to him. Although Socrates describes himself as wavering back and forth between opposite things, simply saying that Socrates promotes SVW and Hippias rejects it does not give us a particularly helpful picture of the discussion around SVW. Something about Socrates' arguments seems compelling to Hippias, and vice versa; if we can determine which points in particular are so compelling, we might be able to see whether there is a way for them to come into agreement and put an end to wavering. It is to this task that we will now turn.

Justice and SVW

Socrates and Hippias often feel that they are in tension, but they also agree with each
other at important points throughout their discussion. Here I will illuminate some of Hippias'
and Socrates' primary shared convictions, which will be recognizable to a reader familiar with
the dialogue but are worth examining in more detail. I take as a "shared conviction" a point that
one of the interlocutors makes to which the other feels drawn to accept, without great hesitation
or outright disagreement.[40] These convictions manifest themselves into three theses proposed by
Socrates (S_1-S_3) and one proposed by Hippias (H_1):[41]

> S_1: In some domains, a voluntary wrongdoer is better than an involuntary wrongdoer.
> S_2: Someone who is wisest and most powerful in a certain domain is also best in
> that domain.
> S_3: Justice is either a power, *epistēmē*, or both.
> H_1: We are inclined to be more lenient to those whom we consider involuntarily unjust.

We will examine each thesis in its context to see how it relates to SVW and where it could

conflict with the others. When we figure out where Socrates and Hippias agree, we can see more

clearly where tensions might arise, and this will leave us better prepared to locate a solution.

To begin, Hippias does not resist SVW in every case; in fact, he agrees that it applies to

many domains.[42] Socrates' first argument concerns running, and Hippias agrees that the runner

who runs slowly voluntarily is better than the one who runs slowly involuntarily. Neither slow

runner is doing a good thing; rather, "the good runner voluntarily does this bad and shameful

---

[40] Hippias is relatively transparent concerning what he accepts easily, hesitates about, and completely rejects.
Therefore, I take his reactions to be a suitable guide to what he thinks could square with his other convictions.
[41] I do not wish to contend that these commitments are the only ones shared by the interlocutors. They are, however,
important commitments concerning SVW and, I believe, sufficient to target where Socrates and Hippias find tension
in need of a resolution.
[42] I use "domain" here not as a technical term but rather to identify whatever area is under investigation with regard
to SVW. Socrates' examples vary widely in kind, from sports to functioning of body parts to animal souls; each of
these is a "domain" for which SVW would hold. To say that one is "better" is to say that one is better in that
domain, for any domain. Such an understanding can be used to counter charges of equivocation concerning moral
and non-moral senses of "better:" SVW is domain-specific all the way down.

thing, but the bad runner does it involuntarily" (373e4-5). Their relative merits, however, are uncontested: the involuntarily slow runner is worse (373e6-374a1). We might be suspicious of the voluntarily slow runner and hesitate to say that she is better *tout court*. Perhaps she was bribed to throw the race or is losing as part of some elaborate joke, and we would be tempted to say she is a jerk or corrupt and certainly not a good person. This does not change the fact that she is good at running compared to the "more worthless" (πονηρότερος) involuntarily slow runner: with regard to the domain of running the voluntarily slow runner is better.

Hippias agrees that this assessment is correct "in a race, at least," applying only to the domain of running (374a1). Socrates thus shows Hippias that they both accept the following thesis:

S$_1$: In some domains, a voluntary wrongdoer is better than an involuntary wrongdoer.

Hippias is clearly hesitant to apply this thesis to any domain whatsoever, so a general version of SVW is not a shared conviction. But he agrees that it holds for running, and he will continue to agree for many other examples, so it holds in at least some cases. The argument for S$_1$ goes by fairly quickly in the text, but we can find support for this line of reasoning by recalling the earlier half of the dialogue. There, Socrates argues that the truth-teller and the liar are the same person. When one is an expert in a certain domain, as Hippias claims to be for many domains, then, and only then, can one consistently speak falsely about its subject matter. Therefore, "you say that liars are powerful and intelligent and knowledgeable and wise for whichever things they are liars" (366a3-4). Socrates asks whether an expert like Hippias is "only wisest and most powerful or also best in the things in which you are most powerful and wisest" (366d3-5). Hippias responds that "I am definitely the best" (366d6). The two agree on the superiority of such a person:

S₂: Someone who is wisest and most powerful in a certain domain is also best in
that domain.

Unlike in the case of S₁, where he must examine each domain in turn, Hippias does not hesitate

to assent to S₂ as a general rule.

Later on, Hippias accepts many analogous arguments about SVW's application to other

domains, from wrestling to tools to crafts. Eventually, however, Socrates goes too far:

> Socrates: Well? Would we not want to have our soul be the best
> possible?
> Hippias: Yes.
> Socrates: Then will it be better if it does bad things and errs
> voluntarily or involuntarily?
> Hippias: But it would be terrible, Socrates, if those doing injustice
> voluntarily are better than those doing so involuntarily. (375c6-d4)

Although Hippias felt comfortable accepting SVW for many domains, he is adamant that

applying SVW to our souls—and particularly to justice, which he brings up unprompted—is

highly problematic. This is the second of two outbursts where Hippias expresses his discomfort

with this position. The first outburst comes about immediately after Socrates states SVW for the

first time:

> Hippias: But how, Socrates, could those voluntarily doing
> injustice and voluntarily scheming and doing bad things be better
> than those doing such things involuntarily, for whom there seems
> to be much forgiveness, if someone unknowingly does injustice or
> lies or does some other bad thing? Surely also the laws are much
> more severe for those doing bad things and lying voluntarily than
> for those doing so involuntarily. (371e9-372a5)

Hippias states his reasons for resisting SVW: in personal and legal matters, we are much more

forgiving if we determine that someone has done injustice involuntarily. Hippias' claims send

Socrates into his long speech about his wavering. I noted earlier that Hippias' outburst causes

Socrates to waver because he notices that something Hippias says seems true to him, and now we

can see precisely what causes Socrates' wavering. Socrates is inclined to accept the following common intuition:

H_1: We are inclined to be more lenient to those whom we consider involuntarily unjust. Although Hippias accepts Socrates' argument for running and other domains without problem, H_1 prevents him from doing the same for the domain of justice. The tension that has been building now becomes explicit in the very last page of the dialogue, where Socrates attempts to run an argument opposed to these intuitions, in support of voluntary injustice.

In the first step of his argument, Socrates makes a surprising move, asking whether justice is "either some power or *epistēmē* or both" (375d8-9). Socrates ventures that it is "necessary" that justice be so categorized, and Hippias says "yes" immediately. Both parties present themselves as committed to this thesis:

S_3: Justice is either a power, *epistēmē*, or both.

Is this categorization compelling? One might think of justice chiefly as a virtue, not necessarily as a power or *epistēmē*. But what it means to be *epistēmē* or power on this picture has not itself been defined. All that Socrates gives us is that "the more powerful soul [is] more just" and "the wiser soul [is] more just" (375e2-5). This would, at least, seem to follow from S_2. We will look at S_3 and its importance more closely later on.

The just soul is the one that is wise and powerful. Therefore, Socrates notes that it "has more power to do both fine *and* shameful things for every action" (375e9-376a1). Socrates' next move connects craft and injustice: "Whenever [the more powerful and wiser soul] does shameful things, then, it does so voluntarily through power and craft; these things appear to be part of justice" (376a2-3). Let us pause briefly on this point. The arguments of the *Hippias Minor* and the Socratic paradoxes both deal with voluntary and involuntary action, but Plato does

31

not spend much time elaborating on the notion of the voluntary itself. Socrates' claim here is

therefore significant in laying out what would make an unjust action voluntary. He names two

components of voluntary action: first, the power to do something unjust, and second, the craft or

know-how to make such an action happen. We should keep this description of the voluntary in

mind, for, as we will see later on, explaining the impossibility of voluntary wrongdoing will

involve getting clear on these components. To return to the point at hand, we see from S_2 that to

have these components means that one is best; the people who are best are those with the power

and craft to do something voluntarily.

In the next lines, Socrates concludes that "the more powerful and better soul, whenever it

does injustice, will do injustice voluntarily, and the worthless soul will do so involuntarily"

(376a6-7). Spelled out more starkly, "the one who voluntarily errs and does shameful and unjust

things, Hippias—that is, if there is such a person—would be no other than the good person"

(376b4-6). Hippias expresses his final hesitation at this conclusion, which he now feels to be in

tension with his earlier commitments; he ends the dialogue unsure of whether he can reconcile

his stance with Socrates' arguments. Something has gone wrong in the conversation; though

Hippias and Socrates agree on many points, they still end up forced into wavering and unease.

Let us take stock quickly of what Socrates and Hippias hope their philosophical views

can include. We have four shared theses, three presented by Socrates and one by Hippias. The

most important thing to note is that these statements do not contradict. It is only when we apply

the S-statements directly to the domain of justice that we run into tension with H_1 and have

trouble explaining our lenience towards the involuntarily unjust. These theses set up the heart of

the problem and make manifest the disagreement between Socrates and Hippias: something

about justice should allow for H_1, despite SVW. What Socrates and Hippias need now is this

"something about justice:" what it is, and how it will work. It is here, I shall argue, that SP1 comes in, though in a very different way from what has previously been assumed.

Hippias and the Socratic Paradox

As we saw earlier, scholars who take a position against SVW usually turn to Socrates' final argument for recourse: Socrates tells Hippias that the voluntarily unjust person would be better "if there is such a person." Many scholars have argued that this falsifies SVW because it indicates that there isn't anyone who would ever do injustice voluntarily. They resort to a version of SP1 as it applies to justice: no one does injustice voluntarily.[43] Those who find SVW unpalatable can breathe a sigh of relief; Socrates is hinting that he is not serious about the whole thing. Now that we have examined Socrates' and Hippias' shared commitments, however, I believe that we will be able to recognize this approach as misguided. Make no mistake: SP1 is very important for understanding the *Hippias Minor*. But we can now make two observations that will allow us to understand its true role.

The first observation is that SP1 does not contradict SVW. If no one does injustice voluntarily, then there is no situation in which we would have to say that SVW is incorrect. SP1 merely provides some preliminary explanation for how SVW might deal with justice; it does not negate it, even in this domain. As such, it would be incorrect to consider SP1 as showing part or all of the *Hippias Minor* to be a *reductio* of SVW. The argument has not produced an impossible conclusion, and as a result we do not have to reject SVW or any related theses. In light of this fact, we are invited to reconsider what the relationship between SP and SVW is supposed to be.

---

[43] We need not trouble ourselves that SP1 does not apply to all wrongdoing. Part of our project will be to trace the scope of SP1 through our dialogues of focus. We can see that Socrates' starting position on SP1 is rather narrow in considering only justice. This does not mean, however, that he considers justice to be its only content, even at this point. Rather, Socrates' focus reflects his purposes for this particular argument.

Recall that Socrates is searching for an end to wavering by finding a way to accommodate theses that he and Hippias both find compelling—theses that we have shown to be potentially reconcilable. Socrates hints at SP1 only after working through many iterations of SVW and considering what would happen in the case of justice. In light of these points, I propose that Socrates' puzzle around SVW is what provides reasons for him to commit to SP1. Socrates does not state SP1 explicitly but only raises its possibility by wondering—indeed even doubting, as the force of εἴπερ might suggest—whether there are any doers of voluntary injustice.[44] He has considered and presented to Hippias so many examples of voluntary wrongdoing being preferable to its involuntary counterpart, to the extent that it would be untenable to deny SVW wholesale. Rather than reject SVW, he can instead consider why it might not be problematic in the case of justice, and this involves showing that there is no situation in which it will apply. On this reading, SP1 is not always lurking in the background, waiting to give color to the arguments of the *Hippias Minor*. Rather, it is a principle which Socrates has not already accepted but now feels drawn towards; he can embrace it as a way to reconcile his and Hippias' arguments. As such, the *Hippias Minor* has an important role to play in Plato's treatment of SP1: it shows a primary motivation for Socrates to commit to it and, perhaps, even the point at which he accepts it. SP1 does not indicate that SVW is wrong; SVW indicates that SP1 is right.

Of course, not all scholars claim that SVW is not to be taken seriously. Jones and Sharma, after all, argue precisely against this point. Though they do not say so explicitly, I expect they would also accept that SVW and SP1 can both be true together. Like other scholars,

---

[44] The absence of a clear statement of SP1 is, in my opinion, another point against considering these arguments a *reductio*. If Plato wishes to show that we have been led into an impossible conclusion, he is indicating as much with only the subtlest of winks.

however, they take SP1 to be one of Socrates' pre-commitments, finding the *Hippias Minor* to support indirectly "the Socratic thesis" laid out by SP1 (128). SP1 is important to them because, as they argue, Socrates' arguments create an obvious immoralist problem about the superiority of the voluntarily unjust. This problem is not so obvious, however, given their interpretation of how Socrates treats SVW and Hippias' objections to it. As we saw, Jones and Sharma take Socrates to waver because he is temporarily blinded by Hippias' arguments, but a larger theory for SVW will be able to put these objections to rest. To see SVW as an immoralist problem, however, is precisely to admit that Hippias' intuitions should not be put to rest. Socrates *should* try to save the phenomena to which Hippias appeals and causes Socrates' wavering. On Jones and Sharma's view, there is no reason why Socrates could not just double down on accepting SVW for justice without qualification—an uncomfortable claim, perhaps, but only a problematic one if we think we need to reconcile it with H_1. The account I present, by contrast, can better explain why Socrates is searching for a solution at all. In doing so, it also enables us to understand how SP1 is really working in the *Hippias Minor*.

One reason why scholars have not so far recognized the role of SP1 is, I expect, that it underlies a startling claim: Socrates is not already committed to SP1 in a "Socratic" dialogue such as the *Hippias Minor*. One might think, after all, that to the extent that Socrates had any prior philosophical commitments of his own, surely a Socratic paradox like SP1 would have been first among them. A full treatment of the "Socratic-ness" of SP1, as I have said, is best saved for further research, but I think these considerations give us reason to doubt that SP1 must have been inherited from Socrates as a pre-commitment present primarily in Plato's early thought. Furthermore, we can find other places in the early Socratic dialogues where Socrates does not seem to espouse SP1. Let us consider two such examples below.

The first example can be found in Socrates' argument against Meletus in the *Apology* that

he should not be punished for corrupting the youth. Socrates argues that "either I do not corrupt

the young, or if I do, it is involuntarily" (25e6-26a1). This dichotomy, of course, does not

indicate strongly one way or the other whether there are any voluntary wrongdoers, only that

Socrates himself is not one of them. His next claim, however, is more telling: "if I corrupt them

involuntarily, the law does not require you to bring people to court for such involuntary

wrongdoings, but to get hold of them privately, to instruct them and exhort them" (26a2-4).

Socrates' statement here is largely in line with Hippias' H_1, that the courts are much more lenient

to involuntary wrongdoers. In fact, Socrates goes further by claiming that involuntary

wrongdoing should not be the concern of the courts at all; this would lead us to conclude,

however, that the law should be concerned with something else—voluntary wrongdoing.

Socrates' argument as a whole is difficult to understand if we think that he is already committed

to SP1 in the *Apology*: he argues that he is not a voluntary wrongdoer and implies that the law is

for voluntary wrongdoers, neither of which is necessary or on target if he already thinks there are

no voluntary wrongdoers.[45] We may try to give a complex account that shows how Socrates can

speak in this way despite a commitment to SP1, but it seems to me more natural to think that

---

[45] See Brickhouse and Smith (2018), who give a detailed analysis of the passage in question and how it presents problems for those who think that Socrates is committed to SP1. They forcefully block one way out of this problem: concluding that there is no place at all for the laws and punishment since all wrongdoing is involuntary. Rather, they argue, Socrates does see a need for some types of punishment, and so the viable options are either to accept that the philosophical commitments in each of Plato's dialogues may be incompatible or to find some new understanding of what SP1 might mean. The fact that the first option is even on the table speaks to the depth of the worry about this passage: the inconsistency is not at all easy to explain.

Plato may not have yet grappled with SP1 to the extent that he concludes voluntary wrongdoing

to be impossible.[46] If he had, he could have just as easily used SP1 to make his larger point.[47]

We can find a second example if we return to the *Crito*. Socrates poses the following

questions: "Do we say that one must never in any way do wrong voluntarily, or must one do

wrong in one way and not in another? [. . .] When one has come to an agreement that is just with

someone, should one fulfill it or cheat on it?" (49a4-5; 49e6-7). It would be strange to claim that

we must never do injustice voluntarily if SP1 is already one of our commitments; it is like saying

(though SP1 does not seem to be about physical possibility) that a human must never bench press

a Boeing 737. It is difficult to see how "must" in these cases could have any normative force,

since what it means to be justice (or a Boeing 737) would already preclude humans from ever

acting in this way. Perhaps Socrates thinks that Crito does not share SP1 and so he frames the

question normatively, but, again, it would be more natural to think that Socrates is not always

operating with SP1 as a background commitment. Consequently, when the *Hippias Minor*

arrives at SP1, it is arriving not at an assumption but at a solution.

---

[46] Brickhouse and Smith choose the first option, building on their (2010), by presenting a difference between a self-harming action (which is always involuntary) and an other-harming action (which may be voluntary); the same action may be both self- and other-harming, hence it is involuntary in one sense and voluntary in another. I would agree that we see places in Plato's corpus where Socrates is concerned with self-harming actions, such as in his arguments about desiring the good in the *Meno* (77e-78b). The discussion in the *Apology*, however, does not seem to me to be concerned with what it means for an action to be voluntary and how that might relate to self-harm. An account that explains voluntary action in these terms is needed primarily if we think that SP is a background commitment in the *Apology*. If we call that into question, as I have, the need largely disappears. It is furthermore not worrisome that the conditional conclusion of SP1 in the *Hippias Minor* is not complex and developed compared to Brickhouse and Smith's version. On the interpretation I am presenting here, Plato is in the beginning stages of determining what such a commitment might mean; he makes preliminary moves towards further explanation, as we see in the next section, but he does not yet give a full theory, and we should not necessarily expect him to do so. I am grateful to an anonymous reviewer for pressing me to clarify these points.

[47] Socrates' defense speech is, of course, mean to be provocative and often *ad hominem* against his accusers. But this would be no reason that he would not have used SP1 to argue for his innocence, since it achieves that task equally well. In fact, if he is aiming for maximum shock value in his striking arguments, using SP1 would do the trick even more effectively.

The proposal that Plato presents reasons to commit to SP1, rather than (merely) inheriting them from his teacher Socrates, strengthens our overall thesis for the importance of the Socratic paradoxes in Plato's thought. To be sure, there is no reason that Plato could not take a claim like SP1 as axiomatic and attach great significance to such a doctrine on those grounds. Without any evidence, however, that Plato himself saw good independent reason to embrace it, SP1 could be as quickly abandoned as adopted. If we take the view that Plato's philosophy comes into its own and moves away from Socratic teaching (whatever we might take that to be), we would not feel great cause to see the Socratic paradoxes as fundamental ingredients in these developments. But if, on the other hand, we see SP1 not as an axiom but at the center of parts of Plato's reasoning, we would do well to consider its role as we move forward.

This brings us to our second observation about SP1: while it can be a solution to the tension in the *Hippias Minor*, on its own it can get us only so far. SP1 has traditionally acquired the name of "paradox" because it is a surprising and unintuitive claim. If Socrates is to commit to SP1 because it will resolve his and Hippias' wavering, he needs to find some plausible grounds on which to do so. The *Hippias Minor* does not take SP1 as a brute fact, and so we should hope to find some indication in the dialogue that Plato is considering how to support it.[48] What makes SP1 hold, allowing it not to contradict SVW while also explaining Hippias' conventional intuitions about justice?[49] The *Hippias Minor* will not give us a full solution, but I argue that we see Plato start to develop an answer.

---

[48] I take Jones and Sharma's move to answer a similar question (though I am skeptical about their ability to make such a move at all, see above), where they argue that one can never unjustly promote one's own interests, to be ultimately unsatisfying. Though accepting SP1 as explanatorily basic provides less of a need to explain its plausibility, it is still troubling that Plato does not give any indication that he is thinking along those lines in the *Hippias Minor*. At the least, I would consider an account that shows the *Hippias Minor* itself to be making progress towards an answer to these questions to be superior to one that does not.

[49] One might object that Hippias does not share any intuitions about SP1 being right and would certainly think that voluntary injustice is possible. I think, however, that committing to SP1 would allow Socrates to explain Hippias' intuitions that what we typically call "voluntary" injustice is worse than conventional involuntary injustice. Recall

Knowledge and Voluntary Injustice

As we have seen, Socrates implies that a larger theory can still accommodate both SVW and Hippias' resistance: justice may be the sort of thing that never admits of voluntary misuse. What is missing is an explanation for how this theory would fit together and what guarantees that SP1 will work as a solution. Evidently, a whole range of domains, from running to crafts, works in the way SVW dictates, without the need for something like SP1. We need to find out what makes justice different.

Hippias has implied that there is something about justice that makes it special and immune to the implications of SVW. But one of the only things Socrates proposes about justice is the shared commitment S_3, that it is a power, *epistēmē*, or both. One proposal could be that S_3 is what we are meant to deny; in order to know that justice evades SVW, we must realize that justice is not a power or *epistēmē* but something else entirely. One might be skeptical of this move, however. First, Hippias and Socrates both accept S_3 without any hesitation. In fact, Socrates is the one who proposes that justice is a power and/or *epistēmē*; if Plato is trying to construct a *reductio* to refute this proposal for justice, it would make more sense to have Hippias be the one to put forward what he thinks justice is. Second, we do not have a clear account of what justice is supposed to be, to the extent that we could deny each part of this threefold

---

that Hippias' resistance to SVW applying to justice stems from H_1, that we are more lenient towards those whom we see as involuntarily unjust. If justice as *epistēmē* should guarantee SP1, as I argue in the next section, then our attitude towards different types of injustice can be explained in terms of the severity of different types of ignorance. Take an example of a person who does not pay for her dinner at a restaurant. In an "involuntary" case, perhaps she thought she had paid and left without realizing it; she would be ignorant of the fact that she had not paid for dinner. In a "voluntary" case, she knew that she did not pay but left anyway; one could argue that she was ignorant of what justice dictates concerning interactions with other members of society. Though both cases display an injustice, one could say the "voluntary" case requires harsher punishment because the ignorance is of a more serious nature, namely of justice itself. Both examples of injustice, however, would be involuntary on this picture, because they come about as a result of ignorance. This proposal does not require voluntary wrongdoing to be explained only in terms of ignorance, but it does involve raising the standards for what counts as the knowledge possessed by an agent for a given action, which we will see more clearly in the following chapters. Thanks to Jeremy Henry and Rosemary Twomey for pressing me on this point.

characterization. Socrates goes on to argue that a soul that has the power to do justice is more powerful than one that does not; likewise, a soul that has *epistēmē* of what is just is wiser than one that does not (375e2-376a2). Power and *epistēmē* are required for doing justice, and given that they are all we have so far posited about it, they should, we hope, point us to how we should understand justice and how it can evade SVW's implications. To state the target more clearly, if justice is the sort of thing that abides by SVW in a special way (H_1 and SP1), and if justice is an *epistēmē*, power, or both (S_3), then something about justice's status as a power or *epistēmē* should explain why it gets this special status.

I propose that the arguments of the *Hippias Minor*, especially those of the first half of the dialogue, suggest that its status as a power cannot do this work. The dialogue begins with a comparison of Achilles and Odysseus, using the comparison as a basis to argue that the liar and the truth-teller are the same person. Hippias assumes that Odysseus' πολυτροπία, or wiliness, makes him worse than Achilles, but Socrates shows that a πολύτροπος person is someone who has the power both to lie and to tell the truth.[50] Socrates first asks Hippias whether it is true that "liars do not have the power to do anything, just like sick people, or that they have the power to do something;" Hippias responds that "I say they have quite a lot of power to do many things, especially to deceive people" (365d6-8). Liars have the power to do many things because having power does not mean that one will do good. Socrates suggests as much in his definition of what it means to be powerful: "each person is powerful, who does whatever he wants, whenever he wants; I mean not prevented by sickness or these sorts of things, but just as you have the power to write my name whenever you want, this is what I mean" (366b7-c3). These conditions for

---

[50] Mulhern (1968) emphasizes that πολύτροπος is often read negatively as inclining "more to one *tropos* than to another. The rendering must suggest no more than that an ability is possessed" (284). Though Hippias treats the adjective as having a negative connotation, Socrates' point is that we should read it (and the translation "wily" above) with less negative baggage that is usually attached to it. At its core, the term is about ability, not inclination.

having power are not particularly strong, and they certainly do not seem to be accompanied by any normative baggage. To have the power to deceive, someone would not only "be the most powerful in lying" but also "be the most powerful in telling the truth" (367b6-c2).

As Socrates argues in the *Hippias Minor*, Achilles does not have the power to lie or tell the truth: he asserts that he will leave the war but he is unable to make that assertion true. Odysseus, on the other hand, does have the power to achieve his goal of eventually returning home, often lying along the way.[51] Unless we want to say that Odysseus' power to deceive makes him the better hero, which Hippias would surely resist conceding, we must agree that power alone does not make someone do good things over bad ones. The power to tell the truth is also the power to lie, and as such it can produce good actions or bad ones.[52] Having the power to do good or bad does not determine whether or not one will, in fact, do good, let alone explain why one acts as she does.[53] Accordingly, the fact that justice is a power cannot explain why no one does injustice voluntarily.

With this in mind, we are left with the other option for solving the puzzle: understanding justice as an *epistēmē* is supposed to explain SP1. In this way, Socrates introduces SP2 and

---

[51] Lévystone (2005) argues that, though Odysseus was not always held in high regard, Socrates' students saw the epic hero in a more positive light, even drawing parallels between Odysseus and Socrates himself. It is not surprising, therefore, that Plato might be keen on showing the more compelling aspects of Odysseus.

[52] The thought that power alone does not determine one's actions is not unique to the *Hippias Minor*. Socrates quickly argues against virtue as a power in *Meno* 78c-79b. If virtue is the power to acquire good things, as is proposed, we cannot leave it at that but must add that it is the power to acquire good things virtuously. Though this is obviously a circular emendation, the point is clear: one cannot just have the power itself. Another, though quite different, example is Socrates' claim at *Gorgias* 466d that orators and tyrants have the least power in the city. Though Socrates argues that having (true) power will mean that one does good things, one only has true power when one has knowledge of what would be the best thing to do. In a sense, one might argue that power will make one do good or just things, but what really is doing the work is knowledge, without which we would not have power at all. In both cases, power itself is not responsible for doing good actions rather than bad ones. See Penner (1991) for a helpful discussion of how to think about power in the *Gorgias* passage.

[53] It is worth noting that this does not mean that justice is not a power of some sort. If voluntary injustice is impossible, then justice clearly cannot be only a power; our remaining options, however, are that it is an *epistēmē* or both an *epistēmē* and a power. It could very well be the latter option; what the above shows is merely that the "power" part of that definition cannot on its own support SP1. I do not wish to consider here whether justice is an *epistēmē* or both an *epistēmē* and a power, though I do find it significant that Plato will elsewhere call *epistēmē* itself a power (see, for example, *Republic* 477d-e).

shows its connection to SP1. In order to understand the impossibility of voluntary injustice, we must embrace the fact that justice is an *epistēmē*. The problem is that it is not yet clear how *epistēmē* helps us understand SP1. Indeed, we do not know much about what *epistēmē* is supposed to be. Before the argument about justice, we see Socrates mention in conjunction with *technē*, or craft, as something that abides by SVW unproblematically: the soul that is "better at playing the lyre and flute and everything else in the crafts and sciences" (κατὰ τὰς τέχνας τε καὶ τὰς ἐπιστήμας) will do bad things voluntarily (375b8-c1). The "crafts and sciences" conception of *technē* and *epistēmē* does not indicate that either *technē* or *epistēmē* would be able to show how justice could act as dictated by SP1.

We see a shift, however, during the final argument about justice. After Socrates establishes that justice must be a power and/or *epistēmē*, he lays out two conditions, as we saw earlier, for voluntary injustice. It is worth examining the full exchange in both English and Greek:

> Socrates: Wasn't this more powerful and wiser soul shown to be better and to have more power to do both fine *and* shameful things for every action?
> Hippias: Yes.
> Socrates: Whenever it accomplishes shameful things, then, it does so voluntarily through power and craft; and these things appear to be a part of justice, either one or both of them.
> Hippias: It seems so. (375e8-376a4)

> ΣΩ. Οὐκοῦν ἡ δυνατωτέρα καὶ σοφωτέρα αὕτη ἀμείνων οὖσα ἐφάνη καὶ ἀμφότερα μᾶλλον δυναμένη ποιεῖν, καὶ τὰ καλὰ καὶ τὰ αἰσχρά, περὶ πᾶσαν ἐργασίαν;
> ΙΠ. Ναί.
> ΣΩ. Ὅταν ἄρα τὰ αἰσχρὰ ἐργάζηται, ἑκοῦσα ἐργάζεται διὰ δύναμιν καὶ τέχνην· ταῦτα δὲ δικαιοσύνης φαίνεται, ἤτοι ἀμφότερα ἢ τὸ ἕτερον.
> ΙΠ. Ἔοικεν.

The more powerful and wiser soul is better (ἀμείνων), but Socrates associates it as having only power (δυναμένη) when doing shameful actions. The mechanism that would do voluntary injustice in particular is power and craft (δύναμιν καὶ τέχνην). This is the first time that justice is described as a *technē*; at this point, *epistēmē* drops out of the argument and does not return. Although talk of power remains constant throughout this part of the argument, once voluntary injustice is under consideration, the wiser soul is changed to be the soul that has craft.

Though this may seem like a small linguistic point, I believe that it is significant for uncovering Socrates' intuitions about justice and *epistēmē*. When Socrates is describing what justice is, he is inclined to think of it as an *epistēmē*, along the lines of SP2. When, on the other hand, he is thinking about how voluntary injustice would be preferable, he is inclined to think of it as a *technē*. Socrates' intuitions about *technē* and voluntary wrongdoing are not unique to this argument: he indicates elsewhere that *technai* do not necessarily incline towards the good.[54] As with power, Socrates does not view *technē* as something that can do the work of supporting SP1.

Socrates' careful treatment of *epistēmē*, however, indicates that he is not quite comfortable thinking of it as something that can allow for voluntary wrongdoing. If justice is an *epistēmē* but does not work as a "crafts and sciences" *epistēmē* would, then *epistēmē* itself becomes our focus: we must determine what sort of thing it is such that acquiring *epistēmē* of justice will ensure that we will not commit injustice voluntarily. As such, having *epistēmē* must involve something much stronger than a typical conception of craft knowledge. The *Hippias*

---

[54] See Hulme Kozey (2019), who argues that *technai* in the *Gorgias* and *Phaedrus* are not necessarily good-directed. I am not the first to suggest that Socrates does not wish to treat justice as a craft in the *Hippias Minor*. As with SP1, however, I wish to argue for the role of the *Hippias Minor* in formulating these views; its arguments are not reductive but themselves provide motivation for a new understanding of what *epistēmē* must be so as for SP1 to work. Socrates' attitude towards *technē*, furthermore, does not mean that he does not develop a notion of *technē* elsewhere that would preclude voluntary wrongdoing and be a suitable candidate for justice. I would simply argue that the *Hippias Minor* gives us no indication that he is looking to make these developments, whereas we do find him looking to develop *epistēmē* as a separate notion in this way. I am grateful to an anonymous reviewer for pushing me to clarify the relationship of *technē* and *epistēmē* in this part of my argument.

Minor does not yet have a full account of *epistēmē* and justice that can do this work. But by the end of the dialogue, Plato has reached two tasks for further inquiry: first, that we must clarify justice's relationship to *epistēmē* in order to explain how justice relates to SVW, and second, that we must rethink *epistēmē* in order to do so. A crafts-and-sciences *epistēmē* is not enough; *epistēmē* that explains how justice works the way it does must be a stronger notion, a notion that now remains to be worked out.

Conclusion

The *Hippias Minor* does not have all the answers. Rather, it represents the start of a project of formulating a theory in which its arguments can reside. Socrates recognizes a number of points that the theory should accommodate, points raised by both Socrates and Hippias that initially appear contradictory and cause both interlocutors to waver. The four shared commitments identified in this paper help guide us towards what is still needed to resolve the tensions that the dialogue presents. Socrates argues that voluntary wrongdoing is preferred in many domains, including crafts, but Hippias counters that we must explain our more forgiving attitude towards those who are involuntarily unjust. In order to reconcile these views about SVW and justice, we should recognize these points as compatible and commit to SP1, that no one will do injustice voluntarily. We thus see Plato providing reasons for accepting SP1 and indicating its philosophical importance.

The task remains to figure out how to support such a proposal. One thesis in the *Hippias Minor* that points us in the right direction is that justice is a power or *epistēmē*; these characteristics should explain why voluntary injustice does not occur. Power cannot provide such an explanation; a person with power alone is not inclined in any particular direction with regard to justice. We need a notion of *epistēmē* that will preclude voluntary injustice, and in the

44

Hippias Minor Plato has set up the project of determining what that notion is. This will involve a further investigation into SP2, the Socratic paradox that virtue is knowledge, in the hopes of finding out what about this virtue-*epistēmē* guarantees that no one does injustice voluntarily.

We see, then, that SP1 and SP2 are key to figuring out how Plato is thinking about the relationships among virtue, wrongdoing, and *epistēmē*. If Plato is keen to develop *epistēmē* along the lines proposed in the *Hippias Minor*, we should expect to find examinations of these topics elsewhere that further the search for a satisfying conception of *epistēmē*. Plato's response to these questions does not come all at once, but as we turn to the *Protagoras*, we see him take significant steps towards fleshing out his solution.

Chapter II: Virtue and *Epistēmē* in the *Protagoras*

Introduction

As we saw in the previous chapter, Socrates seeks a way to explain how SP1, that no one does wrong voluntarily, can hold. The way forward seems to lie in a further exploration of SP2, that virtue is knowledge. A form of SP2 appears briefly in the *Hippias Minor* regarding justice: justice is a power, *epistēmē*, or both. The *Hippias Minor* does not go into extensive detail about what SP2 might mean in this context. Nevertheless, as we have seen, it is very important: something about justice as an *epistēmē* is supposed to make voluntary injustice impossible. The *Hippias Minor* has set up the epistemological project, though it has not provided many concrete answers about what *epistēmē* is supposed to be. Our task, and Plato's, is now to start to find some of these answers.

With these motivations in mind, I turn to Plato's *Protagoras*. The *Protagoras* is a much longer work than the *Hippias Minor* and in many ways more complex. Nevertheless, the two texts share some similarities. Both were supposedly written earlier in Plato's career. Hippias features as a character in both dialogues. He is obviously much more prominent in the *Hippias Minor*, being Socrates' only real interlocutor. In the *Protagoras*, however, he is present at Socrates' and Protagoras' debate, and he interjects (or attempts to interject) at a number of places. Socrates addresses him directly a few times, including in a discussion of SP1. Hippias' presence in the *Protagoras* provides at least preliminary evidence that the dialogues are meant to be read together and that we should have Hippias' arguments in mind as we tackle those of the *Protagoras*.

A more significant connection is the treatment of virtue and knowledge in the two dialogues. The *Hippias Minor*, as I have argued, seeks to accept that no one does injustice

voluntarily, but to do so we need to understand what it means for virtue to be *epistēmē*. The relationship of virtue and *epistēmē* is front and center in the *Protagoras*, as Socrates and Protagoras debate throughout whether the virtues are teachable and how they relate to one another. They consider the unity of the virtues, the strength of *epistēmē*, the relationship between pleasure and the good, and the nature of wrongdoing, in which Socrates promotes what is commonly called the "denial of *akrasia*." All of these aspects are interesting and interrelated, and I believe that they will help us make progress towards understanding *epistēmē* and its relation to the virtues.

In this chapter, I will examine closely the arguments near the end of the *Protagoras* in which Socrates lays out his "denial of *akrasia*." I will argue that the *Protagoras* gives us two important features of *epistēmē* that are essential to understanding its role. The first lies at the heart of the denial of *akrasia*: someone with *epistēmē* of what is good and bad will never do something bad as a result of a more powerful motivation. *Epistēmē* is the strongest thing in human affairs; pleasure, fear, or some other force cannot overcome it. The second feature concerns the workings of this *epistēmē*: *epistēmē* involves measurement and functions by weighing up different aspects of our choices and allowing us to see them accurately. The measuring *epistēmē*, according to Socrates, will "save our lives" by guaranteeing that we act well and are not swayed by misleading appearances.

There is good reason to think that this kind of *epistēmē* would produce only good actions and block wrongdoing, so we have made significant progress on the questions raised in the *Hippias Minor*. The status of these gains is uncertain, however, because it seems that Socrates' entire argument hinges on first accepting that pleasure is the good. Much ink has been spilled on whether Socrates commits himself to some kind of hedonism in the *Protagoras*; many scholars

rightly hesitate to think that Socrates would embrace such a position. I join these scholars in their hesitations, but for different reasons: I argue that the role of the good life in the passage reveals there to be two parallel arguments about the good and wrongdoing. We are thus able to separate the argument's conclusions as they pertain to a good life in general, distinct from the good life specified as the pleasant life. This double argument allows us to save Socrates' proposal about the strength and function of *epistēmē*. Additionally, we are able to see how this *epistēmē* can ensure SP1, by showing what is really good for one's life and guiding one to the right actions as a result. We should therefore take the *Protagoras* as a serious step forward in developing a notion of *epistēmē* that can explain SP1 and SP2.

Overview of the Passage

The *Protagoras* chronicles a contest of sorts between Socrates and the eponymous sophist, who attempts to convince a potential pupil of the merits of studying with (and paying) him. Protagoras expounds on his ability to teach virtue, and Socrates asks in response whether we should consider the virtues to be a unity or different from each other. Protagoras insists on the latter, prompting Socrates to launch a number of arguments in favor of the unity of the virtues. He manages to get Protagoras to agree that wisdom, justice, piety, and temperance are quite similar, although courage, the latter maintains, is still different. Socrates' final aim in the dialogue is to convince Protagoras that courage is indeed to be identified with the other virtues—in particular, with wisdom. Socrates makes two arguments for this claim. The first, from 349e-350c, is unsuccessful; the second, from 359a-360e, manages to convince Protagoras or at least render him unable to deliver a final response.

Our passage of interest comes between these two arguments, from 351b-357e. Though we will not be examining the arguments about courage in any detail, I believe that the placement

of our passage between them is significant in flagging from the start how the Socratic paradoxes

relate to Plato's treatment of *epistēmē*. Plato, as we saw previously, is looking to find what about

justice as an *epistēmē* would guarantee good action. Socrates has just argued for the unity of the

virtues, inviting us to expand the inquiry about justice to the virtues in general. Accordingly, we

should be on the lookout for a conception of *epistēmē* that is associated with the virtues. Since

Socrates' first argument about courage was unsuccessful, whatever comes in our passage is a

necessary setup to show why one of the virtues is to be considered as wisdom.[55] We might hope,

as a result, to get a further exploration of SP2—that virtue is *epistēmē*—and make progress

towards our task of finding a notion of *epistēmē* that can guarantee SP1.

I have (unevenly) divided the passage in question into seven stages. In this section, I will

briefly go through the relevant points of each stage as neutrally as possible, and I will spend the

following two sections focusing on the picture of *epistēmē* that we are able to glean from the

passage.

> Stage 1: Initial discussion of pleasure and the good (351b3-e11)
> Stage 2: Initial proposal of the strength of *epistēmē* (352a1-e4)
> Stage 3: Argument that the pleasant is the good (352e5-355a5)
> Stage 4: First argument for the "denial of *akrasia*" (355a5-e4)
> Stage 5: Second argument for the "denial of *akrasia*" (355e5-356a5)
> Stage 6: Argument for an *epistēmē* of measurement (356a5-357c1)
> Stage 7: Summary of conclusions (357c1-e8)

Following this passage, Socrates also engages in a brief discussion with the sophists about SP1,

which we will examine in the final section of this chapter.

Socrates kicks off Stage 1 somewhat abruptly by proposing a seemingly innocuous claim:

some people live well and others do not. Protagoras denies that living painfully is living well,

and he agrees that someone who lived pleasantly seems to have lived well. When pushed further

---

[55] I will address the identification of *epistēmē* and wisdom in the next section.

by Socrates, however, as to whether living pleasantly is good, Protagoras hesitates and emphasizes that one should only take pleasure in noble things. Socrates tries to press the point, but Protagoras is uncomfortable answering, so Socrates moves to Stage 2, asking instead about *epistēmē*. He proposes that someone with *epistēmē* of good and bad would always do what is right. This view contrasts with that of the Many, soon to be an imagined interlocutor in the discussion, who think that other forces—anger, pleasure, pain, love, or fear—can rule someone with *epistēmē* instead and make that person do the wrong thing. Protagoras sides strongly with Socrates, and Socrates prompts him to help explain how the Many are mistaken about *epistēmē*.

In Stage 3, Protagoras and Socrates "team up" against the Many to explain why *epistēmē* cannot be overcome by pleasure. The Many contend that some pleasant things are bad and some painful things are good, and one can be "overcome" by the pleasantness of the bad thing or the painfulness of the good thing. The former situation occurs whenever one indulges in the pleasures of food, drink, and sex, although knowing they are bad. Socrates draws out from the Many that these things are bad only because they result in the painful long-term states of disease and poverty. On the other hand, things like exercise and surgery are good but painful, because they result in future pleasant things. As a result of this assessment, Socrates determines that pleasure is the only standard that the Many have for evaluating what is good. Because there is no other such standard, the pleasant is the good.

With this identification in place, Socrates shows that the possibility of *epistēmē* being overcome by pleasure is absurd. His first argument, in Stage 4, replaces "pleasure" with "good." On this understanding, one must do something bad, knowing that it is bad, while being overcome by the good. This, apparently, is absurd, to say that one knowingly did something bad as a result of being overcome by the good. The only solution is that the good in the worthless action must

50

be less good than the bad things that it brings about but have been mistakenly thought to be better. The same reasons apply at Stage 5, which deems it absurd that someone does painful things, knowing they are painful, because one is overcome by pleasant things. To explain how someone could choose the bad or painful thing, we must say that someone does not know that the choice is going to produce fewer good things and more bad things; someone does not have *epistēmē* in that situation.

Socrates next deals with an objection that the nearness of a pleasure makes it a different kind of thing, able to overcome someone with *epistēmē*. To meet this challenge, Socrates argues in Stage 6 that we can weigh the immediate pleasure against its long-term results, since pleasure and pain are the only relevant considerations. One can make the wrong decisions because of an immediate pleasure's appearances, which make it seem larger than it is. By contrast, we should try to acquire the art of measurement, which will reveal the true size of the immediate pleasure. Since the pleasant is the good, and pleasure and pain can differ only in relative amount, then the measuring art would save our lives by showing us how to make the best and most pleasant choices. Socrates can now respond fully to the Many in Stage 7: "You have agreed that those who err do so from a lack of *epistēmē* concerning the choice of pleasures and pains—and these are goods and bads—and not only *epistēmē*, but what you earlier have agreed is measurement" (357d3-7). In order to avoid doing wrong, we must embrace an *epistēmē* of measurement.

As might be apparent already from this brief overview, Socrates considers getting clearer on *epistēmē* an important part of his argument about the virtues. As a result, he goes into more detail here than he did in the *Hippias Minor*, and we should look for ways in which we can fill out our initial picture. When we examine this passage more closely, we will be able to pull out two notable claims about *epistēmē*:

Denial of *Akrasia* (DA):[56] It is absurd that a person could have knowledge that an action is bad but be overcome by some other motivation and do it anyway.

Epistēmē of Measurement (EM): An *epistēmē* of measurement will guarantee that we do good actions and avoid bad ones.

As we shall see, each of these claims makes significant progress towards answering the questions about *epistēmē* that were raised in the *Hippias Minor*. Let us examine and clarify each one.

Denial of *Akrasia*

Socrates begins his treatment of *epistēmē* not so much with an argument as with an assertion of its strength. He begins Stage 2 by inviting Protagoras to share Socrates' view of *epistēmē*, against the opinion of the Many: "What do you think about *epistēmē*? Does it seem to you as it does to the Many or not? This sort of thing seems right to the Many about *epistēmē*, that it is not strong nor commanding nor ruling" (352b1-4). The Many do not think *epistēmē* is powerful because "although *epistēmē* is often present in a person, *epistēmē* does not rule him but something else, sometimes anger, sometimes pleasure, sometimes pain, at times love, often fear" (352b5-c1). The Many think that *epistēmē* is weak "like a captive" (352c1). Even if a person has *epistēmē*, this *epistēmē* will not rule them—it will not guide their action—because of a number of more powerful forces. Pleasure and pain are two of the forces Socrates mentions, but they are not the only ones: anger, love, and fear can all drag around *epistēmē* as well. This is not, to be clear, a positive picture, even in the eyes of the Many. It would be well and good if

---

[56] The phrase "denial of *akrasia*" has been almost universally applied to this section of the *Protagoras*, based on the similarities found between the phenomenon described by Socrates as absurd and Aristotle's account of *akrasia*. It is worth noting that Plato himself does not use the term, and although Plato and Aristotle are clearly examining similar ideas, certain differences should be kept in mind. Most notably, the *Protagoras* does not describe the phenomenon in terms of parts of the soul or along the lines of reason versus appetite in the way that Aristotle does. Though other dialogues—most notably the *Republic*—do distinguish multiple parts of the soul, it is not clear that Plato was already making these distinctions in the *Protagoras*, nor that these distinctions will contradict Socrates' account, a point to which we will return in Chapter 3. Moss (2014a) views the *Protagoras* as containing the seeds of Plato's move to tripartition in the *Republic*, but see Wilburn (2014) for an argument against thinking that tripartition is meant to explain *akrasia*. For the sake of consistency, I will continue to refer to the argument as the denial of *akrasia*, though one should read it with these caveats in mind.

epistēmē were strong enough to motivate in the face of these other forces, but it happens not to be the case. Finally, *epistēmē* is not rarified in any sense, reserved for the few or the wise. Rather, it is "often present" in a person—presumably, the Many take themselves to have *epistēmē*. The phenomenon of being dragged around is one they attribute to themselves; they have knowledge but still do the wrong thing.

The view that Socrates presents, on the other hand, is quite different: "*epistēmē* is noble and can rule a person, and if someone knows good and bad things, they would not be overpowered by anything so as to do something other than what *epistēmē* bids, but *phronēsis* would be sufficient to save a person" (352c3-7).[57] Someone who knew what was good and bad would not be compelled to do anything besides what *epistēmē* commands. We thus get a preliminary glimpse of the content of the *epistēmē* in question, that it is of good and bad. Though this content is not incredibly specific, it does indicate that we are on the right track if we are searching for an *epistēmē* that aligns with SP1 and SP2. Socrates' note that *epistēmē* is noble contrasts to the Many's view by implying that *epistēmē* is not only extremely powerful but also not as prevalent as the Many might think. Socrates begins to put this conception of *epistēmē* on a pedestal; in contrast to some ordinary way of knowing, we are inclined to treat *epistēmē* as something loftier.

Protagoras was not yet willing to take a clear stand on pleasure, but Socrates hopes that he will side against the Many on *epistēmē*—as, indeed, he does. He responds to Socrates' proposal by asserting that "it would be shameful for me in particular [i.e. as a sophist] to say that wisdom and *epistēmē* are not the most powerful thing of all in human affairs" (352d1-3). One

---

[57] Plato never gives any systematic discussion of the differences between *epistēmē* and *phronēsis*. As we will see in Chapter 4, Aristotle considers at least Socrates to treat them as the same. *Phronēsis* does not play a role in Socrates' argument apart from this single mention, so I will consider it not to be significantly different from *epistēmē*. (Aristotle will take a much different view of the synonymy of the two terms, as we will see).

facet of Protagoras' response is particularly interesting for our purposes. Though wisdom (*sophia*) did not appear in Socrates' question, Protagoras identifies wisdom and *epistēmē* both as the strongest thing in human affairs. In fact, he identifies them with each other, referring to them collectively as the single most powerful thing (σοφίαν καὶ ἐπιστήμην μὴ οὐχὶ πάντων κράτιστον). The kind of *epistēmē* under consideration here is precisely the kind that relates the virtues. As Socrates has argued earlier in the *Protagoras*, wisdom is a virtue that is united to justice, temperance, and piety. If *epistēmē* and wisdom are identified with each other, and wisdom and the other virtues are to be identified with each other, then *epistēmē* here is the one in question when we say "virtue is knowledge" and the one at issue when we saw SP2 in the *Hippias Minor*.[58]

One might object that it is Protagoras, not Socrates, who is making this identification. The text gives us good reason, however, to think that Socrates accepts this conception of wisdom-*epistēmē* as well. First, it is one of the few places where Socrates endorses one of Protagoras' remarks as true (Καλῶς γε, ἔφην ἐγώ, σὺ λέγων καὶ ἀληθῆ, 352d4).[59] Additionally, Socrates treats *epistēmē* and wisdom as the same at other points in the *Protagoras*. At 330b4, for example, Socrates uses *epistēmē* instead of wisdom when he lists the five virtues under consideration.[60] Near the end of the dialogue, he treats them both as having the same opposite,

---

[58] In making this identification, I am glossing over several issues concerning how the virtues are a unity. Vlastos (1972) distinguishes three potential types of unity that appear to be endorsed by Socrates: (1) "Unity," that is, the virtues are all identical and synonymous, (2) "Similarity," that is, the virtues resemble each other in all of their aspects, and (3) "Biconditionality," that is, someone exhibiting one of the virtues necessarily exhibits them all. Vlastos argues that while (3) is the most straightforward to understand, all three theses can be understood to express essentially the same doctrine, which can be understood by considering "Pauline predication" as a more intuitive way of grasping confusing identity claims. Penner (1973) presses instead for (1), followed by Hartman (1984). In particular, all of the virtues are essentially knowledge of good and bad, as applied to different situations for the different virtues. Given the position of *epistēmē*, I find myself more sympathetic to Penner's reading, but as I am more concerned with an account of *epistēmē* than of the virtues, I do not need to take a strong stand here. Whatever we may say about these issues, my main point is that we have good evidence that SP2 is at play.
[59] I do not see any indication that Socrates' truth endorsement here is ironic, *pace* Goldberg (1983).
[60] Duncan (1978) also picks up on this point and finds it significant.

ἀμαθία: *epistēmē* is contrasted with ἀμαθία at 357e1 and wisdom with ἀμαθία at 358c2-3. We should thus take SP2 to be operative here; *epistēmē*, as wisdom, is the "knowledge" of "virtue is knowledge."

Of course, the Many as Socrates imagines them would not straightaway concede that *epistēmē* works in this way. They must be convinced that *epistēmē* cannot be overcome by pleasure in particular, giving Socrates the opportunity to elaborate on the strength of *epistēmē*. Socrates gets the Many to agree that cases of acting through being overcome by pleasure are bad because they bring pain and diminish pleasure in the long run. Accordingly, knowing that an action is bad involves knowing that it will produce less pleasure and more pain, since that is the only standard for good and bad. Socrates exhorts the Many not to "use many names at the same time, 'pleasant' and 'distressing' and 'good' and 'bad'" (355b4-5). As we saw, the Many would have to think first that a smaller amount of good compelled them to forgo a larger good, and second that a smaller pleasure compelled them to forgo a larger pleasure.[61] As such, they would strictly be acquiring something worse, knowing it is worse, for no other reason than that its smaller good overcame the larger one. This, Socrates argues, is absurd and could only happen to someone without such knowledge.

Socrates has thus taken himself to have proved DA: it is absurd that a person could know that an action is bad but be overcome and do it anyway. DA blocks a prominent type of voluntary wrongdoing—when one knows what to do but is led by another motivation to do

---

[61] There has been a lot of discussion regarding what exactly the absurdity here amounts to—whether there is a contradiction and what it is, or whether it is just seen to be absurd. See Gallop (1964), Vlastos (1969), Dyson (1976), Nussbaum (1986), Weiss (1989), McCoy (1998), Wolfsdorf (2006), and Callard (2014). When we look more closely at how the measuring *epistēmē* deprives appearances of their power, we will see the absurdity to lie in something like this: it is absurd that something that, thanks to the measuring *epistēmē*, has no motivational power will motivate us to do something (bad). Whether this line of thought is spelled out as a logical contradiction or mere absurdity is not really worth worrying about; the point for now is that Socrates feels confident there *is* an absurdity of some kind and so DA must be right.

wrong. We can see how *epistēmē* would connect with virtue: someone with *epistēmē* of good and bad would not fall victim to the many temptations that cause cowardice, intemperance, and the like. We have also gotten at least some brief notes on what this *epistēmē* might involve; Socrates states at the outset in Stage 2 that it involves good and bad, and we see this confirmed in Stage 4.

We might wonder, however, whether and how this type of *epistēmē* would work. In particular, Socrates may be incorrect that these things can outweigh each other only in relative excess or deficiency. If there is some other way to understand the choice to perform a pleasant but worthless action, then the Many might be correct in their analysis, and *epistēmē* would not be the strongest thing in human affairs.

Epistēmē of Measurement

An objection is promptly raised along these lines. Socrates had posited that the good and bad, or the pleasant and painful, can be compared to each other only in terms of relative excess and deficiency. This claim ignores the possibility that "the immediate pleasure differs a lot from both the pleasure and the pain at a later time" (356a5-7). One may accept that indulging in a pleasant thing will bring less overall pleasure than abstaining from it. That does not mean that we can act on this knowledge when we are immersed in the situation itself. Surely the glass of wine sitting right in front of you, say, is just a different kind of thing from the thought of a hangover the next morning.[62] Even though the overall pleasure from the wine is less, the kind of pleasure that it now presents causes one to make the wrong decision. If the nearness of a pleasure is overwhelming in a way that *epistēmē* cannot counteract, then DA is false and

---

[62] One need not be held up by concerns that in the wine example one's faculties have been severely impaired by previous glasses of wine. Suppose the person in question is just, unfortunately, very susceptible to hangovers—it takes only two glasses to bring about long-term pain, and likewise at least two to feel the impairing effects of alcohol. In the overcoming situation, she is deciding whether to drink the second glass.

epistēmē cannot block this kind of wrongdoing. Though Socrates mentions only the nearness of a pleasure as a potential threat, we might think there are many other ways in which pleasures could turn out to be very different. Someone who loves risk might find the riskiness of a course of action—which surely involves the fact that bad consequences may arise—as an irresistible motivation towards a bad action. The novelty of a guaranteed painful experience might also render it enticing. These pleasures loom large: isn't it clear that they can have more power in the heat of the moment than *epistēmē*?

Socrates' response is clear: they cannot. Earlier, he drew out from the Many that pleasure is the only standard in determining what is good. If that is the case, then immediate and long-term pleasures "surely do not differ in some other way than pleasure and pain [. . .] for it is not possible to differ in any other way" (356a7-8). We have one category, pleasure, and Socrates, against the Many, wants to apply that category to both immediate and long-term pleasures, as well as to riskiness and any other feature that could serve to distinguish them. If Socrates denies that immediate pleasure is of a different kind, then he faces two challenges. First, why would someone *think* that immediate pleasure is so different? Second, how can one get past this kind of thinking?

To start to meet both of these challenges, Socrates introduces the idea of weighing. Someone can determine what is most pleasant "just as a person skilled at weighing, putting together the pleasant things and putting together the painful ones, setting both the near and the far on the balance, you say which is more" (356a8-b3). Socrates proceeds to lay out three weighings: (1) the pleasant against the pleasant, (2) the painful against the painful, and (3) the pleasant against the painful. These three weighings will allow us to evaluate any choice, no

matter how many different kinds of pleasure may be at hand.[63] Instead of considering just a

single glass of wine, imagine that one is at a wine shop, choosing whether or not to purchase an

entire Jeroboam of wine—six regular bottles' worth (or four, if champagne is the pleasure of

choice). The associated pleasures and pains would certainly include the pleasures of drinking

each glass and the pains of future hangovers. But these do not exhaust the considerations: there

is an aesthetic pleasure or pleasure of novelty in buying such a comically large bottle, of showing

it off to and sharing with friends, a pain of spending quite a bit of money, of carrying such an

unwieldy item home. Socrates' point is that all of these, whatever labels we might give them,

can be measured on the scale of pleasure and pain.

Suppose we were to attempt to decide whether to buy the bottle. Though Socrates does

not go into detail about the interaction of these three weighings, we could imagine them

proceeding in steps. First, we weigh up the pleasures associated with performing the action

versus not performing the action. It does not matter that buying the bottle produces some

pleasure of novelty and some pleasure of drinking. For each of these, we can ask, "how much

pleasure does it bring?" We might determine that buying produces quite a lot of pleasure and

refraining produces none or very little. With regard to pleasure alone, buying beats out

refraining. In the second weighing, we might find a number of pains associated with buying, and

just a few with refraining. The choice, then, hinges on the third weighing, whether the net

pleasures of buying exceed the net pains. If we determine that they do, then we should buy; if

they do not, we should refrain.

---

[63] Richardson (1990) takes the presence of three weighings to imply that the pleasures and pains are not
commensurable. I would disagree, since Socrates flags so explicitly that there is no way the pleasures and pains
differ than the way in which they can be weighed against each other. The three weighings appear, as a result, as
steps in the same process of weighing; one could think of the process as one extended weighing.

58

Socrates has taken steps towards answering the second challenge: we can overcome the appeal of an immediate pleasure or pain by weighing up the different pleasures and pains speaking for or against that thing. This weighing is performed without consideration for the nearness or variety of the pleasures and pains; all that matters is the relative size of the pleasures and pains. Socrates has not, however, met the first challenge. He has asserted that the nearness of pleasure is not important to weighing, but we will not be able to do this weighing at all if we are still "overcome" by something about the near pleasure. Suppose we conclude that we should buy the Jeroboam, but we are overwhelmed by how heavy it would be to carry and are driven to avoid that pain. We can admit that we *shouldn't* have been overcome by the immediate pain of carrying it, but if we are rendered unable to weigh it correctly when it is necessary to do so, then we have not really come up with a good solution to this kind of error.

At this point, Socrates turns to analogy. He invites the Many to consider whether "the same-sized things appear to the eye bigger from close up, and smaller from farther away" (356c5-6). The answer is clearly yes: a tower half a mile away appears smaller than an identical tower right next to us. Socrates attributes this phenomenon to "the power of appearance," which "causes us to waver and makes us often turn about the same things back and forth and have regret both in our actions and in the choices of big and small" (356d4-7). If we were to judge solely by appearances, we would think that the close tower is larger. If there were some stakes involved in correctly determining which tower is larger, we would regret following our appearances and saying that the close tower is larger. The towers' appearance, furthermore, does not stay fixed: if I walked over to the other tower, that tower would now appear larger. The "power of appearance" is not a good guide either for the true size of the towers or for any actions that depend on knowing their relative sizes.

Socrates' description of the power of appearance recalls his self-description in the *Hippias Minor*: wavering back and forth between two things and unable to determine how to move forward. Socrates wavered because he did not have knowledge: opposite things seemed true to him and prevented him from drawing conclusions about justice and wrongdoing. In the case of the *Protagoras*, we also see an example of opposing things seeming true. In one context Tower A appears larger than Tower B, and in another context Tower B appears larger than Tower A. As in the *Hippias Minor*, we should be able to end our wavering by getting to the heart of the matter. We can determine which of the opposite appearances is false or how, somehow, they do not contradict one another.

Socrates posits something that, in contrast to the power of appearance, would allow us to get to the heart of the matter: the "art of measurement" (356d4). The art of measurement is something that "will make this appearance powerless, and making clear the truth it would bring our soul peace, remaining in the truth, and it would save our life" (356d8-e2). In the case of the towers, we could use some sort of measuring stick to determine both towers' true heights, and we could then compare these heights instead of their relative apparent sizes. Socrates posited that the towers were the same size, so the art of measurement will show that neither of our appearances told the truth about this matter. Socrates considers a case where finding out the truth about the sizes is more important than just a question of getting things right. In a situation where "acting well depended on doing and taking things of great length, and avoiding and not doing small ones," the art of measurement would be "our salvation in life" (356d1-3). Socrates does not, to be clear, think that our wellbeing depends upon choosing the largest tower. The lesson is more general: if our wellbeing depends upon choosing a certain kind of thing and avoiding its opposite, appearances may confuse us and cause us to choose the things we are

supposed to avoid. With an art of measurement, however, we can overcome these appearances and choose what will actually contribute to our wellbeing.

We can apply this line of thinking to pleasure and pain. Just as a nearby tower may look larger through the power of appearance, appearances may cause a near-term pleasure or pain to appear disproportionally large. But this illusion should worry us no more than the fact that a tower appears larger to us when we are closer to it. We can resist being "overcome" by pleasure if we possess a measuring art, which will tell us the true relative sizes of near and long-term pleasures and pains. This measuring art is important not just for finding out which actions are more and less pleasant. Importantly, "salvation in life has appeared to be in the correct choice of pleasure and pains [. . .] so first it appears to be measurement" (357a5-b2). An art that would determine what is, rather than what merely appears, the most pleasant will result in a good life.

Socrates notes that this measuring art would be an *epistēmē* (357b4). He declines, however, to specify further, saying only that "whatever this art and *epistēmē* is, let us investigate some other time; but it is enough that it is *epistēmē*" (357b5-7).[64] Since Socrates has shown that this *epistēmē* of measurement would be our "salvation in life," he can now respond fully to the Many's objection:

> You have agreed that those who err do so from a lack of *epistēmē*
> concerning the choice of pleasures and pains—and these are goods
> and bads—and not only *epistēmē*, but what you earlier have agreed
> is measurement; and the action made in error without *epistēmē*
> even you yourselves know, I suppose, that it is done from
> ignorance. (357d3-e1)

---

[64] Socrates uses both *technē* and *epistēmē* initially to describe the measurement, and then he continues to talk of it only as an *epistēmē*. I refer to it as the measuring *epistēmē* both for shorthand and to emphasize its importance; as a *technē*, it is not on par with other *technai* but is a skill of a different sort, particularly because it is divorced from wrongdoing. Insofar as it bears similarities to measuring length or number size, it has a "technical" component, but it is primarily an *epistēmē*. I also find it significant that the measuring art is called a *technē* when Socrates introduces it as a measurement of pleasure and pain, but it is only an *epistēmē* in Socrates' primary argument for the measuring *epistēmē* in general. I believe that this echoes my point about the specialization of *epistēmē* in the *Hippias Minor* near the end of the previous chapter.

Those who supposedly have *epistēmē* but are overcome by pleasure turn out not to have *epistēmē* at all. If they had the *epistēmē* of measurement, they would be able to weigh up the near and long-term pleasures and pains; since they have not been able to do so, they must be swayed by appearances and not have *epistēmē*. It is not surprising that the Many and Socrates differed on whether people who are overcome have *epistēmē*; recall that the Many considered *epistēmē* to be widespread while Socrates considers it to be "noble." By laying out what *epistēmē* might involve, Socrates shows the Many that it is not something found in themselves but rather a prized but rare thing that would indeed save their lives.

One point worth emphasizing is that we can see clearly how a measuring *epistēmē* of this kind has practical value. Whatever *epistēmē* itself may involve, we would be able to use it to make better choices and perform better actions. *Epistēmē* measures things that are relevant to these actions and choices—even the Many can see its value. The fact that *epistēmē* is robust, "noble," and not widely possessed does not change its worth. As we further explore the workings of *epistēmē* and flesh it out more in the next chapter, it will be important to keep this point in mind.

Let us take stock of what we have learned about *epistēmē* in the *Protagoras*. Socrates has provided us with two important features that show how *epistēmē* can guarantee good action. First, what I have called DA states that it is absurd that a person could have knowledge that an action is bad but be overcome by some other motivation and do it anyway. Second, what I have called EM states that an *epistēmē* of measurement will guarantee that we do good actions and avoid bad ones. Put together, these claims present us with a robust notion of *epistēmē* that not only tells us what to do but blocks other forces that attempt to guide us otherwise.

If we are able to take DA and EM seriously, then Plato has presented us with a way that someone with *epistēmē* would avoid a serious and prevalent type of voluntary wrongdoing.[65] Someone who had *epistēmē* of justice would know the just thing to do, and with this *epistēmē* in place she would not be drawn by pleasure to commit injustice instead. Accordingly, this passage of the *Protagoras* can answer the questions raised in the *Hippias Minor*; we can see what about SP2 explains SP1. Before we can fully accept this account, however, we must deal with a potentially fatal problem with accepting Socrates' arguments at face value.

The Pleasant is the Good: Hedonism in the *Protagoras* and Recent Interpretations

The problem is that the arguments for DA and EM appear to depend upon first accepting a hedonist thesis that the pleasant is the good. As is likely already clear from our examination above, pleasure is a constant presence in our passage of interest. Our brief overview identified two stages of the argument, Stage 1 and Stage 3, that focus primarily on identifying pleasure and the good. These stages appear to be necessary setup to get the arguments for DA and EM off the ground.

Take DA first. While Socrates' initial description of *epistēmē* in Stage 2 does not hinge on pleasure, Stages 4 and 5 are set up as a proof for Stage 2, to respond to an objection from the Many. If they depend on hedonism, then it seems that accepting the account in Stage 2 will as well. In setting up the first argument for DA, Socrates notes that "if you cannot say that anything is good or bad other than whatever results in pleasure and pain, listen to what's next" (355a3-5). His refutation of the Many thus hinges on having established a commitment to hedonism in Stage 3; he claims, furthermore, that "all the demonstrations depend on this" (354e7-8). The

---

[65] We will be able to reach the conclusion that this is the *only* relevant candidate for voluntary wrongdoing, but I will address this point at the end of the chapter.

absurdity of *epistēmē* being overcome becomes clear precisely when "we stop using many names at the same time, 'pleasant' and 'distressing' and 'good' and 'bad,'" since "these have appeared to be two," i.e. pleasant = good and distressing = bad (355b4-6). Without being able to substitute "good" for "pleasant" at Stage 4 and "painful" for "bad" at Stage 5, it is not clear how the two arguments for DA would succeed.

Likewise, the argument for EM at Stage 6 is precipitated by a need to elaborate on how *epistēmē* and pleasure interact. The ability to weigh near and far things purports to explain how immediate and long-term pleasures would be measured, as we saw in the example of the Jeroboam. Socrates notes that the function of the measuring *epistēmē* itself "has appeared to be in the correct choice of pleasure and pains" (357a5-7). The measuring *epistēmē* that we seek must be an *epistēmē* that tells us which pleasures and pains are greater and lesser. Socrates sums up in Stage 7 by once more reminding us that pleasure and pain are the good and bad (357d5-6), further cementing their connection to DA and EM.

Both DA and EM, then, apparently rely on first accepting some hedonist thesis. This has led a number of interpreters to conclude that Socrates himself endorses hedonism in the *Protagoras*.[66] The biggest advantage of these interpretations is that we can take the text, and accordingly Socrates' conclusions about *epistēmē*, at face value. The biggest disadvantage is that Plato's other dialogues hold little evidence of supporting hedonism and much more evidence against it.[67] In order to reconcile the *Protagoras* with Plato's thought as a whole, one must either

---

[66] See Adam and Adam (1905), Hackforth (1928), Gosling and Taylor (1982), Nussbaum (1986), Rudebusch (1999), Taylor (1991), Irwin (1995), Rowe (2003), Moss (2014a). Penner (1997) does not think the point to be of much importance but also considers Socrates to endorse hedonism.

[67] *Phaedo* 68e-69c contains a fairly direct attack on the kind of hedonism that Socrates seems to promote in the *Protagoras*: keeping away from some pleasures to get more pleasure is described as licentiousness, and Socrates suggests instead a purging away of pleasures in favor of wisdom. See Weiss (1989) for the disconnect between the *Protagoras* and the *Phaedo* (and, to some extent, the *Apology*). The *Gorgias*, *Republic*, and *Philebus* also offer critiques on thinking of pleasure as the good.

portray Socrates' position as a brief flirtation by Plato or representative of a kind of hedonism that's not, in fact, incompatible with his attacks elsewhere.[68] The first strategy is a major blow against thinking DA and EM to be proposals that Plato considered serious or stable parts of his thought. The second strategy loses the advantage of being a straightforward reading of the texts; we will examine what kind of hedonism is in play more closely later on, but it is already relatively clear that drawing out an "enlightened" or "sophisticated" hedonism from the *Protagoras* will require a bit more work than the letter of the text affords us.

These worries are primary motivators for a second camp of scholars, who argue that Socrates is not committed to hedonism in the *Protagoras*. Rather, his main goal, whatever his commitments, is to show that Protagoras and/or the Many are hedonists or persuade them of hedonism for some larger purpose.[69] These views range from considering Socrates to be an active anti-hedonist to simply not being committed either way; what they tend to have in common is a focus on the *ad hominem* nature of the text. Socrates is not so much trying to put forward a theory of his own as he is examining the views of others, so we should not take his proposal of hedonism to be serious.

One upshot of this line of thinking is that if we do not think that Socrates is endorsing the claims about pleasure, then it would be strange to think him serious about the conclusions that follow or at least to have provided an argument for them. This puts the status of DA and EM in serious jeopardy, and we must confront the worry that the entire last section of the *Protagoras* is

---

[68] Most of the above interpreters hold some form of the first position. Moss (2014a) most clearly holds the second, along with Penner (1997).

[69] See Grube (1933), Sullivan (1961), Sesonske (1963), Santas (1966), Vlastos (1969), Dyson (1976), Duncan (1978), Zeyl (1980), Frede (1986), Weiss (1990), Hemmenway (1996), Kahn (1996 and 2003), Annas (1999), Russell (2000, 2005), Wolfsdorf (2006), Dimas (2008), Vasiliou (2008), Shaw (2015), Callard (2016), Wilburn (2016), Mann and de Harven (2018), and Pasnau (2021).

little more than an exercise in sophistry. Some scholars may well be happy to bite that bullet,[70] but we already saw that this move left much to be desired in the *Hippias Minor*. We left the *Hippias Minor* with a task of finding out more about *epistēmē* in the hopes of understanding the Socratic paradoxes. Socrates is now giving us some of these answers in conjunction with things like SP2 that he does seem inclined to endorse, not only on our reading but on a general understanding of Plato. A purely *ad hominem* reading would render this passage a mishmash of serious and unserious claims and leave Plato's overall purpose a bit too mysterious. If we can find some way to make sense of what Socrates is up to here, we might hope to bring more clarity to his attitude in and our takeaways from the passage.

The thorny debate about hedonism in the *Protagoras* has raged on for some time, with little hope of a consensus.[71] Though I do not intend to answer all of the relevant questions surrounding these issues, my hope is to provide a reading that takes a new path forward: I wish to show that DA and EM are put forward as serious conclusions, ones that furthermore do not depend on hedonism or any commitments that are ascribed only to Socrates' interlocutors. While a couple of attempts have been made in this general vein, none have been satisfying, especially in making sense of the structure of the passage as a whole.[72] Additionally, scholars have tended to focus on DA at the expense of EM. As my primary interest is in the conclusions

---

[70] Weiss would strike me as an especially eager contender, particularly given her (2006) view of the Socratic paradoxes that we saw in the Introduction.

[71] The most extreme pessimism might come from Taylor (2003), who modified his original pro-hedonist position to suggest instead that Plato *intentionally* wrote in such a way as to make the question of Socrates' hedonism unanswerable.

[72] I take Morris (2006) to argue for something close to this line of thought: he separates, as I will, two versions of DA in the text to argue that only one relies on hedonism. However, he has less to say about the purpose of the two arguments, does not take them to apply to EM, and seems rather skeptical about the success of the passage as a whole, compared to what he takes to be the *Republic*'s view. Clark (2012) argues that a principle of psychological eudaimonism, rather than psychological hedonism, is at play in the argument; I believe that this point picks out something very crucial about the passage, as we will see. Clark does not, however, spend much time on the upshot of this view and what it might mean for the conclusions that Socrates draws.

that can be drawn about *epistēmē*, I intend to show how both DA and EM are afforded serious

consideration without a commitment to hedonism.

Pleasure and the Good Life

In order to understand the role of DA and EM in the *Protagoras*, I propose to focus on a

crucial but insufficiently recognized aspect of the passage in question: the role of the good life

in framing and propelling Socrates' arguments.[73] Socrates' aim in introducing hedonism is to

bring to light that his interlocutors conceive of the good life as the pleasant life. We will see

what kind of life this turns out to be: one comprised of conventional, but decidedly un-Socratic,

goods. Socrates himself does not endorse this conception of a good life as the correct one.

Rather, he focuses on this specific good life because it is the one that his interlocutors embrace,

while setting up his arguments so that they apply to the good life in general, whatever that may

turn out to be.[74] This strategy allows Socrates to set the stage for DA and EM in a way that does

not rely on hedonism.

Our first indication of the importance of the good life is that it kicks off the entire

discussion. Stage 1 starts in the following way: "Do you say, Protagoras, that some people live

well, and some badly?" (351b3-4). In context, Socrates' question is a jarring change of subject.

Protagoras has just finished a small speech about his conception of courage, arguing that it

---

[73] Scholars have, on the whole, not taken Socrates' points about the good life to be important in the passage. Vlastos (1969) holds "all men desire welfare" as one of Socrates' commitments, but he takes this as a general Socratic assumption and does not dwell on what sorts of "welfare" might be under consideration. Clark (2012), as mentioned in the note above, is the first to give due consideration to Socrates' treatment of the good life in the *Protagoras*. However, his treatment is general and avoids fleshing out how exactly the good life fits into the passage and bolsters Socrates' conclusions. He worries that Socrates is not explicit enough about its role, noting a bit timidly that "[p]erhaps Plato is insinuating that the good life somehow matters to the argument" (253). I intend to continue his project by pursuing this line of thought more deeply.

[74] In this way, my position is "anti-hedonist" and in line with scholars who consider hedonism a commitment of the Many and Protagoras but not Socrates. I would not, however, consider my position to be *ad hominem* in any strong sense. Though Socrates is certainly using his arguments to set up an attack on Protagoras' conception of virtue, his point is not only to consider his interlocutors' views but also to draw conclusions of his own.

involves nature and nurture rather than wisdom. Without much indication of moving on to a new line of thought, or even acknowledging Protagoras' speech at all, Socrates proceeds to ask whether Protagoras considers some lives to be good ones.[75] Let us examine what the function of such a beginning might be.

The framing of Socrates' question invites us to consider two points. The first is that the jarring change serves to emphasize the question itself: the abrupt shift indicates to us that Socrates' question is not just filler but something meant to be considered in its own right. At the same time, we are invited to try to see how the question of living well *does* relate to the previous discussion: in what ways is it not a shift but a related point?[76] Here could be one way to understand the connection: in order to deliver a verdict on what courage is and whether it is wisdom, we must first grasp living well in general. Both interlocutors might think that living well involves living courageously, but that is no help if the former is undefined and the latter is under dispute.[77] If Socrates wishes to show Protagoras that the latter's views on virtues are mistaken, he could employ two strategies. First, he could show that Protagoras' conception of living well is faulty. Second, he could show that, even for Protagoras' conception of the good life, courage is to be understood the way Socrates proposes.[78] This strategy leaves open whether Protagoras' conception of living well is a good one and whether it is shared by Socrates. While

---

[75] As Weiss (1990) points out, Socrates' use of δέ at 351b3 indicates that it is not a completely new start, even though the subject matter seems to be unrelated. I think that this fits well with my two points below: even though the question is jarring, it still invites us to consider how it might relate to what came before it.
[76] Weiss (1985a) also sees the framing as trying to emphasize a connection.
[77] See also Goldberg (1983), though he makes the point in relation to knowing the good itself rather than the good life.
[78] Zeyl (1980), though directly concerned with DA, thinks Socrates employs this strategy: "If I can defend a view of mine by either of two arguments, only one of which I accept as sound but whose premises may be hard to defend, while I regard the other as valid, depending on premises some of which I do not accept, and I realize that the latter argument would have greater cogency against someone who does accept these premises than the former, I may have excellent reason to use the latter argument to defend my view. This, I believe, is exactly the position of Socrates in the Protagoras" (260). Where I depart from Zeyl regarding is that I take Socrates actually to employ both arguments (of a sort, as we will see), rather than just one that Protagoras would see as sound.

neither party has yet given content or further specified what "living well" means, Socrates

clearly deems it important to set out straightaway that everything under consideration will relate

to this idea.

Socrates prods Protagoras precisely on what his good life involves. In particular, he asks

whether Protagoras thinks that living pleasantly counts as living well. Protagoras agrees that

living in pain is living badly and living pleasantly is living well.[79] But he hesitates to say that

everything pleasant is good, qualifying that this is the case "only if one lives taking pleasure in

noble things" (351c1-2). Socrates attributes a similar view to the Many: "Surely you too don't

call some pleasant things bad and distressing ones good, as the Many do? I mean, in the respect

in which things are pleasant, are they not good, if nothing else results from them?" (351c2-5).[80]

For Socrates' interlocutors, there appears to be a gap between living pleasantly and pleasant

things themselves, since Protagoras and the Many are inclined to designate some of the latter as

bad. Protagoras picks up on this gap, telling Socrates that "it seems to me safer not only to give

an answer for the present, but rather for my life as a whole" (351d2-4). He reveals himself to

resist a strict hedonist thesis not because there is necessarily something inherently bad about

certain pleasures but because they are bad when considered in the context of a life. We thus

have to reconcile two things about Protagoras' view: first, that living pleasantly is living well,

and second, that some pleasant things are bad. One solution could be that Protagoras does not

---

[79] Zeyl (1980), followed by Weiss (1990), Hemmenway (1996) and Shaw (2015), takes Protagoras' initial assent to show that he is a hedonist, although he may be ashamed to admit it upon further initial prodding. The argument above is sufficient because if one considers living pleasantly as living well and living painfully as living badly, then nothing but pleasure and pain can be the deciding factor (a point Socrates clarifies later in the argument). I am inclined to agree with this reading: Protagoras sees the pleasant life as the good life, then worries about the upshot of this conception, which later ceases to bother him. However, whether Protagoras initially accepts the pleasant life as the good life or is only comfortable with this identification later on does not matter much for our purposes. Wolz (1967), Russell (2000), and Dimas (2008) argue against concluding that Protagoras already accepts hedonism here.
[80] Throughout the argument, Socrates uses both ἀνιαρός and λυπηρός to refer to painful things. To track the terms, I use forms of "distressing" for the former and "painful" for the latter, although I do not think that much hinges on the difference.

actually endorse the former claim. If he does not have a conception of living well, however,

further discussion of virtue becomes much more difficult. Socrates proceeds to take up the latter

point and examine further, through his imagined dialogue with the Many, what it means for some

pleasant things to be bad.

Socrates begins Stage 3 by laying out the Many's claim against the strength of *epistēmē*:

"you are often overpowered by food and drink and sex because they are pleasant, knowing that

they are worthless" (353c6-7). The characterization of a pleasant thing as worthless prompts an

additional clarification from Socrates: "why do you say these things are worthless? Is it because

[1] they provide pleasure itself immediately and each of them is pleasant, or because [2] later on

they make sickness and poverty and procure many other things like these? Or if any of these

pleasures procures nothing later on, but only causes enjoyment, would it nevertheless be bad,

because [3] it causes enjoyment in whatever way and for whatever reason?" (353c9-d6).[81] The

Many answer that the reason [2] is the source of badness: pleasures can be bad if they produce

things like sickness and poverty. Socrates asks whether "they cause distress when they cause

sickness, and they cause distress when they cause poverty," and the Many answer that pain

would indeed be a necessary part of these states (353e3-4). Socrates prompts that "it seems to

you, as Protagoras and I are saying, that these things are bad on account of nothing other than

that they result in distress and deprive one of other pleasant things" (353e5-354a1). The

pleasures of food, drink, and sex are bad not in themselves but only whenever they contribute to

or produce disease and poverty. The distinction between short-term and long-term conditions

---

[81] I follow Bizoń and Sokołowski (2012) in treating μαθόντα as adverbial and refer the reader to them for more
thorough discussion. They make the larger point that Socrates here provides an alternative to the way in which
pleasure could be bad, though it is not accepted by the Many.

and the specification of future bad states to avoid make one thing clear: the Many have a certain conception of what makes for a bad life.

The Many state that they avoid some pleasures of food, drink and sex because they hope to avoid disease and poverty. They do not give any further states that disease and poverty would bring about, merely noting that disease and poverty themselves are painful. These states "bottom out" regarding what sort of life to avoid—a live of disease and poverty is the ultimate bad life. Now it is not unreasonable to think that the pleasures of food, drink, and sex can cause one to end up in disease or poverty. Though an occasional indulgence may not make much of a difference, one who pursues them "often," as the Many claim, may indeed bring upon oneself a much lower standard of health and wealth in the long run. It is also reasonable to see both states as painful, or at least as accompanied by pain. But is living in disease or poverty living badly? Though the Many are happy to take this step, Socrates does not join them in his endorsement, proposing only that there is no other reason that these states would be bad apart from pain.[82] There is no indication that Socrates would consider disease and poverty to be emblematic of a bad life.[83] The life that the Many hope to avoid, however, has turned out to consist of these things, which they consider to be bad just insofar as they are painful. The Many's stance

---

[82] Taylor (1976) and Irwin (1995) take this endorsement as a sign that Socrates is committed to hedonism, but he says no such thing. To claim that pain would be the only reason to call something bad does not mean either that the thing *is* bad inasmuch as it is painful or that that thing is part of a bad life. Furthermore, as others have noted, this "endorsement" can be read as Socrates confirming the Many's view, not expressing his own. See Zeyl (1980), 255. One of the only places in this passage where Socrates explicitly endorses something as true is, as noted before, Protagoras' agreement that *epistēmē* is the strongest thing.

[83] There seems to be, in fact, more evidence that he did not, if his own life is to be any indication. Though Socrates did not appear to live in extreme poverty, he also was not very wealthy by most accounts and made much of the contrast between himself and sophists like Protagoras who sold their craft. As the discussion terminates in what is certainly a pointed exhortation to pay the sophists, the inclusion of poverty here is especially worth noting. On Socrates' poverty and its philosophical legacy, see Schaps (2003). Whether Socrates suffered from disease is less clear, though he was certainly considered ugly (see e.g. *Theaetetus* 143d); for a(n over)diagnosis of Socrates' physical afflictions, see Papapetrou (2015).

indicates that they have a conception not only of a bad life but also of a life they do want to pursue—the good life.

The Many's conception of a good life comes into focus when Socrates argues from the other direction, concerning good but painful things like surgery and military training. Such things are good not because they are painful but "because later on health comes about from them and good conditions of bodies and the safety of cities and rule over others and wealth" (354b2-5). The Many choose some painful things in order to have a life of health, wealth, and political power.[84] Once more, the Many go further than just saying that these things are good: these states "top out," indicating that they comprise living well. It does not take much familiarity with Plato's Socrates to see that this list would be a very strange one for him to uphold as a good life. For one, there is no mention of the condition of the soul, only that of the body, quite contrary to the usual Socratic spirit.[85] Power over others seems much more to be a Thrasymachean good than a Platonic one. Socrates points out that these things would be considered good "because they result in pleasant things and relief and preventions of painful ones" (354b6-7). As with the

---

[84] The Many's collection of goods in their good life can be used, I would argue, to block interpretations of an "enlightened" hedonism in the *Protagoras* (see Gosling and Taylor (1982), Penner (1997), and Moss (2014)). The thesis that the pleasant is the good is understood on these terms: what is under consideration is a life of conventional (and un-Socratic) goods, and there is no indication of other goods that would themselves be good only insofar as they are pleasant. Shaw (2015) forcefully argues this point against a hedonist interpretation. A few other scholars make note of the collection of goods but do not attach to it much significance. Vlastos (1969) makes note of the fact that Socrates is more permissive than usual in terms of what count as goods, but he says only that Socrates hopes to meet the Many on their own terms (74-5). Sullivan (1961) mentions the absence of typical Socratic goods but gives no upshot thereof. Weiss (1989) points out the collection of goods primarily to draw a contrast with the more philosopher-friendly stance in the *Phaedo*. Callard (2016) takes the goods as pretty typical for Greeks (but does not compare them to Socratic goods).

[85] See e.g. *Apology* 29d-e on Socrates' "usual way" of exhorting others to be good: "Good Sir, you are an Athenian, a citizen of the greatest city with the greatest reputation for both wisdom and power; are you not ashamed of your eagerness to possess as much wealth, reputation, and honors as possible, while you do not care for nor give thought to wisdom or truth, or the best possible state of your soul?" Socrates explicitly flags this line of thought in the beginning of the *Protagoras* as well, in which he admonishes the young Hippocrates to be discerning before flying off to study with a sophist. One would be careful before going to any old physical trainer, and since the soul is of so much greater importance than the body, one should be even more careful before putting one's soul in the care of another (313a-c).

states comprising the Many's bad life, there is no indication that Socrates follows the Many and sees these goods as necessary or sufficient for a good life.

From this argument, we see that the Many have a number of goals for their life long-term: achieve good and healthy bodies, wealth, power, etc. and avoid things like disease and poverty. All of these things, Socrates points out, would be good or bad in the eyes of the Many for no other reason than from how much pleasure or pain accompany them.[86] Socrates challenges them to put forward a different *telos* for calling things good (354b7) but is pessimistic about their ability to do so. His constant request for a *telos* other than pleasure and pain is also a challenge for the Many to come up with further requirements for a good life.[87] If health and wealth are in the service of some more ultimate good, then it may not be the case that pleasure is the only standard for what to choose and avoid. Through their inability to specify what is good in any other way, they have given content to their conception of living well, and that content has turned out to involve nothing more than acquiring pleasure and avoiding pain.

Although Socrates does not specify some alternative content here, he is careful to note how the Many's consideration of the pleasant as what is good has depended upon what the Many *do* consider to be the content of a good life. This conception affects not only the Many's view of their life but also their actions. Because they have pleasure as their *telos*, "you pursue pleasure because it is good, and you avoid pain because it is bad" (τὴν μὲν ἡδονὴν διώκετε ὡς ἀγαθὸν ὄν, τὴν δὲ λύπην φεύγετε ὡς κακόν, 354c3-5). The Many's *telos* explains *why* they aim to act as they do; they pursue what they see as good. We should note two points about Socrates'

---

[86] Dimas (2008) drives home the point that the Many pursue pleasure just because they consider it to be good. The goodness of pleasure is the source of motivation (269). The fact that considering something to be good is the primary or only motivation will be important for separating pleasure out of the main argument below.
[87] Sullivan (1961) sees the stress and repetition of *telos* to point to a hesitancy to accept the proposed hedonism: "were Plato genuinely making the assertion that there really *is* no other τέλος, this repetition would be pointless and laboured" (27). I would expand this point to the constant mention of pleasure and pain as good and bad, as shown below. Bizoń and Sokołowski (2012) also take the repetition as pointed and indicating other options.

73

statement here. First, the explanation underlies a more general principle: we pursue what we see as good, and we avoid what we see as bad. The ὡς plus the participle ὄν indicates that something's (allegedly) being good causes us to pursue it.[88] Second, the ὡς further indicates that the good is indeed alleged.[89] There is no sign that Socrates endorses what he takes to be the Many's principle of action.

We can thus conclude that Socrates' interlocutors have a hedonist conception of the good life, which informs their view of pleasures as the good and the thing to be pursued. Socrates does not himself endorse this conception, but his discussion of living well and what we pursue makes clear that he is nonetheless advancing claims about action and the good life. This allows us to separate two things, the (unspecified) good life, and the good life *qua* pleasant life:

Good Life	Pleasant Life
Everyone strives to live a good life and pursue what is good.	Everyone strives to live a good life and pursue what is good.
The good life consists in [not further specified].	The good life consists in a collection of things that are good only insofar as they are pleasant.

Socrates has established that the Many specify the good life in the latter way by 355a, right before he begins the arguments for DA and EM. He asks if "passing your life pleasantly without pains is sufficient for you," knowing that the Many's answer must be yes (355a2-3). He has completed the task of determining that the Many's good life is just the pleasant life. Because this is just one conception of a good life, however, it may not be necessary for further arguments that concern the unspecified good life. From this point in the text, I argue that Socrates puts forward two arguments in tandem—one for the good life understood in each of these ways.

---

[88] Thanks to Jessica Moss for helping me to see the importance of this point.
[89] See Smyth (1920), §2086.

The Double Argument

Socrates has set up the pleasant life as one conception of a good life; he has focused on it because it is the one that his interlocutors embrace. As we have seen, however, we should keep in mind how Socrates' arguments apply not just to this conception of the good life but to living well in general. Our task is to determine whether DA and EM hold only for the specific good life under consideration or whether Socrates makes an argument that would apply without this specification. As we take a look now at how the passage proceeds, I will argue that the latter is the case.

The clearest indication that Socrates has two arguments in play when he starts upon his conclusions is that he, in fact, makes two arguments when he argues for DA, in Stage 4 and Stage 5 of the passage. After establishing that the Many see the pleasant life as the good life, Socrates notes, as we saw, that the absurdity of the Many's position will become clear if we stop using so many different names. Instead, since they consider pleasure to be the good, "let us call them by two names, first by the good and bad, then again by pleasant and distressing" (355b6-c1). He then makes an argument with good/bad before making one with pleasant/painful, which we can separate in the same way that we separated the good life and the good life *qua* pleasant life. The main points of each line of thought are similar:

Good Life	Pleasant Life
People pursue what is good and avoid what is bad.	People pursue what is good and avoid what is bad.
The good is what is good.	The pleasant is what is good.
For the good to overcome the bad, it would have to outweigh the bad.	For the pleasant to overcome the painful, it would have to outweigh the painful.
It is absurd that one can know that something is bad on the whole but do it anyway because one is overcome by good.	It is absurd that one can know that something is painful on the whole but do it anyway because one is overcome by pleasure.

Why would Socrates feel the need to make the same argument twice with different names?

Perhaps he wishes to make the same point in two different ways for extra clarity. In light of the above distinction between the good life and the good life qua pleasant life, however, we might see these two arguments as making different points. The first one states that no one can do something bad, knowing it to be bad, because one is overcome by the good. The second one states that no one can do something painful, knowing it to be painful, because one is overcome by pleasure. If we specify the good life as the pleasant life, then these two theses will amount to the same thing. If the good life is unspecified, or is specified in a different way, then the first argument will still stand, for whatever the good and the bad may turn out to be. My reading thus has the advantage of explaining Socrates' purpose in constructing two stages for DA and their fit within the passage as a whole: Socrates shows how DA holds for the interlocutors' hedonism while also holding for any conception of the good life.[90]

For DA, Socrates does a curious thing: he seems to emphasize throughout that the pleasant and the good are the same, but he makes his first argument without reference to pleasure at all. A closer look at his emphasis shows that he stops short of a full endorsement, going only so far as to say that the pleasant and good have "appeared" (ἐφάνη) to be the same. The constant emphasis without endorsement, I take it, is a sign that Socrates continues to bring to light what his interlocutors think and how his argument is working for their own commitments. But by his own lights, we achieve DA without a commitment to hedonism: it is absurd to think one could know what is bad and do it anyway by being overcome by the good. This can happen only if we

---

[90] As I mentioned earlier, my reading follows Morris (2006) in taking these to be two separate arguments, only one of which depends on hedonism. Morris, however, does not give an explanation for the presence of the two arguments, noting only how they are related to each other: "it seems as if we have one argument for what we might call Socrates' *overall conclusion* – that it is impossible to do what one knows to be wrong – which is dependent on psychological hedonism, and one which is independent of it" (204; emphasis in text). My reading aims to provide an explanation for this relationship by considering the function of the arguments in the context of the passage as a whole.

misjudge how much good a particular action will bring about. We have secured one important aspect of *epistēmē*: it cannot be "overcome" in a way that leads us to do bad knowingly.

If we can separate the good life and the pleasant life for DA, we might hope to do so as well for EM. Socrates' setup of the importance of measurement indicates that this is indeed the case. He first raises two examples of how measurement could save us: if our salvation in life turned on finding larger things, we would need a measuring stick, and if it turned on finding greater numbers, we would need a calculator. These examples serve to illustrate a general point: if our salvation in life depends on choosing more of one thing and avoiding more of its opposite, then an *epistēmē* that measures those things will be essential for living a good life. Socrates' description of the measuring *epistēmē* is itself a general one: "the art of measurement would have made this appearance powerless, and making clear the truth it would have brought our soul peace, remaining in the truth, and saved our life" (356d7-e2). What will save our life itself depends on what we understand to be the good life. Once more, we are reminded that the life under question matters for understanding the function of Socrates' argument.

This general point is already in place by the time we get to Socrates' "solution," wherein he notes that "salvation in life has appeared to be in the correct choice of pleasure and pains" (357a5-7). The Many's conception of the good life has pointed us to look for an *epistēmē* of measuring pleasure and pain, but, as with the first conclusion, Socrates stops short of saying that this is indeed our salvation, noting once more that this "has appeared" to be the case.[91] If our

---

[91] One might worry whether rather too much is being made of ἐφάνη at 355b6 and 357a6: the verb can sometimes be read more strongly as "has turned out" rather than "has appeared." I would argue that the context surrounding EM especially would incline us towards the latter reading. Socrates' discussion there centers around the difference between appearances and the truth (and indeed begins with the word φαίνεται, when Socrates notes how towers appear larger when closer at 356c5). A measuring *epistēmē* is meant to counteract the power of appearance (ἡ τοῦ φαινομένου δύναμις, 356d4). It is noteworthy, then, that EM only has *appeared* to be of pleasure and pain—not only are the Many confused about being overcome by pleasure, they are also confused about what would save their lives. Thanks to Abigail Breuker for discussion of this point.

salvation in life turns out to be the right choice of something else, then an *epistēmē* suited to

measuring that thing would save us. This hesitancy helps explain his apathy towards further

investigating what the measuring *epistēmē* for pleasure and pain would be; if the good life is not

really to be specified as the pleasant life, then there is little need to determine how to measure

pleasures and pains. When Socrates concludes that making bad choices comes from ignorance

rather than being overcome, he once more uses both good/bad and pleasure/pain: "those who err

do so from a lack of *epistēmē* concerning the choice of pleasant and painful things—*and these*

are good and bad things" (357d4-6; emphasis mine). Before, during, and after discussing the

measuring *epistēmē*, he mentions both dichotomies, an indication that we should keep both in

mind. Even though the Many consider them to be the same, we should not be so quick to

conclude that the real measuring *epistēmē* involves the measurement of pleasure and pain.[92]

If Socrates argues for a measuring *epistēmē* concerning the good life in general, then

there is an obvious answer to the question of what it measures: good and bad things, as opposed

to pleasures and pains.[93]

Good Life	Pleasant Life
An *epistēmē* that will measure what makes up a good life will be our salvation.	An *epistēmē* that will measure what makes up a good life will be our salvation.
Good and bad things can all be measured against each other.	Pleasant and painful things can all be measured against each other.

---

[92] I mentioned earlier that the hedonism at issue is not "enlightened," of the sort that pursuing philosophy and being virtuous as Plato understands it is really the most pleasant thing. It may very well be that what the measuring *epistēmē* produces turns out to be the pleasantest choice, or at least be a part of the pleasantest life. This does not entail that what we are measuring is pleasure and pain itself. Even if Plato thinks the best life is the most pleasant one, the measuring *epistēmē* for that life is not what Socrates lays out when he outlines what the Many think would save their lives.

[93] Frede (1986) is sympathetic to taking things in this direction though is more skeptical of the measuring *epistēmē* as a whole. Pasnau (2021) also finds EM to work for any value—most likely anything that would fill in the left side of the chart. What Socrates is on the lookout for is a homogeneity of value, which would come about under a unified good. He takes this point of the *Protagoras* to be a bridge between the Socratic view and Plato's metaphysics of value; while we may disagree about the Socratic-Platonic connection, I am, of course, very sympathetic to the thought that Plato is developing something substantial about goodness and wisdom here that we see come to fruition in Plato's later thought. We may still be left feeling that, in the absence of hedonism, any satisfying answers about what to measure and how to live are missing, as Vasiliou (2008) notes. We will return to this point at the end of the chapter.

The measuring *epistēmē* measures and weighs what is good and bad.	The measuring *epistēmē* measures and weighs pleasures and pains.

One immediate worry is whether good and bad things can be measured in the way that Socrates thinks pleasure and pain, not to mention large and small, are. I do not think Socrates gives a fully fleshed-out answer here, and I will return to this point at the end of the chapter. However, I do think that Socrates indicates that good and bad do not differ from pleasure and pain in this regard. Towards the end of his argument for DA, he asks how a bad thing could possibly have overcome a good one. Surely it cannot be the case that the good aspects of a bad action were "worthy" to overcome not doing it, for then the person "would not have erred"—the "bad" action would have turned out to have been the right one (355d5-6). Socrates thinks that raises a further question: "According to what are good things worth bad ones or bad things worth good ones? Is it not according to anything else than whenever one is bigger, and the other is smaller? Or one is more, and one is less?" (355d6-e1). Some particular action has both good and bad things that result from it, and avoiding that action also has good and bad results. The action is worth doing whenever the good things are bigger or more than the bad ones, and the bad ones are smaller or less.

The upshot is that Socrates does indicate here that we can measure good and bad things, or at least what makes for a good or bad choice, against each other; they, like pleasure and pain, differ only in regard to their relative size or quantity. We would thus be able to weigh them with a measuring *epistēmē*, if we possessed one. As with DA, EM can go forward without a necessary attachment to pleasure and pain; we should consider the measuring *epistēmē* to be a serious proposal from Socrates that is not just an *ad hominem* attack against Protagoras and his fellows. Whatever the good life may turn out to be, *epistēmē* of how to measure those good and bad things would be our salvation in life. The good and bad things do not have to be pleasures

and pains to be objects of the measuring *epistēmē*.[94] As a result, we can still hold onto the features of *epistēmē* that are developed in the *Protagoras*. In particular, we have a measuring *epistēmē* whose function is to cause us to make the right choices and lead a good life. Such an *epistēmē* would certainly be a very good candidate for the kind of thing that would guarantee virtuous action, of the sort that we sought in the *Hippias Minor*. If this is the case, we might hope to see Socrates address the connection to SP1. Fortunately for us, the text proceeds by doing exactly that.[95] Let us turn now to the relevant passage.

The Socratic Paradoxes and the Measuring *Epistēmē*

If what I have outlined above is correct, then the double argument should be able to inform Socrates' argument going forward. After Socrates finishes addressing the Many, he turns back to conversing with the sophists—not only Protagoras, but Prodicus and Hippias as well. In this final section, we will look at their brief conversation, which marks a shift both from the preceding argument and from the subsequent final pages, in which Socrates turns again to Protagoras alone to discuss courage and wisdom. Though this division is not typically taken to be important, the shift in characters highlights a shift in the subject matter. Socrates now seeks to drive home to the sophists the impossibility of voluntary bad actions.

---

[94] Using the measuring *epistēmē* in this way also helps avoid potential problems with a measuring *epistēmē* of pleasure, whatever one's stance on hedonism may be. Dyson (1976) finds there to be something fundamentally problematic about using the measuring *epistēmē* for pleasures: the model is supposed to measure actual pleasures, not expected ones, but the person trying to use the measuring *epistēmē* for her own life has access only to the latter, not the former. I am not convinced that there would be a difference between actual pleasures and expected pleasures for someone with the measuring *epistēmē*; though it would certainly be a very powerful type of *epistēmē* that can show how much pleasure will actually come to pass, that is, in a way, the point. But whatever we might say about pleasure, the worry does not seem to be so pressing when we consider how the measuring *epistēmē* might apply to something other than pleasure and pain. On the other side of the coin, Dimas (2008) contends that measuring pleasure and pain would be relatively easy to do—most adults have a good sense of what produces pleasure and what does not, more clearly than what is good and what is not. Why, then, are there so few courageous people? Shouldn't they know well that going to war is pleasant (if, in fact, it is, which Dimas doubts)?

[95] It is also worth mentioning that Socrates brings up SP1 right before his first argument for courage as well, in his notorious critique of Simonides' poem. He does not offer any sort of argument for it but simply asserts it, and there are several peculiarities with that passage that would incline us away from drawing too many conclusions. See Frede (1986) on the relationship between it and the rest of the *Protagoras*.

80

I noted at the start of this chapter that the inclusion of Hippias in the *Protagoras* is one reason to read it in tandem with the *Hippias Minor*, and the discussion here highlights that connection. In the *Hippias Minor* Socrates came to the conditional conclusion that no one does injustice voluntarily, but he sought further elaboration and explanation of that thesis. The direct mention of both Hippias and SP1 in the *Protagoras* indicates that we might find such elaboration and explanation here. Socrates takes himself to have arrived at SP1 on account of the preceding argument, so the *epistēmē* that we have encountered there is not only one that is not overcome by other forces but also one that can support SP1.

Socrates lays out the following points in favor of SP1:

(1) The pleasant is good and the painful is bad. (358a5-6)
(2) All actions for living painlessly and pleasantly are noble. (358b3-5)
(3) A noble deed is good and beneficial. (358b5-6)
(4) If the pleasant is good, then no one knowing or thinking that something else is better than what he does, and is able to do the former, still does the latter, although it is possible to do the former. (358b6-c1) [96]
(5) Being overcome by oneself is ignorance, and mastering oneself is wisdom. (358c1-3)
(6) Ignorance is to have a false opinion and to be deceived about worthwhile matters. (358c4-5)
(7) No one goes voluntarily for bad things or what he thinks are bad things, nor is it in human nature to want to go for what one thinks is bad instead of good. Whenever someone is compelled to choose one of two bad things, no one chooses the greater if it is possible to choose the lesser. (358c6-d4)

(4) and (7) both bring to mind SP1: no one chooses worse things or goes towards bad ones. It is certainly exciting for our purposes that Socrates feels himself able to wheel out SP1 after his earlier arguments about *epistēmē*. We have been looking for something that can guarantee and explain SP1, and we get here not just a hint but an explicit statement of commitment to it. The

---

[96] Given that we have so far been focusing on *epistēmē*, it might be surprising, if not troubling, to encounter the phrase "knowing or thinking" (εἰδὼς οὔτε οἴομενος) instead of speaking of *epistēmē* alone. I would maintain, however, that the inclusion of οἴομενος does not affect our conclusions about *epistēmē*. EM solely concerns an *epistēmē* that would save our lives; there is no mention of an analogous "measuring belief," nor is there any reason to think such a thing might exist. Even regarding DA, whether or not one pursues what one believes best is of little importance if one is arguing about the power of knowledge. See Penner (1997), Callard (2014), and Pasnau (2021) for perspectives on the importance of knowledge despite belief-*akrasia* cropping up at this point in the dialogue.

trouble with accepting (4) and (7), similar to the trouble we had before, is that it seems to hinge on (1), that the pleasant is good. (4) in particular constructs this conclusion as dependent on the identification of the pleasant and good, putting the status of SP1 here in question.

I would maintain that considerations of the double argument can be used to alleviate these worries as well. Whatever one's standard for living well might be, whenever one employs the measuring *epistēmē* for that standard, one will be able to see whether the goods of a particular action outweigh the bads or not. For the sophists, as for the Many, the standard has turned out to be living pleasantly. It is striking that Socrates asks the sophists whether the pleasant is good, as we see in (1), if he thinks that he has already provided a convincing argument. Rather, it makes more sense to think that he is confirming that the sophists' conception of a good life is the same as the Many's. Indeed, Socrates describes good and noble actions as those "for living painlessly and pleasantly." By stating that all actions that contribute to living pleasantly are good, we are confirming the actions that are constituents of a good life.

(4) and (7) tell us a couple of things about motivation: no one chooses a worse action when one is able to choose the better, and no one voluntarily goes toward what is bad. These claims come in response to the objection that one can be overcome and go towards what one knows is worse. Recall that when Socrates was questioning the Many, it was important for him to establish that they viewed the pleasant as the good. This was not merely a normative claim but also an explanatory one: they pursue pleasure *because they see it as good*. Understanding the source of the Many's motivation for pleasure highlights an important fact: the possibility of something having the power to overcome does not even get off the ground if nothing about it appears good. Socrates emphasizes early on that the worthless pleasant things are bad because of their consequences, not in themselves; the Many still consider their pleasantness itself to be

good.[97] Socrates promoted DA by showing that the worthless pleasant thing was chosen wrongly not necessarily because it was a mistake to think it had any goodness whatsoever but rather a mistake to think it better than the alternative course of action. Though the smaller amount of goodness still has motivating power, when weighed properly it will no longer be seen as the good thing to do on the whole (no matter its residual goodness). Socrates appeals to the fact that one can only be motivated by what one sees as good. Hence one must misjudge the size of goodness if one acts badly—evidence of a lack of *epistēmē*.

The way to understand the protasis in (4) is to recognize that Socrates is flagging the sophists' standard for a good life. If all people aim to live well, and the standard for living well has been established, then any actions that may appear to be living well but, through the measuring *epistēmē*, are recognized not to be will no longer be actions that should hold sway over a person. One is tempted by a pleasant thing only because one sees it as good; when one no longer sees it as good, then one is no longer tempted. Because the pleasant is the sophists' standard for a good life, the measuring *epistēmē* will allow the sophists no longer to be motivated by a less pleasant thing, and so they will agree that (4) describes their action. (4) does not, however, preclude other standards of a good life. For these other standards, what seems to be good may differ, but the principle is the same: what seemed to be good will not overpower us if we grasp that it is outweighed by something better overall.

If we think back to the double argument, we recall that Socrates hopes to explain DA and EM for whatever one's conception of the good life might be. Things that seemed to be able to

---

[97] See also Dimas (2008), who thinks that the argument against Protagoras and the Many cannot go through unless pleasant things are considered good *qua* pleasures: "Someone who feels a strong psychological pull toward a particular pleasure might act to get it. If this is bad also *as* pleasure, nothing about it, not even its pleasantness can be considered as being good. Everything about it is bad and the agent would not even have one single reason to go for it." (268).

overcome us were thought powerful because doing or avoiding them appeared disproportionally good. When one has the measuring *epistēmē*, however, one is able to determine just how good something *is*, not how good it appears to be, for one's life. Things are only attractive insofar as they are good, so using the measuring *epistēmē* can remove this attraction. No one will choose to do the worse thing if the worse thing is shown not to be good. We can now see how SP1 can be introduced: a measuring *epistēmē*, through bringing to light what is really good for one's life, will prevent us from being overcome, by removing the power from bad but (once) attractive things.

I have argued that DA and EM can successfully support SP1 in the face of the objection that someone with *epistēmē* could be overcome by other motivations. I will close this section by noting that this advancement has gotten us further than may be expected, as it is not clear whether there are any other contenders for voluntary wrongdoing. If one does wrong without being "overcome," it would be in one of two ways: (1) one follows bad reasoning to do what one thinks is the right thing (à la Aristotelian vice), or one chooses the bad thing with full knowledge and for no other reason, a sort of clear-eyed evil. There is debate in contemporary ethics over whether the latter notion is coherent.[98] Whatever one's intuitions may be about the issue today, this kind of "guise of the bad" thinking is pretty clearly absent from ancient philosophy.[99] As we saw, people avoid what they see as bad; choosing a bad thing in the absence of any apparent good is not a particular threat—certainly not in Plato's eyes—to SP1.

---

[98] See Velleman (1992) for the most notable recent argument in favor of the "guise of the bad." Vogt (2017) offers arguments against this position that she deems in line with the general ancient view.

[99] See *Meno* 77c-78b, in which Socrates takes himself to show very briefly that no one desires bad things while knowing they are bad. It is notable that Socrates' argument also depends on the good life: a key premise is that no one wishes to be miserable and unhappy. Aristotle does not have a space for this idea in his ethical thought either. Either one knows what to do (with the various murky qualifications from *Nicomachean Ethics* VII.3) but is overcome by appetite or spirit (*akrasia*), or one's reason has been corrupted (vice); there is no clear-eyed action against the good and one's own happiness.

Aristotelian vice does not pose a threat either. A vicious person would not have *epistēmē* of good and bad at all, not to mention an *epistēmē* of measurement; such a person is deeply misinformed about what is good and bad. And if we are to take SP2 seriously and think virtue is knowledge, someone acting viciously is ignorant. The *Hippias Minor* indicated that voluntary action in the realm of virtue involved power and know-how—the latter of which is *epistēmē*. Under this conception, a vicious person would not be acting voluntarily. Socrates' argument in the *Protagoras*, then, attacks the only plausible candidate for objecting to SP1: that of knowing but being overcome by some other motivation. If that candidate is defeated, we might reasonably conclude that we have shown voluntary wrongdoing to be impossible.

Conclusions

The *Protagoras*' famous denial of *akrasia* provides us with fruitful ground to explore the development of Plato's epistemology. A notion of *epistēmē* that stays strong in the face of pleasure and measures up our choices accurately is one that can guarantee virtuous action. Socrates' arguments that pleasure is the good show that his interlocutors have a certain conception of the good life, although he is also presenting arguments about the good life in general. Because of this distinction, we can see two parallel arguments concerning *epistēmē* and wrongdoing: one that involves pleasure and pain and one that does not. We can thus take Socrates as presenting a serious proposal that informs and advances our conception of *epistēmē*. The *Protagoras* answers the *Hippias Minor*'s questions by presenting an epistemology that responds to and secures the Socratic paradoxes.

We have not, however, gotten a very clear grasp of this notion of *epistēmē*. One significant remaining task is to explain what, exactly, the measuring *epistēmē* would measure if it is not measuring pleasure. We know in general that it has to do with the good and bad, such that

we will make good choices and lead a good life. The *Protagoras*, however, has not offered up

an alternative to hedonism or an alternative content for the measuring *epistēmē*. How does the

measuring *epistēmē* work, such that having such an *epistēmē* will guarantee right actions and will

save our lives? Though the *Protagoras* does not tackle this question, I believe that we will find

answers in the *Republic*.

Chapter III: Truth and *Doxa* in the *Republic*

Introduction

We have seen so far that Plato's treatment of the Socratic paradoxes has led him to develop a notion of *epistēmē* that guarantees virtuous action by using measurement to determine the right thing to do and take away any motivation not to do it. In the *Protagoras*, Socrates described this *epistēmē* as measuring pleasures and pains, but we saw that hedonism is not necessary for employing the measuring *epistēmē*. What the *Protagoras* did not tell us is what might be able to replace pleasure and pain as the content that would be measured. We saw in general terms that *epistēmē* involves good and bad and is able to show us the truth beyond appearances. But without a further explanation of what and how *epistēmē* measures, we have not yet acquired a satisfying picture of how *epistēmē* can support the Socratic paradoxes.

In this chapter, I propose that Plato's *Republic* fills out this account in the following way: *epistēmē* measures truth itself. Making virtuous decisions involves discerning which choices have more truth and which ones have less. When one acquires *epistēmē*—in particular, *epistēmē* of pure intelligibles, the objects with the fullest truth—one will be able to measure up truth and falsity about things in the world. Knowing how much truth certain things possess will put our souls in the proper condition to do well, and only well—in the words of the *Protagoras*, it will "save our life."

This proposal may be true but is far from clear. Before we can accept it as a satisfying part of Plato's philosophical picture, we must deal with several worries, three of which loom especially large. The first worry concerns the relevance of truth: Plato is not usually (if ever) taken to consider truth as important or relevant to the Socratic paradoxes. Though we usually take *epistēmē* to involve truth in some sense, it is not clear how much truth itself is a focus of

Plato's epistemology. I have been arguing for the connection between the Socratic paradoxes and Plato's epistemology, but this picture does not seem important to either, rendering a turn to truth strange and unwarranted. The second worry concerns the measurement of truth: we do not talk about weighing truth, as we can "weigh" the sizes of towers or (supposedly) pleasures and pains. My proposal above mentions deciding between things that have more and less truth, but how could truth come in these relative amounts? And what exactly are the "things" that have more or less truth?

The third worry is not specific to my proposal but is instead the long-standing Problem of the Gap. To recapitulate, the *Republic* posits a "Two Worlds" epistemology, in which *epistēmē* and *doxa* (belief or opinion) deal with completely separate objects. The objects of *epistēmē* are the intelligible Forms; the objects of *doxa* are the sense-perceptible things in the world as we conceive it. I have argued that Plato means for *epistēmē* to be used to guarantee virtuous action to its possessor. But on the Two Worlds view, it is unclear how *epistēmē* can guarantee anything at all. We have no *epistēmē* of things in the sense-perceptible world, so how can it improve our actions? The *Republic*'s picture of *epistēmē* appears to be bereft of any practical benefit.

This chapter will proceed by tackling these worries in turn; in doing so, we will be able to elaborate upon the picture of *epistēmē* as a measurer of truth. The *Republic* is a very long and rich text, and we will be focusing in on two main themes that will help us understand Plato's treatment of *epistēmē*. The first is Socrates' initial discussion of education, which comes primarily in Books II and III. There Socrates gives his most detailed views on truth and its relationship to acting and living well. The second is Socrates' account of *epistēmē* and *doxa* in the "middle books," Books V through VII. In these books, we will be able to consider *epistēmē* and *doxa* themselves in some detail and see how they fit in with Socrates' views on truth and

88

measurement. By the time we have examined the relevant passages, it is my hope that we will see the connections between these two themes and their importance for our project as a whole.

We shall start by noting a new role for the Socratic paradoxes in the *Republic*. SP1 is not absent from or rejected in the *Republic* but instead comes up in a different form: Socrates now speaks of voluntary bad and the impossibility thereof in terms of accepting falsehoods and being deprived of true *doxai*. We will then discover how truth is something that can be measured by looking Socrates' account of lying, in which we will draw out two different types of truth. One type concerns how much a *doxa* will turn our soul towards or away from *ta onta*, or the most real things; the other type concerns how much a *doxa* maps onto the events and objects of the changing world. Each of these types of truth will appear in different amounts in a particular *doxa*; the *epistēmē* of measurement will weigh these types of truth to determine which statements we should accept in our souls and which we should not (and, as rulers, which statements to impart in citizens).

This brings us to our major epistemological upshot: *epistēmē* improves our *doxai* by being able to measure the truths found in each *doxa* and thus determining whether we should accept them. *Doxai*, as describing aspects of the changing sense-perceptible world, are never fully true or false but instead have both truth and falsity. The fact that *doxai* have an incomplete share in truth brings to light an important insight: *epistēmē* and *doxa* can be connected to each other on account of each one's relationship to truth. With this in mind, we can make sense of Plato's discussion of *epistēmē* as calculative, as well as a model for virtuous behavior. As a result, we are able to show how Plato can bridge the Problem of the Gap: measuring truth allows us to use *epistēmē* to shape our *doxai*, despite the two states remaining completely separate in their objects. Finally, the connection between the measuring *epistēmē* and having the best *doxai*

shows us how *epistēmē* entails virtuous action, thus giving us a concrete understanding of the Socratic paradoxes.

Truth and the Socratic Paradoxes

Since the Socratic paradoxes have so far been a helpful guide to Plato's epistemology, let us begin our study of the *Republic* by investigating what they might teach us here. In both the *Hippias Minor* and the *Protagoras*, we saw that SP1 was an important part of Plato's thinking about wrongdoing and *epistēmē*. The *Hippias Minor* showed Socrates first focusing on one virtue, justice, faced with the task of explaining how no one does injustice voluntarily. In order to provide this explanation, he expanded the investigation to all the virtues in the *Protagoras*, searching for an understanding of SP2 that would guarantee virtuous action in general. The *Republic* deals with many of the same themes, seeking to explain in detail, like the *Hippias Minor,* why doing justice is superior to doing injustice and, like the *Protagoras*, how the virtues are related to one another. SP1 and SP2 figured prominently in these earlier discussions, so we should expect that they will also be important in the *Republic*'s arguments and the theory of virtue that Socrates builds therein.

A first glance at the *Republic* threatens to disappoint us greatly in this regard. Early in the dialogue, Socrates' interlocutors praise injustice as superior to justice, either sincerely or as an attempt to elicit Socrates' response. Adeimantus, ostensibly doing the latter, proposes that everyone would be unjust if they could get away with it; they do just things only because they cannot. This line of thought culminates in the following claim: "no one is just voluntarily, but they find fault with doing injustice because of cowardice or old age or some other weakness, being unable to do it" (366d1-3). Adeimantus thus argues for the opposite of SP1: no one does justice voluntarily. As Socrates is set to provide a response, we might think it likely that he will

90

end up wielding SP1 and explaining that injustice is what is involuntary. But this explanation never comes. Socrates never states in the *Republic* that no one does injustice voluntarily, nor does he state SP1 as applying to virtuous action in general. Adeimantus' challenge is similar to SVW from the *Hippias Minor,* and it could likewise be answered by SP1. So why does it not come up?

One response could be that Plato has now abandoned the Socratic paradoxes and so cannot appeal to them in arguing for the superiority of justice. If that is the case, however, we might hope to see it stated explicitly as well, which we do not. Recall that in the *Hippias Minor* Socrates gave an account of voluntary wrongdoing as wrongdoing that is done with both power and know-how in a certain domain (which, in the case of virtues, is *epistēmē*). The *Republic* does not give us a different account of the voluntary, and we do not get any indication from Socrates either that there is voluntary wrongdoing or what it would consist in.[100] It appears, then, that Socrates does not respond to Adeimantus' tantalizing challenge at all. In fact, we might wonder whether voluntary wrongdoing is even a concern here: the entire work contains only a handful of uses of ἑκών, most of them not laying out any sort of general rule but rather some specific instance of acting voluntarily. Instead of leading us to believe that Plato has abandoned his project, however, the use of ἑκών in the *Republic* allows us to focus in on a

---

[100] I believe that one can take this line of thought to counter the objection that Book IV's account of tripartition nullifies the Socratic paradoxes. According to the objection, the division of the soul into an appetitive, spirited, and rational part means that each part can control an agent. Unlike Socrates' "denial of *akrasia*," one can be akratic in the usual sense and do wrong voluntarily as a result. The problem with this objection is that there is no indication that being led by appetite or spirit would constitute voluntary wrongdoing for Plato. He never discusses these cases as being voluntary, and there does not seem to be any reason to think that he has changed his mind of *epistēmē* being a requirement for voluntary action in the domain of virtue. Furthermore, as we shall see towards the end of this chapter, someone who does have *epistēmē* is naturally going to have the other parts of her soul in the right condition. SP1 and SP2, therefore, are not obviously affected by Plato's introduction of tripartition. See also Gerson (1989) on the insufficiency of this line of thought.

couple of key instances where it is employed—instances that signal an important transition and reformulation of the Socratic paradoxes.

There are two places in the *Republic* where Socrates asserts something that no one would do voluntarily. They both come in Socrates' outline of education in the just city. They have an interestingly similar structure: in a short space, Socrates asserts a general thesis about everyone's relation to the good as well as a more specific claim about our attitudes towards truth and falsehood.[101] The first of these comes near the end of Book II, where Socrates is outlining the stories to be told about the gods during the education of the youths. Socrates first asserts that a god would never make himself worse, because no one would do so: "surely no one among gods or humans, Adeimantus, would make himself voluntarily worse in any way" (381c3-4). A corollary of not making oneself worse, according to Socrates, is that the gods would not put on deceptive guises. This stems from the fact that both gods and humans have a certain relationship to the truth: "Don't you know that all gods and humans hate a true lie, if one can call it that? [. . .] no one wishes to lie voluntarily to the most important part of oneself and about the most important things, but one fears to possess it there most of all" (382a4-9). Though Socrates' focus in this passage is on the gods, he also lays out a universal, SP1-like fact, that no one voluntarily lies in a certain way.

Socrates' statement about voluntarily worsening oneself finds a parallel in Book III, right before the (in)famous Noble Lie passage. In this case, instead of making oneself worse, being made worse is Socrates' focus: "don't you think that people are deprived of good things involuntarily, but of bad things voluntarily?" (413a5-6). The reason this principle is important is

---

[101] In this chapter, I will be speaking of and translating "lies" and "falsehoods" (and related parts of speech) indiscriminately. The same word is used in both cases in Greek, and so one should try to read the ensuing discussion with as little baggage inclining towards one or the other as possible.

that Socrates wishes to make a distinction concerning our attitudes towards true and false *doxai*.

He claims that "a *doxa* seems to me to depart from the mind either voluntarily or involuntarily—

the false *doxa* departs voluntarily when one learns better, and every true *doxa* departs

involuntarily" (412e9-413a2). True and false *doxai* depart involuntarily and voluntarily,

respectively, because "it is bad to have a falsehood about the truth, and it is good to have the

truth" (413a6-7). One would never voluntarily give up a true *doxa* because having the truth is

good and having falsehood is bad.

It is no accident that these points have come up in this way. Socrates' goal in this part of

the *Republic* is to outline how the citizens of his ideal city should be educated; this education is

vitally important for shaping the souls of young people so that they act well and lead good lives.

Most importantly, having the good and being deprived of the bad are intimately connected with

truth and falsity. Instead of focusing on the fact that no one does injustice voluntarily, Socrates

asserts that no one accepts falsehoods voluntarily and, on the other side of the coin, no one

abandons the truth voluntarily. Living well, on this account, involves having the truth and

eschewing falsehood. At the end of the *Protagoras*, we were faced with the question of what

epistēmē could measure such that we would live well. If people are deprived of truth only

involuntarily, then an *epistēmē* that measures truth itself would guarantee that people continue to

have truth and avoid falsehood. Having the truth is, in some way, important enough for living

well that being deprived of it is universally despised.

Nevertheless, the passages I noted above raise as many questions as they answer. The

first one describes our attitude not towards lying in general but to a certain kind of falsehood,

puzzlingly called a "true" falsehood. It comes in the middle of a strange discussion about

different types of lying, some of which Socrates calls useful. It is troubling to see Socrates not

only accept lying but actively advocate for it, from telling false stories to children to having society at large believe the so-called Noble Lie.[102] Getting clear on Socrates' attitude towards truth and lying will thus be necessary for determining how we should understand the new focus of the Socratic paradoxes and also getting to the heart of some strange and unsettling parts of his view.

True Lies and Useful Falsehoods

In order to make sense of Socrates' treatment of truth, and in so doing to understand how truth could be a focus of *epistēmē*, we must take a closer look at these passages about truth and lying. The first, as mentioned before, comes as Socrates is describing what stories to tell about gods in the ideal city. Gods never deceive, because they have no reason to lie, but for humans it is necessary to distinguish between two kinds of lying. One of these kinds is such that all humans would never do it voluntarily, the "true" lie that was picked out earlier. Socrates' interlocutors are, perhaps understandably, confused at his description, and so he explains further:

> You think I am saying something profound. But I'm saying that
> everyone would least of all welcome lying about *ta onta* in the soul
> and having lied and being ignorant and having and possessing a
> falsehood there—they hate it in that kind of place most of all.
> (382b1-4)

The lying that is never done voluntarily is specified as lying about *ta onta*, or "the things that are," which Socrates calls the most important things (382a8). Whatever *ta onta* might turn out to be, which we will consider in the next section, harboring a falsehood about them—being deprived of their truth in one's soul—is hated by everyone.

---

[102] I will not be addressing the Noble Lie specifically, but the discussion of truth that is to follow should apply to it relatively easily. There has been a good bit of recent work in seeing the Noble Lie in a more positive light; Rowett (2016) regards the status of the Noble Lie as not a lie in any important sense. Her consideration of the important truths that it contains is, in my view, on the right track, though probably somewhat more progressive than Plato actually was. See also Page (1991), Gill (1993) and Baima (2017a) on the benefits of the Noble Lie.

In contrast to true lies, Socrates describes another type of falsehood: lying "in words" (ἐν τοῖς λόγοις). Whereas no one voluntarily embraces a true lie, lies in words are "sometimes useful, so as not to be worthy of hatred" (382c7-8). Humans sometimes use lying in words because there are situations where that course of action would be best or necessary. Socrates gives three such situations: (1) "against our enemies," (2) "some so-called friends, whenever they attempt to do something bad on account of insanity or some ignorance," (3) "in the cases of stories we were just talking about, on account of not knowing how it truly was in those olden days" (382c8-d2). In these circumstances, we may have reason to say something false because of the particular details of our situation. Socrates goes on to argue that the gods never have reason for these sorts of falsehoods. The situations described above are only issues for humans: our condition in the world necessitates the embrace of some kinds of falsehoods.

As a result, true lies and lies in words are importantly different. Humans have use only for the latter, whereas gods have use for neither. This difference seems to stem from the subject matter of these falsehoods: there is a class of things that only humans, not gods, deal with and sometimes tell lies about, but *ta onta* are of concern to both gods and humans and must always be represented truly. Lying in words is an "imitation" of true lies: the falsehoods that are sometimes useful are imitations of *ta onta*. The subject matter of the two types of lying thus seems to be different, if one is an imitation of the other. We may accordingly be able to separate true lies and lies in words as two *types* of falsehoods: their conditions and objects are different, so as lies they are different.

If Socrates is outlining two different types of falsehood, we may expect to find truth corresponding not just to falsehood in general but to each of these types of falsehood. We would thus be able to separate out two types of truth: one that concerns *ta onta* and is embraced by

95

everyone, and one that is an imitation of this type of truth. We have some initial indications here that truth may not be a strict binary matter: Socrates is considering different ways and conditions in which our souls may have truth and falsity. I am not the first to suggest that there are different types of truth in the *Republic*, as we will see presently. Our task at hand, however, is to figure out how to conceive of these types of truth, especially the kind that no one embraces voluntarily.

We see further confirmation of this picture of truth in our second passage at 413a. After claiming that being deprived of the truth is a bad thing, Socrates clarifies what it means to have the truth: "Doesn't it seem to you that to believe the things that are is to have the truth?" (ἢ οὐ τὸ τὰ ὄντα δοξάζειν ἀληθεύειν δοκεῖ σοι εἶναι, 413a7-8). We should notice two things. First, *ta onta* appear once more in conjunction with the truth. In this case, believing *ta onta* is simply having the truth, as opposed to earlier where some falsehoods were not deserving of hatred. Having the truth about *ta onta* thus seems to be of utmost importance, compared to any other type of truth. Second, Socrates uses *doxa*-language to speak about truth: having the truth is *doxazein ta onta*. Recall that *doxa* concerns things in the sense-perceptible world; for Plato, it is not the highest level of cognition, but here it is concerned with the sort of truth afforded it. Since we know that the truths at issue have to do with *doxa*, we can focus our search for understanding the two types of truth accordingly.

Let us take stock: SP1's new formulation in the *Republic* points us towards truth as at the center of the connection between *epistēmē* and right action. Truth and lying themselves, however, are not so clear cut, and we must determine how to conceive of the different types of truth and their relative importance. The presence of multiple types of truth is promising for the thesis I stated at the beginning, that *epistēmē* measures truth; a complex and non-binary picture would indicate that truth is something that has the potential to be measured. We are still left with

96

the question of filling in the gap between *doxa* and *epistēmē*. But before we can offer an answer, we must get clear on what Socrates means by *doxazein ta onta* and never lying voluntarily about them.

Truth, Forms, and the Gods

The most promising path forward involves looking to *ta onta* themselves: how should we understand *ta onta* in *Republic* II and III? A direct translation, "the things that are," is not much help. We might at first take it at face value: a metaphysically unambitious "what is the case," or something of that kind. The context, however, indicates that *ta onta* is more elevated. All gods and humans hate to harbor falsity about it, and it concerns "the most important things." Furthermore, lying "in words" seems itself to include falsehoods about what is the case: each of the situations Socrates mentions would be lying about *ta onta* on this reading. As Socrates is trying to separate lying about *ta onta* from these situations, we should expect the type of truth that corresponds to *ta onta* to be of a loftier sort.

As I mentioned above, a few scholars have, like me, taken Socrates' discussion of true lies and lies in words to indicate two types of truth in the *Republic*. To this end, they have offered proposals for how to understand *ta onta* in Books II and III and the type of truth associated with it. There have been two main proposals so far. The first is that *ta onta* are the Forms; the second is that they are ethical truths. I will examine each proposal in turn, and we will see that neither is sufficient in the context of the discussion. This point should lead us to embrace a third, more expansive proposal for *ta onta* and its corresponding truth.

97

The prevailing view about this type of truth has been that *ta onta* refer to the Forms.[103] The strongest evidence for the Form view of *ta onta* is that *ta onta* describe the Forms in later parts of the *Republic*. In the middle books, for instance, *ta onta* are unchanging, fully intelligible, and deficient in no way. Socrates speaks of the Beautiful itself and the Just itself, the Forms that are the concern of our highest cognition. It is no surprise that both gods and humans would hate falsity around them, since metaphysically they are the most important things. Even though Socrates does not yet speak of Forms in Books II and III, they occupy such a central place later on that it would not be remiss to read them back into this early discussion of education.[104] We should think, therefore, that the importance of Forms runs throughout the *Republic*, and they as *ta onta* are worth having truth about in every way.

Importing the later metaphysics, however, will itself cause problems for this reading. Recall that one thing we know about *ta onta* here is that *doxazein* them will provide the truth. But if we take *ta onta* simply to be the Forms, then we do not have *doxa* about them, let alone *doxa* that has the truth.[105] The truth about the Forms would involve having *epistēmē*, which is clearly not the focus in Books II and III. Socrates' initial distinction between true lies and lies in words, to be sure, does not mention *doxa*. But in order to reconcile the claims about *ta onta* in

---

[103] Woolf (2009) calls these truths "philosophical truths": "roughly, those truths expressed by the accounts of Forms that it is the task of the philosopher in the *Republic* to discover" (15); Simpson (2007) also sees this kind of truth along these lines. Brickhouse and Smith (1983), though drawing distinctions between kinds of truth in a different way, also seem to view these truths as of the Forms and attribute the distinction to Vlastos (1965).

[104] See Woolf: "When one first encounters [these passages], it is natural enough to read them non-technically. Once we have digested the later metaphysics, however, it is unlikely that we are intended to read phrases such as 'the things that are' with wholly innocent eyes" (20).

[105] As mentioned in the Introduction, many Two Worlds-ers do not take *doxa* about the Forms to be impossible. But this does not vindicate a Form view here. *Doxai* about the Forms are at best blind, as Socrates says in Book VI; *epistēmē* of the Forms is what would really have the truth. Even if *doxa* about the Forms were possible, such cognition is deficient. Vogt (2012) is illustrative: "knowledge and belief have each their own kinds of objects. This does not mean that they cannot—deficiently—be directed at what they are *not* adequate for. Belief can engage with matters that it 'does not live up to,' and since people have views about all kinds of matters, this happens all the time. Similarly, knowledge could be directed at what it is not made for, and this, too, would be a deficient kind of approach" (64).

Books II and III, we would need to say that *ta onta* were the Forms in Book II, but then not in

Book III (since there we are supposed to *doxazein* them), and then again in Book V. This strikes

me as implausible, or at least much less plausible than taking Books II and III, given the

similarity of context and subject matter between them, to be talking about *ta onta* in the same

way. It is more likely that the middle books would be the odd ones out.

We are provided with further reasons to hesitate about the Form view from Baima

(2017b), who attacks this proposal for *ta onta* and provides an alternative. First, he argues,

Socrates warns his interlocutors that he is not "saying something profound" about lying about *ta*

onta (382b1). We should thus incline away from thinking he has in mind the Forms, which are

the most metaphysically profound things out there. Second, the discussion in this passage is in

the context of education and good character development, not in the context of metaphysics and

what is most real. Baima takes a "progressive" reading of *ta onta*: we should not assume that

the metaphysics of the later books is already in place, and it would be strange if it were. Rather,

we should understand *ta onta* in light of the surrounding discussion: Socrates is trying to form

the souls of youths to be ethically upright. Truths about *ta onta* are instead ethical truths, like "I

should not hate my fellow citizens" or "I should care for my relatives."[106] These are the sorts of

things that we would hate to get wrong, because we all want to live ethically.[107]

---

[106] See Baima pp. 14 and 16 for these and other examples of what he takes to be ethical truths.

[107] Baima gives a third reason for not thinking of *ta onta* as Forms, based on later evidence in the dialogue: he argues that in the Cave discussion Socrates says philosophers should lie to non-philosophers about the Forms. But if true falsehoods should never be accepted by humans, then truly lying is not about the Forms. Overall, I'm not convinced that the Cave passage says we should lie about the Forms. Baima quotes the following text: "Soc: Wouldn't it be said of him [viz., the philosopher] that he'd returned from his upward journey with his eyesight ruined and that it isn't worthwhile even to try to travel upward? And, as for anyone who tried to free them [viz., the non-philosophers] and lead them upward, if they could somehow get their hands on him, wouldn't they kill him?" (517a3-6, Baima 11). He takes this passage to imply that the philosopher will lie to the citizens about what the Forms are really like, but this passage does not say anything about lying to the citizens or even withholding truths from them. Rather, it serves to emphasize that someone who comes straight from outside the cave may at first seem to be crazy and should not be hastily dismissed. This passage is not describing the ideal city but what happens to philosophers in typical cities (probably an allusion to Socrates' less than stellar treatment by the citizens of Athens). In these cities, it is an unfortunate fact that philosophers will be derided; it does not even seem possible for them to

I believe that Baima's second point is on the right track, but not his first, so let us start with the latter. It is certainly undeniable in the text that Socrates claims not to be saying anything profound, and we might therefore be tempted to think that he means that the subject matter of *ta onta* is what is not profound. But Socrates' focus in his elaboration is where the true lie would be located, not what its content would be. He mentions three times that the soul is where everyone would hate a true lie, finishing his explanation by saying that it would be an especially hated place for such a thing. The main thrust of Socrates' point is that the soul is an area of the greatest concern for us. As we saw in the previous chapter, the importance of the condition of the soul is not a mysterious thing for Socrates to emphasize and a pretty consistent part of his philosophical agenda. Socrates' reassurance, then, does not preclude *ta onta* themselves, the things we want in our souls, from being of the utmost cognitive and metaphysical importance. Because of this, we do not have good reason to think that the Forms do not pertain to what is under discussion.

What Baima is right to point out, however, is that the discussion has been primarily an ethical one, and focusing further on the context of the passage will allow us to grasp what truths about *ta onta* are. As I have mentioned, Socrates is considering how the guardians of the city may be well brought up, in order to see how a just city might come about. This upbringing involves both musical and physical education, the former of which involves stories told to children. The stories that are the worst in this regard—to the point that they must not ever be told—are those that "it is necessary to censure first and foremost [. . .] whenever someone makes a bad likeness in a story about how the gods and heroes are" (377d7-e2). Much of ancient Greek

_____

lie in this condition. On the other hand, when philosophers in the ideal city are in a position to lie about Forms, there is no indication that they would do so. Simpson (1996) thinks that it would be incoherent for Plato to think they would lie at all; this picture will be complicated, however, by *doxa*'s status with regard to truth and falsity.

poetry, primarily in Homer and Hesiod, should be prohibited for this reason. Instead, we must "ensure that the first stories [children] hear about virtue are the best ones for them to hear" (378e2-3).[108] Socrates has thus mentioned two areas of concern with regard to these stories: divinity and virtue.[109]

Though Socrates' attack on the traditional poets may at first seem to be advocating for telling lies about the gods, full stop, he maintains that our depictions of the gods must represent things that are *true* about them. Socrates' starting principle is that "however the god happens to be, one must always describe him as such," not just how we might like to represent him (379a7-8).[110] The first guiding truth about the divine is that "the god is really good and must be spoken about thus" (379b1). A god should be represented as good not simply because it will educate the young well but because the god is, in fact, good. This premise leads Socrates to make an argument that a god cannot do any harm and so cannot be the cause of anything bad, necessitating the exclusion of a number of traditional stories. A second major point about the gods is that "it is impossible for a god to want to change himself, but as it seems, since he is as

---

[108] Because of the many examples of ancient poetry that would not be allowed in the city, this passage has been cause for no small amount of concern over the centuries that Plato wants to get rid of poetry and censor the arts. I think, as will become clear presently, that this concern has clouded the larger purpose of the passage: to make arguments about the god(s) and what we must and must not believe about divinity. Though I do not have the space to discuss the worries around Plato's treatment of poetry and the arts, I do think that remembering the context of educating very young citizens can be helpful in alleviating some of these worries. Socrates recognizes, for example, that young people cannot easily separate allegory from history; furthermore, beliefs learned while young tend to hold fast. Socrates does not dismiss the arts; rather, he recognizes how much power they have.

[109] Here I am using divinity as shorthand for what has to do with gods and heroes insofar as the latter are divine. In the Greek tradition, heroes are divine, usually because they are the children of gods. Though one might argue that they are not fully divine by virtue of their share in humanity, Socrates appears to treat their close link to the gods as sufficient for applying to them at least many of the attributes he applies to the god(s), though there are certainly key differences. In any case, divinity as a whole, whatever its extension may be, is the focus of the present argument.

[110] Socrates does not make a strong statement on whether there is one god or many, and I do not intend to discuss the matter further here. I do find it interesting that at 379b1, when stating that the god is really good, Socrates opts for the singular ὁ θεός rather than the plural. Reeve in a note on his translation comments that "the definite article is almost certainly functioning as a universal quantifier, as in 'The swallow is a migratory bird,' which means (all) swallows migrate" (59). Given the prevalence of the singular in these passages and the points about the inability of a god to be altered, I would not be so certain that Socrates thinks there are multiple gods. What is certain is that he thinks the god or gods is/are nothing like the gods of the ancient tradition and Homeric epic, and that is enough for our purposes.

beautiful and good as possible, each of them will always remain in his own form" (381c6-8).

Socrates reaches this principle through an argument stemming from a god's goodness, and he uses it to present the discussion on lying that we have already examined. A god, who is good and changeless, would not want to lie and would have no reason to do so.

When we reach the discussion on lying about *ta onta*, we are thinking about how to educate the citizens on divinity so that they are well brought up. Socrates deems these truths vitally important to get right, both because they concern matters of extreme importance and (relatedly) because they are essential to good ethical development. What this context and background indicate is that truths about the gods are to be included in *ta onta*. I do not think that the Form view nor the ethical truth view as they stand can accommodate divinity as part of *ta onta*. Gods are not Forms, nor are they ethical facts of the sort laid out above. They do, however, have things in common with both of them. Like the Forms, the gods are unchanging and highly metaphysically important.[111] Like ethical truths, having the right conception of the gods is essential to forming virtue in the souls of citizens. Truths about *ta onta*, then, concern the highest realities and also develop our souls in the right direction.

With these points in mind, I propose that we adopt a more expansive reading of truths about *ta onta*: a truth about *ta onta* is something that turns one's soul towards the most real things.[112] The language of turning is something that Socrates often employs in the context of the

---

[111] I do not think we have to go so far as to conclude a Plotinian picture of the gods here, in which God is the Form of the Good, for which there seems to be scant evidence in the *Republic*. What I do think we can say is that the gods are included among the highest realities, being good and unchanging, and having the right conceptions of them will turn our souls closer towards understanding both divinity and the Good, whatever the relationship between the two might be.

[112] What sort of "degrees" of reality must we have in mind if we speak of the most real things? Allen (1961) considers Plato to have an understanding of degrees of reality insofar as sense-perceptible things serve as copies of Forms. Vlastos (1965) considers what is "most real" to mean cognitively dependable and undeceiving, since thinking of less real things as "half-existing" would be absurd. I would maintain that all we need posit is that there is a realm of things that are completely unchanging and eternal, and these are the things to which our souls should orient. These features are often discussed in terms of being fully real, but one should not worry unduly about it.

soul's development and education. Too much of an attachment to bodily pleasures, for instance,

"turns the soul's sight downwards," and one should "get free from those things and turn around

towards true ones" (519b2-4). Our souls are able to turn in both directions; a falsehood

concerning *ta onta* would turn our souls further away, in the wrong direction. The direction of

our soul is not merely a descriptive feature; it also carries normative weight, in that we *should*

turn towards these things.[113] Truths that pertain to the Good, for example, will draw our souls

closer to what is fully good and thus develop our souls in a good way.

These truths may very well include the ethical truths mentioned by Baima, but many

other things would also be candidates for this kind of truth. These truths would also, in some

sense, concern the Forms, which are certainly included among the most real things. The

important point, however, is that these are not, strictly speaking, truths *about* the Forms. One

need not have a grasp of the Forms—indeed, the youths being educated certainly do not have

such a grasp—or even mention the Forms at all to have this sort of truth. What this truth does

instead is draw our souls closer to and turn them in the direction of the Forms. Let us call these

truths ontological truths.[114]

Doxai and the Truth

This account of ontological truth needs further clarity. We have spoken about what is

contained in stories told to youths as "truths" and "something," but what exactly is this

"something"? Recall from our earlier discussion of the Form view that we hesitated to say that

Cooper (1986) promotes a similar deflationary reading that does not, in my view, thereby give up on Two Worlds in any way.

[113] Vogt (2012) also connects the language of turning to the stories told by poets: the stories are in some sense true because they turn the souls of children towards intelligible things.

[114] Szaif (2007) also uses the term "ontological truth" and draws some comparisons with assertible truth. As far as I can understand his account, he takes ontological truth to refer largely to the Forms in general, and he is certainly on the right track by distinguishing truth along these lines. However, as we shall see presently, ontological truth is not merely reserved for the Forms but pertains to *doxa* and is important in evaluating its merits (which Szaif considers to be few and far between, a bit too pessimistically).

these truths were about the Forms because Book III described having the truth as *doxazein ta onta*. We are now in a position to take this description to heart: the things with ontological truth are *doxai*. Let us now explore further what this means and how it relates to the other kind of truth mentioned in Book II.

To start, we find further evidence that *doxai* are at issue in the context of the passages about truth and lying. The stories that Socrates is proposing be told to citizens concern things that the gods did or did not do, the actions of heroes, the founding of the ideal city—things that happen or happened in the world. None of these involves the sort of cognitive grasp that would be characteristic of *epistēmē*; as I noted above, the focus of these stories is primarily on young citizens, who certainly have not undergone the required process of education to gain *epistēmē*. Socrates notes explicitly that the goal is making sure citizens have the right *doxai*, since if "we allow children to hear any stories whatsoever," they will "take *doxai* in their souls that are for the most part the opposite of those that we think they should have when they grow up" (377b4-7). Turning our souls towards the highest realities involves acquiring *doxai* that put them in the right condition.

Let us take an example of such a *doxa*. Socrates is clear that the Homeric epics, though occasionally imparting some good lessons, are filled with stories that will not develop our souls in the right way. Throughout the *Iliad*, for example, Achilles—a divine hero who descended from the gods on both sides of his family—displays all sorts of faults and vices. Taking into our souls that Achilles in the *Iliad* was a model of virtuous behavior will leave us with many falsehoods about divinity and will not develop our souls in the proper way. If we are to believe that Achilles was divine, then we should believe that he acted in a way consistent with the divine qualities laid out above. As someone fully good—like the gods—he would not quarrel with

other divine beings. A *doxa* with ontological truth about Achilles could thus be, "Achilles never quarreled with the gods."

We can see how this would be a *doxa*: it involves how a human (though admittedly a divine one) acted while in the world. We can also see how accepting such a *doxa* would give us the right conception of divinity. But is this, in fact, how Achilles acted? Socrates admonishes Homer at 391a-b for saying that Achilles threatened to injure Apollo and tried to fight a river god. The stories told by Homer and Socrates' proposed alternatives both fall into category (3) of falsehoods in words: stories of long-ago events where we cannot know precisely what happened. There are several things we do not know about Achilles: whether he even existed, whether he was the child of gods, whether he quarreled with the gods, whether he did the other noble and shameful things attributed to him. When we say, "Achilles never quarreled with the gods," we may be saying something that is a falsehood in words.

To be very clear, the reason this *doxa* would be false is not that Homer's account is right. Socrates rules out Homer's Achilles on philosophical grounds: because the gods must be perfectly good, and something perfectly good will produce only good things, the children of gods would not be full of vice. Poets such as Homer must "say either that these deeds were not [the heroes'] or that they were not the children of gods" (391d5-6). What this does not rule out, however, is that Achilles never existed, or that someone named Achilles existed but was fully human and acted largely as Homer tells it. In those cases, "Achilles never quarreled with the gods" is false (or, if we are worried about the referent, "There was once a hero named Achilles who never quarreled with the gods" is false). Even if Achilles the hero never existed, however, a story about his heroic deeds and piety towards the gods could very well turn our souls towards

the highest realities. The *doxa* would still have ontological truth, even if it is false in some other way.

I will call this other type of truth, the one that corresponds to lying in words, particular truth. Recall that one key difference between true lies and lies in words is that the gods would never lie in words, but humans would sometimes have occasion to do so. Socrates emphasizes that it is not just that the gods would be unable to lie in these circumstances; the gods would never find themselves in such circumstances at all. Occasions like not knowing the details of ancient events arise only in the human world, where things are constantly changing and we can never get a secure grasp on them. Particular truth concerns whether a *doxa* maps onto what is happening or has happened in the world. The conditions for particular truth and ontological truth are different, and it is possible to have the latter without the former.[115]

When we are faced with some *doxa*, we might ask whether we should *doxazein* it, admitting it into our souls. This task is not, we can see, a simple matter of determining whether the *doxa* is true or false. For one, a *doxa* might be true for one truth type, both types, or neither. Furthermore, once we have dispensed with a strict binary for truth, we might wonder whether each type of truth is itself an all-or-nothing affair. We may think, for instance, that some *doxai* do a better job of turning our souls than others, even though both may have ontological truth to some extent. We shall see in the following section that particular truth will also be variable, but even if it were not, Socrates puts greater importance on ontological truth. Determining whether a *doxa* is to be accepted or not, then, is quite a complex endeavor.

---

[115] Is it possible to have particular truth without ontological truth? It certainly seems likely that some *doxai* have particular truth and are neutral regarding ontological truth—e.g. "the sky looks very blue today." As I noted above, there do not seem to be *doxai* about the divine that would be particularly true but ontologically false. Things get murkier in other areas, such as the founding of cities; "our ancestors once lived in discord" could be one example. But as we shall see in a moment, speaking simply of ontological and particular truth does not capture the full complexity of the picture.

Two necessary conditions for making such a determination come to mind. The first is to have a grasp of the highest realities themselves. If one is to discern whether a *doxa* will turn a soul towards *ta onta*, one must already know *ta onta*. The second is to be able to conclude how much truth and falsity a *doxa* has for each type, in order to deliver a final verdict on whether a *doxa* is sufficiently true. It seems that the path towards reaching a verdict on a *doxa* would proceed in the following way. First, determine how much ontological truth and falsity the *doxa* has. Second, do the same for particular truth and falsity. Finally, compare these results to each other to conclude whether the *doxa*'s truth outweighs its falsity and by how much. We need not ourselves determine now what a *doxa*'s standard of truth must be. Perhaps we just need its truth to outweigh its falsity. Perhaps we need an even higher standard in order for it to be a *doxa* worth having (it seems doubtful that a lower standard would suffice).[116] The point is that someone who grasps the highest realities will be able to determine, through weighing *doxai*, what *doxai* are worth acquiring with these realities in view.

These conditions probably sound familiar, and with good reason: they mirror the process of the measuring *epistēmē* that we saw laid out in the *Protagoras*. The *Republic* has shown us what we need in order to measure—a cognitive grasp of the most real things, *ta onta* as we have come to understand them metaphysically. Combined with the *Protagoras*, we have a preliminary picture of why grasping *ta onta* is important: it allows us to choose the right *doxai* to have in our souls. In order to grasp the highest realities, we must have *epistēmē*, or the highest cognition of the highest things. *Epistēmē* is crucial for measuring the truth of *doxa*.[117]

---

[116] Brickhouse and Smith (1983), though not describing the measurement of truth, think that the option of rulers telling citizens something that has both types of truth (as they conceive of them) is preferable, and only in unfortunate cases do we need to sacrifice one for the other. This line of thinking, I expect, would be right in the cases where we can choose among *doxai* to accept; whatever our threshold for *doxazein* must be, if we are choosing among similar *doxai* we should choose the one that ranks highest on our scale.
[117] It may be important to emphasize here that we need not take *epistēmē* itself as what measures the truth of *doxa* (although I have not entirely ruled that out). Rather, *epistēmē* is an essential feature of the process of measurement,

I have proposed the following picture: the truths of *doxai* are measurable, and *epistēmē* is uniquely suited to allow us to measure it. One big advantage of this picture is that is shows more concretely how *epistēmē* can be of important practical benefit along the lines of what we saw in the *Protagoras*. The *Protagoras* gave us a general account of a measuring *epistēmē*; now we are able to see more clearly how it functions and is beneficial. Having the right orientation to *ta onta* will develop us well, so the ability to measure truth is an important component of our cognitive and ethical development.

Nevertheless, three major questions remain. First, we still have the Problem of the Gap. It would certainly be nice if *epistēmē* could have this effect on *doxa*, but aren't the two still set over completely different types of objects? Second, we have thus far focused on Socrates' description of truth and *doxa*, and the role of *epistēmē* has been largely speculative. What evidence do we have that Plato was thinking of *epistēmē* in these terms? Third, even if *epistēmē* could measure truth and Plato conceived of it as doing so, the most that seems to get us is that *epistēmē* would *allow* our souls to develop well. But recall that we are searching for how *epistēmē* would *guarantee* the Socratic paradoxes. We would want *epistēmē* to lead to virtuous action necessarily, not just potentially—can Plato's account get us our desired conclusion?

Two Worlds and the Measuring *Epistēmē*

In the remainder of this chapter, I wish to answer these important questions by turning to Books V through VII (with a couple of brief detours along the way), in which Socrates directly tackles *epistēmē*, *doxa*, and the metaphysics that accompanies them. What Socrates lays out in these middle books will show us three things. First, *epistēmē* can measure the truth of *doxa*

---

to the extent that it would not be beyond the pale to speak of a measuring *epistēmē*. What ends up using *epistēmē* to measure, however, might be a different sort of cognitive power or activity.

because truth is a feature they have in common: *doxa* is akin to *epistēmē* because of its share in truth but it is never fully true, even regarding particular truth. Second, mastering calculation is essential to acquiring *epistēmē* and using the best part of our souls, and *epistēmē* presents a model that a philosopher uses to establish truth and virtue in a city. Third, acquiring *epistēmē* is such a soul-altering event that virtue necessarily follows along with it, and the education its measurement provides develops characters unfailingly in the right way. All of these point towards a picture of *epistēmē* as something that measures and imparts truth. In doing so, we see how *epistēmē* improves our *doxai*—thus solving the Problem of the Gap—and guarantees virtuous action—thus satisfying the Socratic paradoxes.

Socrates distinguishes *doxa* and *epistēmē* most explicitly in two places: first, in describing the philosopher at the end of Book V, and second, in the analogy of the Divided Line at the end of Book VI. Let us take these in reverse order and look first at the Divided Line, in which Socrates separates out four cognitive states that correspond to four kinds of objects. *Doxa* concerns "the animals around us and everything planted and the whole category of artifacts," i.e. sense-perceptible objects in the world (510a5-6). *Epistēmē* concerns "what reason itself grasps through the power of dialectic, making hypotheses not starting points but a hypothesis as it is in reality," i.e. the Forms (511b3-4).[118] *Epistēmē* is a loftier and clearer state than *doxa*, owing to the difference in their objects. The reason for *epistēmē*'s higher clarity is not just that the objects are different but the precise nature of the difference: "inasmuch as the objects that [the different

---

[118] It is worth noting that Socrates does not use the term *epistēmē* in this instance but instead νόησις, which could worry us if we are considering his treatment of *epistēmē*. Nevertheless, we have very good evidence that the conception of *epistēmē* we have been following is what is at issue here. In Book VII, Socrates recalls the Divided Line and claims that they should "call the first part *epistēmē*, just as before" (533e3-4). It is perhaps a strange error that Socrates said he called the part *epistēmē* before when he did not, but that should not concern us here; what is clear enough is that the use of νόησις in Book VI and *epistēmē* elsewhere are tracking the same thing. Indeed, it is a couple of lines earlier that Socrates notes not to worry too much about the precise names. On tracking the term "*epistēmē*," see the Introduction.

states] are set over have a share of truth, so does one find these [states] to have a share of clarity" (511e2-4). Truth is one thing that distinguishes the objects of *doxa* from the objects of *epistēmē*: the former have less of a share and the latter have more.

The different sections of the Divided Line are thus separated by the amount of truth each one has. This indicates further that *doxai* are not fully true, a point to which we will return below. What I take to be the crucial upshot of the Divided Line, however, is that the differing degrees of truth provide not only a means of separating *doxa* and *epistēmē*: they also provide a standard of comparison between the two. Although the objects of *doxa* and of *epistēmē* do not overlap, they have a feature in common: both of them possess some amount of truth.[119] Socrates makes clear that the cognitive powers and their objects can both be compared to each other in this way: "would you wish to say, in dividing up by truth and falsity, that as the believable (*doxaston*) is to the knowable (*gnōston*), so the thing likened is to what it is like?" (510a8-10).

If we are looking for a way in which *epistēmē* can affect *doxa*, the fact that both of them share a feature is a telling sign. We may very well accept that the objects of *doxa* and the objects of *epistēmē* do not overlap whatsoever and are completely separate from each other, a fact that I would take to be sufficient to adopt a Two Worlds view of Plato's epistemology. We need not take the further step, however, of concluding that the objects of *doxa* and *epistēmē* are completely different. The image of the Divided Line suggests that there is something they both share—a relationship to the truth. If *epistēmē* has the greatest share in truth, and *doxa* has a lesser but variable share, then *epistēmē*'s grasp of the truth could affect and improve *doxa*'s.

The Divided Line tells us that *doxa* has a lesser share of truth, and Socrates' distinction between ontological and particular truth provides us with some evidence that truth is variable

---

[119] Socrates' account of the Divided Line thus indicates that *doxa* have truth to some degree, *pace* Ketchum (1987), who argues that truth is to be found only in the Forms.

110

even within the realm of *doxa*. I maintain that we see further evidence of *doxa*'s condition in his

other comparison of *doxa* and *epistēmē* from the middle books. At the end of Book V, Socrates

distinguishes between philosophers and the "lovers of sights and sounds" (476b4). The former

desire to gain *epistēmē*, which Socrates names as the power that is set over *to pantelōs on*, what

completely is. They seek to know the Beautiful itself, the Form of Beauty. In contrast, the

lovers of sights and sounds pursue many beautiful things but have no thought or desire for the

Beautiful itself.[120] Socrates names the many beautiful things as comprising a separate category

from *to pantelōs on*; they are instead "what both is and is not" (ἔχει ὡς εἶναί τε καὶ μὴ εἶναι,

477a6). Socrates explains how things that the lovers of sights and sounds consider beautiful are

different from the Beautiful itself:

> Socrates: "My very good man," we will say [to the lover of sights
> and sounds], "will there appear any one of these many beautiful
> things that is not ugly? And of the just things, any that is not
> unjust? And of the pious things, any that is not impious?"
> Glaucon: No, but it is necessary that somehow the same beautiful
> things will also appear ugly, and similarly for the other things you
> asked about. (479a5-b1)

Socrates proceeds to assign these things as the objects of *doxa*, which has been shown to deal

with both what is and what is not.

What we should take from this interestingly contradictory phrase is that there is no case

where "X is beautiful," where X is not the Beautiful itself, is true. There will not be any

beautiful object of *doxa* that is not also ugly. This phenomenon does not apply just to things like

beauty, which one could argue are too normative or subjective to be fully realizable to begin

with. Socrates also names halves and doubles as things in the world that would both be and not

---

[120] I believe that the distinction between *doxa* and *epistēmē* that is important here can be made without ruling on whether the many beautiful things are beautiful particulars or beautiful kinds or types. See Gosling (1960) and Buckels (2018).

be. The objects of *doxa*, the things in the sense-perceptible world, are rolling around between

being and non-being, as Socrates says. Accordingly, the *doxai* we accept into our soul will be a

in a similar condition: they will never, or at least very rarely, have full particular truth,

representing the world as it actually is.[121] Instead, they will be *both* true and false, and some will

have more particular truth than others.[122]

Before we proceed from this point, however, I believe it is important to get clear about

what this interpretation does and does not have to say about some of the more contested aspects

of this passage. The end of *Republic* V is the primary locus of concern in the debates about

whether to accept a Two Worlds reading of *doxa* and *epistēmē*. Socrates distinguishes between

epistēmē being set over "what completely is" and *doxa* being set over "what both is and is not."

One might take these uses of "is" in several different ways. Three different readings of "X is…"

have typically been proposed: (1) an *existential* reading that "X exists," (2) a *predicative* reading

that "X is F," and (3) a *veridical* reading that "X is true."[123] The question of the sense of "is"

relates to the question of whether *propositions* or *objects* are at issue here, regarding both *doxa*

and *epistēmē*. A veridical reading tends to take a "propositions" view, in which the difference

between *doxa* and *epistēmē* is that the former deals with a set of propositions that are true and

---

[121] One might wonder whether we could construct a *doxa* that is so specific, or so tentative, that it would have full particular truth. Examples could be, "this painting appears beautiful in these respects at this point in time," or "either it is the case or it is not the case that Helen is fully beautiful." These examples would lead me to conclude that we can find a select few cases where there is nothing particularly false about a *doxa*. They are, however, not very interesting and very rare. Even the first example is likely to fall victim to some degree of falsity. Schwab (2016) notes that our status in the world means that we can never know the full merits and drawbacks around an individual *doxa* (in his example, a law to be enacted): "given the messiness, complexity, and interconnectedness of concrete matters in the perceptible world, seemingly remote factors can affect the appropriateness of applying a predicate to a particular perceptible object. Quite often, indeed perhaps always, the full set of circumstances will elude any actual philosopher ruler (or even a whole cadre of rulers)" (68). Woolf (2013) also takes this feature to show that we can never have knowledge of perceptibles.
[122] The language of "rolling around" indicates that *doxai* are not fixed regarding their particular truth, giving us further evidence of the importance of measurability.
[123] See Fine (1978), 124.

false, and the latter deals with propositions that are only true.[124] Predicative and existential

readings tend to take an "objects" view, in which Socrates' aim is to show that *epistēmē* concerns

the Forms and *doxa* concerns sense-perceptibles themselves, rather than propositions about

them.[125] A propositional view would typically reject Two Worlds, while an objects view would

embrace it.

The first thing to say is that although I have been speaking extensively about truth, I do

not think that confines me to a veridical reading.[126] What I have been chiefly concerned with

regarding truth and falsity is the measurability of our *doxai*, and by extension how acquiring

epistēmē can benefit *doxa* through measurement. The *doxai* that are measured and benefitted are

clearly propositional judgments of some kind, and elsewhere Socrates explicitly mentions

thinking of *doxai* in this way.[127] They are thus what we might think of as the *doxai that one*

accepts into one's soul, which I take to be importantly different from what we might call the

objects of *doxa*. Both of these, however, are at issue here: a *doxa* in our soul contains content

about an object of *doxa*. As with most Two Worlds-ers, I can accordingly take this passage to be

---

[124] See Fine (1978, 1990). Fine is very clear that she does not adopt a "degrees of truth" view, in which *doxai* can be both true and false or something else outside the strict binary; rather, some *doxai* are true and some are false. Annas (1981) deems degrees of truth as silly as degrees of existence, though I hope to have given it much greater plausibility so far.

[125] I believe that it has been sufficiently argued that the existential reading does not have much power on its own and should either be abandoned or adopted as part of the predicative reading, so I will treat them here as a unit. See Gonzalez (1996), as well as Sedley (2007) against an existential reading in Greek in general. Cross and Woozley (1964) take Plato to mix the two though fault him for doing so.

[126] A full discussion of the drawbacks of the (purely) veridical reading is not necessary here, but just to say a few words: one primary point usually mentioned in favor of this reading is that it meets a "dialectical requirement" that Socrates must argue in a way understandable and acceptable to the lover of sights and sounds. See Gosling (1968), Fine (1978, 1990), and Stokes (1992), the last of whom takes this requirement to show that no serious conclusions are meant here at all. I believe that it has been shown adequately that the dialectical requirement is no special help to the veridical reading; Gonzalez (1996) argues this point especially forcefully, and see also Baltzly (1997) and Vogt (2012). One of the other major motivations for a veridical reading is that it avoids the Problem of the Gap, but I hope that my reading can deflate those worries as well.

[127] See *Theaetetus* 190a: "And when [the soul] arrives at something definite, either by a gradual process or a sudden leap, when it affirms one thing consistently and without divided counsel, we call this its *doxa*. So, in my view, to judge is *doxazein*, and a *doxa* is a statement which is not addressed to another person or spoken aloud, but silently addressed to oneself."

speaking about the status of objects of *doxa* and their difference from the Forms. What I take to be the upshot for our purposes, however, is an insight about the *doxai* that we *doxazein*, which are propositions. I see this reading as predicative primarily but also importantly connected to the veridical reading.[128]

We can thus connect this reading to our discussion of particular truth. The relationship would be something along these lines: the fact that a certain object of *doxa* X is both F and not-F means that (as in, we can conclude that) the proposition "X is F" is both (particularly) true and false. This does not mean that the proposition is equally true and false, or that all propositions have an equal share of particular truth. We might take the lovers of sights and sounds to have *doxa* with significant particular truth.[129] Even though "Helen is beautiful" and "this piece of mud is beautiful" are both not fully true, we can conclude on this reading that the former has more particular truth than the latter. We cannot immediately conclude that we should *doxazein* one and not the other, for we still have ontological truth and a more complex weighing process to worry about. What Book V shows us, however, is that the status of the objects of *doxa* reveals content about them to be measurable regarding particular truth.[130]

In sum, the picture of *doxa* in Books V and VI cements two things: *doxa* and *epistēmē* are alike in having a share of truth, and *doxai* are never fully true. We can thus answer the first

---

[128] In this way, I am largely following Baltzly (1997) in taking both the predicative and veridical readings to be at issue and importantly connected, though his focus is more on the side of *epistēmē* than *doxa*. Baltzly argues that knowing an object of *epistēmē* involves grasping a *logos* that is completely true, while any *logos* for *doxa* would be no more true than false. I do not follow him on the neutrality of *doxa*'s truth value, but I generally agree with his picture of the relationship between the contents and objects of *doxa* and *epistēmē*. Cooper (1986) also combines the different readings in this vein. Taylor (2008) thinks that Plato did not hold there to be a strict difference between objects and propositions at all.

[129] This would be one way to understand Socrates' claim that they are "like philosophers" (475e2). It is certainly possible that this reading is too optimistic and they have no special insights whatsoever; the point still stands in any case.

[130] I take it that the content/object relation will also apply to the Forms as objects of *epistēmē*, though whether our *epistēmē* of the Forms is strictly propositional is not something we need concern ourselves with here. On our cognitive grasp of the Forms (especially the Good), see Gerson (2003), Gentzler (2005), and Butler (2006).

question of solving the Problem of the Gap. *Epistēmē* and *doxa*'s similarity with regard to truth shows us that *epistēmē* would be able to measure the truth of *doxa*. The differences between *doxa* and *epistēmē* show us that *doxa* is able to be measured with regard to its truth. I believe we can conclude that *epistēmē* and *doxa* can have this relationship. Let us turn to considering some evidence that they do.

Epistēmē, Calculation, and Modeling

As Plato develops his philosophy, his conception of *epistēmē* develops into something so robust that it deals only with what completely is and is in no way deficient. The process of acquiring *epistēmē* is neither short nor easy: it takes fifty years (more or less) of different kinds of instruction before one can have the best and purest cognition. But the process is worth it: those who achieve *epistēmē* are uniquely suited to be leaders of the ideal city and develop the souls of its citizens in virtue. I believe that Plato's description of the parts of the soul and his re-examination of the philosophers' education makes clear that he conceived of *epistēmē* as something that essentially involved measurement. We will see this through his emphasis on the role of calculation: calculation is essential to the best part of the soul, as well as to the process of putting it in the best condition.

Let us turn briefly from the middle books to Book IV. There, when Socrates has finished laying out the creation of the ideal city, he compares the finished product to the individual soul in order to identify its parts and best states. There are three different parts of the soul, but one is the best and would most properly rule the other two: the *logistikon*, which is commonly translated as the rational part.[131] What this translation overlooks, however, is that there is a slightly different word in Greek, *logikon*, that is a more fitting term for reason or rationality in general.

---

[131] I am grateful to Jessica Moss for bringing up the importance of the term *logistikon* to me. See also Moss (2008).

Logistikon is more specific: it normally refers to calculation or computation of some sort. Plato explicitly flags this use in a different context in Book I, in an outburst from Thrasymachus: "whoever makes an error in calculating [*logismō*], do you call a calculator [*logistikon*], whenever he makes an error and with respect to that error?" (340d4-6). When the term returns in Book IV in conjunction with the soul, we have been primed to think of it along these lines, and we would be remiss not to think of calculation as important, if not essential, to the best part of the soul.[132]

In Book VII, Socrates lays out how a philosopher would cultivate this part of her soul through the extensive process of education. We are by now well aware that we humans dwell in the world of *doxa* and becoming, in which there is no real stability. If we want to ascend to higher states of cognition, it is necessary to find something that will "drag the soul from what becomes to what is" (521d3-4). The first step in doing so is to find the one thing that "all *technai*, thought-processes, and *epistēmai* use," which is necessary for further cognitive ascent (522c1-2). We may not be surprised to find out that it is "number and calculation [*arithmon te kai logismon*]" (522c6-7). The function of calculation in the philosopher's education is to bring the philosopher to consider Being, rather than just the ever-changing sense-perceptible world. In doing so, what the philosopher comes to comprehend are not just the numbers and apparent objects of calculation but something much more important: "this subject really does seem necessary for us, since it appears to compel the soul to use thought itself [*noēsis*] on the truth itself" (526a8-b3). Learning calculation is learning about truth itself; the path to *epistēmē* involves grasping an art of measurement and developing a deeper understanding of the truth.[133]

---

[132] It is important not to read the calculating part of the soul as merely a number-crunching device of some sort. While Plato's divisions may tempt us to think that the calculative part cannot have its own desires and motivational power, we should be careful not to conclude that there is such a separation. See Büttner (2006) for the "non-cognitive" aspects of the *logistikon* and Lorenz (2006) for some more general views of rational motivation.
[133] A full discussion of mathematics in Plato will take us too far afield, but see Burnyeat (2000) for possible connections between mathematics and the Form of the Good, including the soul-changing aspects of studying mathematics. Mueller (1992) points to outside evidence of Plato advocating for "metrologia" or a theory of

We saw earlier what *epistēmē* would be in a position to measure. Socrates' emphasis on the *logistikon* now makes it clear that he indeed conceives of *epistēmē* as measuring along these lines. Seeing *epistēmē* in this light helps make sense of how *epistēmē* is beneficial both for those who acquire it—the philosopher-rulers who lead the ideal city—and for others who benefit as a result—their citizens. At the beginning of Book VI, Socrates describes how those with *epistēmē* are the only ones who are really fit to rule. He denounces those without *epistēmē* as "blind," because they "have no clear model [of *ta onta*] in their soul and are not able, like a painter looking upon what is most true and always referring to it and beholding it most precisely, thus to establish here customs about beautiful and just and good things, if it is necessary to establish them, and to save and guard what has been established" (484c4-d2). Looking at "what is most true" is essential for establishing the best possible city on earth; a philosopher must establish things that are as close to the truth of *ta onta* as possible.

The philosopher's work is not mainly focused on just making laws or setting out decrees that contain the most truth. Rather, the work is at the level of souls, in "establishing what he sees there [in the realm of *ta onta*] into the characters of people both in private and in public, and not only forming his own" (500d5-7). By grasping the highest realities, the philosopher-ruler establishes the best *doxai* in the souls of both herself and her citizens. Plato tells us here something important about the measuring *epistēmē* that hearkens back to the "noble" *epistēmē* of the *Protagoras*: the measuring *epistēmē* will not be possessed by everyone. If that were the case, everyone would need to undergo the same process of education, an infeasible requirement given the time and difficulty involved. Rather, the measuring *epistēmē* is needed only by whoever is responsible for shaping the character of others.

---

measurement within the Academy (the word appears nowhere else and would seem to be unique to Plato). See also Sedley (2007).

We can see, then, that Plato connects *epistēmē* and the best reasoning strongly to calculation. Furthermore, this picture makes good sense of the role of the philosopher-ruler. Though more work is likely needed to discern the specifics of how the philosopher employs the measuring *epistēmē*, I hope to have shown there to be good evidence that Plato has such a view in mind in the *Republic*. We can now respond to the second question: not only is *epistēmē* able to measure a measurable *doxa*, but we can also conclude that Plato thought it did so.

The Socratic Paradoxes and the Measuring *Epistēmē*, Again

We are now left with the third question, the one that has been running in the background throughout our project: can this notion of *epistēmē* support the Socratic paradoxes? In order for it to guarantee SP1 and SP2, it appears that acquiring *epistēmē* must have some necessary connection to virtuous action, to the extent that voluntary wrongdoing is impossible. In this final section, we shall see that *epistēmē* produces virtuous action in two ways: someone who gains *epistēmē* loves virtue alone as a result of nature and mixing with the Forms, and those on whom the measuring *epistēmē* is used will be virtuous as well.

In emphasizing the importance of the philosopher's grasp of *epistēmē* and truth, Socrates shows how the right attitude towards the truth leads to a life of virtue. The philosopher who comes to grasp *ta onta* will never be beset by vice in her soul:

> Socrates: Will we not reasonably argue that a real lover of
> learning by nature strives towards what is [*to on*]? [. . .] when he
> approaches and mixes with what really is, begetting understanding
> [*nous*] and truth, will he know and truly live and fully mature and
> cease his labor pains, but not until then?
> Adeimantus: That is the most reasonable of all.
> Socrates: Well? Will it be part of his nature to love any falsehood
> or, quite the opposite, to hate it?
> Adeimantus: To hate it.
> Socrates: And with truth leading, I don't think we would ever say
> that a chorus of bad things will follow it.
> Adeimantus: How could it?

Socrates: Instead, a healthy and just character, which temperance
also follows.
Adeimantus: Absolutely. (490a8-c7)

There seem to be two elements at play regarding this connection. First, a philosopher is naturally inclined to love the truth. Though this nature can be corrupted and fail to ascend to the highest cognition, it naturally starts out on the right foot, so to speak, and inclines toward ontological truth especially. Second, interacting with the Forms is a soul-changing activity. As Socrates describes it, mixing with what truly is causes a philosopher truly to live and mature, and evils could never follow from that state.[134] Recall that a voluntary wrongdoer in the domain of virtue would have to have *epistēmē* of good and bad and still do the bad thing. Together, the features of the philosopher and her acquisition of *epistēmē* show that to be impossible.

The good character that comes from truth is not, however, unique to the philosopher but also manifests in those she rules. Though citizens without *epistēmē* would by default not be able to do wrong voluntarily, it is still noteworthy that these citizens would rarely do wrong even involuntarily. We have seen that one of the most important ways that a philosopher-ruler can form citizens is through education when they are young. An improper education can be so detrimental because of the profound effects on one's soul and conduct: "there is never, nor has there been, nor will there be a human character that differs regarding virtue from the way that it has been educated by these [improper] things" (492e3-5). Every soul raised poorly will act poorly, and every soul infused with truth will naturally be free of vices and act rightly. The truth that comes with *epistēmē* is so enlightening that it guarantees virtue to those who achieve it and implants *doxai* that adhere to it in others. When one's soul has been raised to have *doxai* with

---

[134] Burnyeat (2000) advocates for a sort of soul-assimilation in encountering the Forms: one takes on the structure of the things one studies. We do not need to subscribe fully to thinking of the Good in mathematical terms (though as we saw above, something like that is important to the picture) to think that the Forms would change our souls.

the most truth, the entire character of the person will be turned towards goodness. There will thus be a striking absence of anything that would lead such a person to do wrong. Acquiring *epistēmē* of this kind is no easy feat, and one might worry that the philosopher's task is nearly impossible. But Plato has at least found what he has been seeking since he set up the problem in the *Hippias Minor*: a conception of *epistēmē* that renders voluntary wrongdoing impossible and guarantees a life of virtue.

One final note about virtue and the measuring *epistēmē*: I have shown how *epistēmē* improves the city's *doxai* by calculating how they stand regarding truth. Even philosophers, however, dwell in the world of *doxa*, and though their ability to measure truth is extremely beneficial, they are still liable to make mistakes. I would argue that we can accordingly make good sense of how even the ideal city can fall from grace and turn into a timocracy, before degrading further into oligarchy, then democracy, then tyranny, as described in Books VIII and IX. Socrates claims that this process starts with mis-timing births within the city. One might doubt the importance thereof, but the point is that Socrates clearly puts the mistake in terms of calculation: "those who are wise, whom you educated for leading the city, will not bring [well-timed births] to bear through calculation with sense-perception, but will pass over [the right times] and bear children when they shouldn't" (546b1-4). When one is working with sense-perceptible things, even the wise will fall into error. Though a city ruled by philosophers will be better than any alternative, there is still room for bad things—brought about involuntarily, but bad nonetheless—to creep in. *Epistēmē* can render a city as good as we can hope for, but it cannot render a city perfect.

Conclusions

In the *Republic*, we have reached the culmination of Plato's search for understanding the Socratic paradoxes and what they demand on his epistemology. His project of finding a robust notion of *epistēmē* has resulted in the metaphysically-infused picture presented by the Two Worlds theory, in which *epistēmē* concerns only the most real and true objects. The importance of this picture of *epistēmē* for the Socratic paradoxes is that it shows the function of *epistēmē* through revealing the Socratic paradoxes' new focus: no one harbors falsehood voluntarily. Socrates' picture of truth and falsity is complex and involves grasping two kinds of truth: ontological truth, which turns our souls towards *ta onta*, and particular truth, which maps onto what happens in our changing world. To have true *doxai* is to have truth on the whole, putting both of these types of truth under consideration and determining how much of each a *doxa* has.

The need to determine how much truth a *doxa* has shows us the role of *epistēmē*: having *epistēmē* of *ta onta* is what provides us with the means to measure truth. Plato separates *epistēmē* and *doxa* according to their share of truth, which provides us with an important insight: truth is something the two have in common, and accordingly one can influence the other on these grounds. *Doxa* by its nature is not fully true or false; *epistēmē* by its nature involves mastering calculation and being used as a model to establish the right *doxai* on earth. We thus can understand the measuring *epistēmē* from the *Protagoras* in its capacity for guaranteeing good action and causing us to live well: *epistēmē* measures truth and in doing so causes us to have the *doxai* that create a virtuous character and a good life.

We have seen, then, how Plato's treatment of the Socratic paradoxes has deep roots in the development of his philosophy as a whole, and in his epistemology in particular. He feels drawn to commit to the Socratic paradoxes, which precipitates further changes that hope to

121

accommodate them. If one has a different attitude towards the paradoxes, however, then we might expect differences in one's epistemology as well. It is with this in mind that we will now consider a contrasting account from Aristotle, one that responds to and reconsiders the Socratic paradoxes, in order to consider their legacy after Plato.

Chapter IV: The *Eudemian Ethics* on Why Virtue Is Not *Epistēmē*

Introduction

In this final chapter, we will be turning to Aristotle's *Eudemian Ethics*. Compared to the

Nicomachean Ethics, the *Eudemian Ethics* has traditionally garnered little attention, although the

last few decades have brought increased interest.[135] One reason for such interest is that the

Eudemian Ethics was likely written earlier than the *Nicomachean Ethics*, and so it represents an

earlier period of Aristotle's thinking that can be used to trace the development of his philosophy.

As a result, it is a helpful treatise for thinking about comparisons between Aristotle and Plato.

As a former student of Plato's, Aristotle would have been steeped in Plato's thought, including

the epistemological and ethical views that we have laid out above. In the earliest periods of his

thought, Plato's views and arguments would have been fresh in his mind. As Aristotle begins to

formulate his own views, we will see him responding to Plato and often trying to distance

himself from his former teacher.

I wish to focus on one such response by embarking upon a close examination of Book

VIII, Chapter 1, the first of three short chapters that conclude the treatise as we have it. VIII.1

has received very little attention in the literature. Most of the discussion surrounding it has been

philological in nature, and with good reason: the extant manuscripts are terribly corrupted, and

at least some emendations are required to make any sense of the text.[136] These many textual

---

[135] A primary motivation for renewed study of the *Eudemian Ethics* traces back to Kenny (1978), who uses statistical and philosophical means to argue that the *Eudemian Ethics*, rather than the *Nicomachean Ethics*, is the original source of the three books common to both treatises (an issue we will return to later), as well as the philosophically superior text. Although Kenny speculates that the *Nicomachean Ethics* is earlier than the *Eudemian Ethics*, he is not fully committed to this position and most scholars follow Jaeger's (1934) influential discussion of the timeline of Aristotle's ethics in putting the *Eudemian Ethics* first chronologically. Jaeger is also largely to credit for the prevailing view that the *Eudemian Ethics* is a work of Aristotle's, edited rather than written by his pupil Eudemus.

[136] For some of this work, see Jackson (1913), Mingay (1971), Moraux (1971), and Woods (1992), as well as the preface and notes in the Oxford Classical Text.

difficulties notwithstanding, the theme of the chapter is relatively clear. It chiefly considers whether the virtues, and *phronēsis*—practical wisdom—in particular, are forms of *epistēmē* and concludes that they are not. What is less clear is how Aristotle's argument establishes this conclusion and why he thinks it important to do so.

Here I hope to shed some light on the argument in VIII.1 by considering how Aristotle conceives of *epistēmē* in the chapter and how this conception departs from that of *phronēsis* and virtue. In particular, I argue that the notion of *epistēmē* Aristotle employs here is one that we will first examine in *Metaphysics* Θ: a contrary-use capacity that can be used to produce opposite states of affairs. The main thrust of Aristotle's argument in VIII.1 is first to point out some key features that identify *epistēmē* as a contrary-use capacity and then to show that *phronēsis* does not share these features and therefore cannot be *epistēmē*. As a result, we see Aristotle putting forward a conception of virtue and *epistēmē* that differs considerably from SP2. Virtue is explicitly not *epistēmē*, and this is because Aristotle denies SP1: someone with *epistēmē* can and does misuse it. *Epistēmē*, as a capacity, has the same sort of status as the voluntarily misused domains in the *Hippias Minor*, while *phronēsis* is what is strong and accompanied by virtue.

With this in mind, we might be tempted to conclude that Aristotle has dispensed with the Socratic paradoxes and modified his philosophy accordingly. Upon closer examination, however, we shall see how Aristotle also develops his philosophy in response to their pull. The argument of VIII.1 is surprising because if we look at the *Eudemian Ethics* as a whole, it is superfluous: in earlier chapters that appear in both the *Eudemian* and *Nicomachean Ethics*, Aristotle has argued much more simply that *phronēsis* is not *epistēmē*. What these arguments indicate together is that Aristotle is operating with two different conceptions of *epistēmē* in his

ethical thought, similar to Plato. This discovery will allow us to conclude two things. First, the simple arguments in the shared books belong in the *Nicomachean Ethics* and indicate the development of Aristotle's views on these matters. Second, this development indicates that Aristotle still felt the connection between knowledge and virtue, giving *phronēsis* a role that Plato held for *epistēmē* and designating *epistēmē* itself as a different kind of virtue.

Eudemian Ethics VIII.1: Overview and Puzzles

Let us start by giving a general sketch of *Eudemian Ethics* VIII.1 and its components. The chapter can be divided (unevenly) into two main parts.[137] The first, briefer part examines the difference between using something correctly and misusing it "otherwise" (1246a26-27). Aristotle posits two ways that one can use something "otherwise": one can misuse it by distorting its proper use, or one can make use of it in a way that does not count as a genuine use of that thing. For example, an eye can see properly, or it can mis-see "otherwise" (1246a28-9). Both the proper use and the first "otherwise" use, however, count as genuine uses of the eye as an eye, as opposed to "otherwise" uses that do not use the eye as an eye. In this respect, *epistēmē* is like the eye, since, Aristotle argues, one can use and misuse it in a similar way (1246a31-5).

With this background, Aristotle turns to the second part of the chapter, where he investigates whether the virtues and *phronēsis* are *epistēmai*. If the virtues are *epistēmai*, then they also should work in the way described above; if they cannot be misused, then they would not be *epistēmai* (1246a35-b1). *Phronēsis*, indeed, would need to work in this way to be *epistēmē* (1246b4-8). Now, whenever *epistēmē* is misused, there is some factor that controls it

---

[137] Woods, in his commentary, divides the chapter into three parts, with one argument for why justice is not *epistēmē* and one for why *phronēsis* is not *epistēmē*. For reasons that will become clear later on, I do not think these two arguments should be separated and will consider them together as one part of this chapter.

and guides it towards misuse, so if *phronēsis* is *epistēmē*, it must also be able to be controlled by something into misuse (1246b8-10). Aristotle rejects a number of possibilities for what could control *phronēsis* and then considers a version of *akrasia* that might explain how *phronēsis* could be misused. This conception of *akrasia*, however, would have the undesirable consequence of permitting a reverse situation, in which one can act wisely despite ignorance in the rational part of one's soul (1246b10-27). The failure of this case brings to light further differences between *epistēmai* and the virtues (1246b27-32). Following this discussion, Aristotle ends the chapter by drawing three conclusions:

> So it is clear that (i) those with *phronēsis* also have those good states of the other part [of the soul], and the Socratic point was correct, that (ii) nothing is stronger than *phronēsis*. But he said that it is *epistēmē*, which is not correct; for (iii) it is a virtue and not *epistēmē*, but another type of cognition. (1246b32-36)[138]

With (iii), Aristotle wraps up the second part of the chapter: *phronēsis*, and the virtues, are not *epistēmai*.

As I mentioned before, Aristotle has helped us by stating clearly what he has taken himself to have proved in this chapter. Nevertheless, there are a number of puzzles about how Aristotle arrives at these conclusions. The opening discussion of use and misuse obviously is important for the subsequent argument, but it is not quite clear how. Part of the obscurity lies in an unexplained shift from considering justice as an *epistēmē* to considering *phronēsis*. Regarding *epistēmē* itself, Aristotle makes a number of points at the beginning and end of the chapter about it, so we might hope to uncover what he has in mind by it. Getting clear on the

---

[138] ὥστε δῆλον ὅτι ἅμα φρόνιμοι καὶ ἀγαθαὶ ἐκεῖναι αἱ <τοῦ> ἄλλου ἕξεις, καὶ ὀρθῶς τὸ Σωκρατικόν, ὅτι οὐδὲν ἰσχυρότερον φρονήσεως. ἀλλ᾽ ὅτι ἐπιστήμην ἔφη, οὐκ ὀρθόν· ἀρετὴ γάρ ἐστι καὶ οὐκ ἐπιστήμη, ἀλλὰ γένος ἄλλο γνώς<εως>, reading the manuscripts' ἐκεῖναι and ἄλλου instead of the OCT ἐκείνων and ἀλόγου, respectively. Translations of the *Eudemian Ethics* are my own. Because providing a coherent argument of VIII.1 involves tackling difficulties within the letter of the text, I will include the Greek text in the footnotes, noting any emendations from the OCT.

contrast between *epistēmē* and *phronēsis*, two things that Aristotle names elsewhere as intellectual virtues, will help us understand how this chapter fits into his larger epistemology and ethics.

One reason the chapter is of interest to our project as a whole is that Aristotle is explicitly tackling the Socratic paradoxes. VIII.1 claims to separate virtue/*phronēsis* from *epistēmē*, in contrast to SP2's claim to identify them. The way his argument proceeds also appears to deal with SP1, as he begins by speaking of the possibility of misusing *epistēmē*. Aristotle's concluding lines call out Socrates explicitly and note that he was partially right about *phronēsis*, though not in a way that would save SP2. What we seem to have, then, is Aristotle grappling with the Socratic paradoxes in an attempt to distinguish his own picture from them. *Eudemian Ethics* VIII.1 would thus be able to give us some insight into the reception and legacy of the Socratic paradoxes after Plato.

Before we can reach a verdict on Aristotle's treatment of the Socratic paradoxes, we will also need to address some issues surrounding the position of VIII.1 in the *Eudemian Ethics* itself and in Aristotle's thought as a whole. This will involve examining the connections to and differences from the other parts of the *Eudemian Ethics*. As a preliminary note, we should dispense with any attempts to connect VIII.1 as flowing seamlessly from what precedes it. Book VIII follows a discussion of friendship in the previous book, and a couple of the manuscripts show the chapter beginning by asking whether one can use and misuse each friend, rather than each thing. The addition of "friend" to the text, however, is pretty clearly an artificial attempt to connect VIII.1 to Book VII; nothing in the subsequent parts of the chapter mentions friends or friendship, and the start of the chapter more naturally starts a significantly different thought that

permeates the rest of Book VIII. We thus have license to take the chapter on its own terms in this regard.

A more serious issue concerns the relationship between VIII.1 and an earlier (in the treatise) discussion of the intellectual virtues, taking place in the second of three "common books" shared between the *Eudemian* and *Nicomachean Ethics*. There Aristotle lays out, much more clearly and in more detail, some key features of *epistēmē*, *phronēsis*, and the difference between the two. We might wonder why Aristotle returns to the topic later on in the same work and, as we shall see, in such a different way. Although one could surely posit lacunae or a piecemeal assembly of the text to account for the peculiarity of VIII.1 (and this may well be part of the story), it is my hope that we can find some philosophical grounds to explain its status, especially given the difference in its focus.

With these points in mind, two primary tasks emerge. The first is to provide a coherent interpretation of VIII.1 that illuminates the argument and how the whole chapter fits together. The second is to determine what to make of VIII.1 in light of the rest of the *Eudemian Ethics* and other aspects of Aristotle's philosophy. We will consider VIII.1's place within the *Ethics* and some philosophical upshots thereof towards the end of this chapter. For now, we can make a start on uncovering the connections with other parts of Aristotle's thought by turning briefly from the *Eudemian Ethics* towards a relevant discussion in the *Metaphysics*.

Contrary-Use Capacities in *Metaphysics* Θ

Eudemian Ethics VIII.1's main aim is to compare *phronēsis* and the virtues to *epistēmē* to show why they are different. To that end, it will be helpful to get a grip on what Aristotle has in mind when he is speaking of *epistēmē*. If we can find a description of *epistēmē* elsewhere that would make good sense of this argument, then that may illuminate what Aristotle is trying to do.

A promising discussion of *epistēmē* for this purpose appears in Book Θ of the *Metaphysics*. The first few chapters of Book Θ focus on explaining capacities (*dunameis*) and the features of different kinds of capacities. In the passage below, from Θ.2, Aristotle distinguishes a special kind of capacity:

> Since some starting-points of these sorts are present in inanimate things, others in animate ones, and in a soul, and in the part of the soul that has reason, it is clear that some of the capacities will be non-rational, whereas some will involve reason. That is why all the *technai*, that is all the productive *epistēmai*, are capacities. For they are starting-points of change in another thing or in the same thing insofar as it is other.
>
> And all the capacities that involve reason are such that the very same one is a capacity for contraries, whereas the non-rational ones are such that one capacity is for one of them. For example, the hot is for heating only, but medicine is for both disease and health.
>
> The cause of this is that *epistēmē* is an account [*logos*], and the same account makes clear both the positive thing and its lack, except not in the same way—that is, in a way it is of both, but in a way it is rather of the positive thing. And so it is also necessary that these *epistēmai* should be of contraries, but of one intrinsically and of the other non-intrinsically. For the account too is of one intrinsically and of the other, in a way, coincidentally. For it is by denial and removal that it makes the contrary clear. For the contrary is the primary lack, and this is the removal of the other [and positive] contrary. (1046a35-b14)[139]

Aristotle first makes a distinction between rational and non-rational capacities. The capacity to heat is an example of a non-rational capacity: a stove's ability to heat a pot of water does not depend on reason in any way. *Epistēmai*, on the other hand, are categorized as rational capacities. In order to have the capacity to practice medicine, one must have the rational faculties required to understand what health is and what produces it (a stove cannot be a doctor).

---

[139] Translations of the *Metaphysics* are taken from C. D. C. Reeve with some modifications, chiefly to highlight the use of *epistēmē*.

Rational capacities are distinguished from non-rational capacities primarily by what they can produce. While non-rational capacities can bring about only one thing, rational capacities can be used to bring about contraries:

(RC): Rational capacities are contrary-use capacities.

Something hot, such as a fire, can make the surrounding environment only hotter, not cooler. By contrast, although a doctor is commonly known as someone who heals, she can produce two contrary states: health and disease, or a healthy person and a sick person.

Epistēmē's status as a contrary-use capacity stems from the relationship of the capacity to *logos*. *Logos* is what makes clear the "positive thing and its lack"—that is, how to bring about both contraries.[140] Having an explanation for how certain chemicals bind to the nervous system, for example, could allow someone to know how to cure a neurological disease or how to create a neurological poison. Altering the *logos*, or corrupting it in some way, will alter or corrupt the capacity to use one's *epistēmē*. With the *logos* intact, however, one has the tools necessary to produce the different outcomes.

Up to this point, we have talked about the two contraries as if they are on equal footing, since someone with a certain *epistēmē* has the capacity to produce both of them. Our intuition, however, is that someone with *epistēmē* of medicine is a doctor, not a poisoner. *Logos* makes clear both contraries, but not in the same way; it is, as Aristotle says, rather of the positive thing. Medicine shows one how to produce health, and by doing so, it also incidentally shows how to produce a lack of health. This lack is a further step: *epistēmē* is intrinsically of only one of the

---

[140] I follow Reeve's translation in thinking of *logos* as closer to an account or explanation (as opposed to speech or argument) that allows one to use one's *epistēmē*. Moss (2014b) is helpful for understanding the relationship of *logos* to virtue and how one might think about *logos* in this context.

contraries. A doctor does not primarily learn how to produce disease, although she is uniquely

capable of doing so.

The fact that *epistēmē* is primarily of one of the contraries might lead us to believe that it

would produce this contrary in some "natural" state and be enough to bring it about.

Aristotle makes it clear, however, that the capacity on its own is not sufficient to produce either

contrary. Because medicine can produce both health and sickness, something outside of the

epistēmē must determine which one it, in fact, produces. In Θ.5, Aristotle explains that contrary-

use capacities need something else to guide them:

> For all these non-rational capacities are such that one is productive
> of one thing, whereas the rational capacities are productive of
> contrary ones, so that [if they produced them in the way the non-
> rational ones do] they would produce contraries at the same time.
> But this is impossible.
>
> There must, then, be something else that is the controlling factor
> [*to kurion*]. I mean by this desire or deliberate choice. For
> whichever of two alternatives an agent desires in a controlling
> way, this it will do, whenever it is such as to be capable and meets
> up with what is capable of being affected. (1048a8-13)

Though medicine is primarily about health, a doctor's *epistēmē* alone does not cause her to heal

someone, but a desire to help her fellow humans or a deliberate choice to minimize suffering

causes her to use medicine to heal rather than sicken her patients. Likewise, a desire to have

fewer humans on the planet could guide her to poison and kill a patient. These desires and

deliberate choices can be manifold, but they demonstrate how contrary-use capacities need a

controlling factor:

> (CF): Something other than the contrary-use capacity controls which of the contraries the
> capacity will produce.

Epistēmai, insofar as they are contrary-use capacities, need something to guide them, or they will

not be able to produce either action.

131

To sum up, *Metaphysics* Θ provides us with a description of *epistēmē* as a contrary-use capacity (RC). *Logos* is the starting-point that shows one how to produce these contrary states, one of which is more primarily the focus of the *logos*. Despite this inclination, *epistēmē* on its own does not produce either contrary. It needs a controlling factor—Aristotle here mentions desire and deliberate choice, though there could be others—to motivate it towards one of them (CF). With this background in mind, let us turn again to the *Eudemian Ethics* to see how its argument matches up to that of *Metaphysics* Θ.

Epistēmē as a Contrary-Use Capacity in *Eudemian Ethics* VIII.1

Eudemian Ethics VIII.1, as I stipulated before, has two parts: the first concerning use and misuse, and the second inquiring whether *phronēsis* and the virtues are *epistēmai*. I believe that each part of the chapter, and the interaction of the two parts, can be understood through the lens of contrary-use capacities. In particular, as I will argue in this section, the first part sets up that *epistēmē* is a contrary-use capacity, and, as I will argue in the following two sections, the second part shows that *phronēsis* and the virtues are not. Contrasting *epistēmē* and *phronēsis* along these lines allows us to see how Aristotle establishes the three conclusions we saw above.

The first part of VIII.1, to refresh our memory, begins with the example of an eye to show the different ways to use and misuse something. An eye has one proper use, seeing correctly, and one improper use, seeing double or incorrectly by squinting or crossing one's eyes so as to distort one's vision.[141] Although the latter is not the correct use of an eye, both of them count as genuine uses—they are both using the eye as an eye, even if one is doing so incorrectly (1246a29-30).[142] By contrast, selling or eating an eye is using an eye as currency or foodstuff,

---

[141] It is pretty clear from how Aristotle frames his example that the misuse of the eye is supposed to be voluntary, a fact that will become important when we turn to *epistēmē*, rather than mis-seeing because of injury, disease, or some other cause.

[142] αὗται μὲν δὴ ἄμφω ὅτι μὲν ὀφθαλμός ἐστιν ᾗ ὀφθαλμός

not as an eye (1246a30-31).[143] Though the eye might look like the same thing in each of these cases, it is not in any true sense that the consumer of the eye is making use of an *eye*, rather than just a piece of matter to serve her purpose.

Let us look back to *Metaphysics* Θ. An eye is a sense organ, so an eye itself is clearly not a contrary-use capacity, or indeed a capacity of any kind. But the distinction between using an eye as an eye and using it as something else does depend on the presence of a capacity: using an eye involves using its capacity for sight, while using it as food does not require—and likely will prevent—this capacity.[144] When we consider how to use an eye as an eye, we are considering the ways in which it might exercise its capacity for sight. There are two ways to exercise that capacity: to see an object correctly (as one, as it is in the world), and to see an object incorrectly (as two, not as it is in the world). These two ways of seeing are contrary to each other, but they are both produced by the eye *qua* eye, and thus by the capacity for sight. Though Aristotle does not yet mention capacities here, he is beginning VIII.1 by considering how one thing—a thing that can employ its natural capacity—can produce contrary results.

Aristotle proceeds by immediately comparing the eye to *epistēmē*: "*epistēmē* is similar; for one can use it both truly and err, such as whenever one voluntarily writes incorrectly" (1246a32-3).[145] Just as one can see correctly, one can use *epistēmē* truly, such as using one's grasp of grammar to write properly. However, just as one can mis-see, one can use *epistēmē* to err. Aristotle contrasts these uses with "using [*epistēmē*] as ignorance" (1246a33).[146] Although he does not give a clear example of what such a use would be, it is clear that such a use would be

---

[143] ἄλλη δὲ κατὰ σθμβεβηκός, οἶον εἰ ἦν ἀποδόσθαι ἢ φαγεῖν
[144] Woods (1992) worries that the eye is not a good case because it is not a capacity itself, but I take the connection shown above to be clear enough not to fault Aristotle much for it.
[145] ὁμοίως δὴ καὶ ἐπιστήμη· καὶ γὰρ ἀληθῶς καὶ ἁμαρτεῖν, οἶον ὅταν ἑκὼν μὴ ὀρθῶς γράψῃ
[146] <καὶ> ὡς ἀγνοίᾳ δὴ νῦν χρῆσθαι

akin to selling or eating an eye—that is, not using *epistēmē* as *epistēmē*. Using *epistēmē* as ignorance resembles "whenever dancers use their foot as a hand and the latter as a foot" (1246a34-5); again, the use of *epistēmē* here would not be as *epistēmē*.[147] The two examples can be categorized thus:

	Using X properly	Misusing X	Using X as Y
Eye	Seeing (correctly)	Crossing one's eyes to see double	Eating an eye
Epistēmē (of grammar)	Writing correctly	Voluntarily misspelling words	"Using" grammar as ignorance

By setting up the parallel with the eye, Aristotle illustrates the different uses of *epistēmē*. Some situations may not count as using *epistēmē*, although *epistēmē* is present in some way. The more important takeaway is that there are two main uses that do count as genuine uses of *epistēmē*: using it truly or correctly and misusing or making errors through it. One needs to have *epistēmē* of spelling, grammar, and so forth to be able to write a correct sentence, and one needs to have the same *epistēmē* to write the sentence incorrectly voluntarily. Without *epistēmē*, someone may try to write an incorrect sentence, but since she is unaware of proper spelling and grammar she may accidentally write a grammatical sentence, despite her best efforts. Furthermore, if she knows absolutely nothing about spelling and grammar, she will not be able to write any sentence at all, so writing an incorrect sentence will be impossible. Many people may write incorrect sentences, but to be able to do so voluntarily, according to Aristotle, requires *epistēmē*.

---

[147] Aristotle gives a brief "example" of this phenomenon as "turning one's hand" (ὥσπερ μεταστρέψας τὴν χεῖρα, 1246b34). What he means is not at all clear, but here is my best guess: one takes everything that is necessary for writing—one's hand, pencil, paper, holding the pencil and making contact—and produces something not in the area of writing at all. Perhaps one merely scribbles on the paper or uses the pencil to tear some holes in it. One is making use of the things involved in writing, but there is no use of the *epistēmē* of grammar as such. I do not think that it is crucially important to determine what Aristotle means, because he himself seems unsure of what it would involve. He hints that it may not be possible to be ignorant from *epistēmē* but only to make errors (εἰ μὴ ἔστιν ἀγνοεῖν ἀπὸ ἐπιστήμης, ἀλλ᾽ ἁμαρτάνειν μόνον, 1246b1-2). In any case, my focus here is on the uses of *epistēmē* that are definitely genuine, so I will leave the question of whether one can use *epistēmē* as ignorance to the side.

The upshot of Aristotle's description so far is that *epistēmē*'s uses are those of a contrary-use capacity. It is not difficult to see that using grammar truly and using it incorrectly are contraries—just like using medicine to heal and to poison. Like the capacity for sight, *epistēmē* as Aristotle lays it out fulfils (RC). Even though we might consider grammar to be about spelling and good syntax, not about misspelling and poor syntax, *epistēmē* gives us the unique ability to produce both. We will discuss the role of (CF) in the following section. For now, it appears that the start of VIII.1 is showing *epistēmē* to be very much in line with what we saw in *Metaphysics* Θ.

Another thing to notice is that *epistēmē* contrasts strongly to what we saw in Plato. The Socratic paradoxes drove Plato to find a notion of *epistēmē* that would make voluntary misuse impossible. Here, on the other hand, Aristotle is leaning into voluntary misuse as a key feature of *epistēmē*. We seem to have returned to the *Hippias Minor*, in which figuring out what admits of voluntary wrongdoing is front and center. If Aristotle is to argue that some things cannot be voluntarily misused, we would do well to consider how those things relate to Plato's conception of *epistēmē*. For now, it is sufficient to note that *epistēmē*'s role as a contrary-use capacity seems to be important for Aristotle's argument: the first part of VIII.1 is geared towards establishing that *epistēmē* works in this way. With this part so understood, we are prepared to turn to the rest of the chapter and see what sense of it we can make.

Virtue, *Phronēsis*, and Control

I have argued above that Aristotle starts VIII.1 by painting *epistēmē* as a contrary-use capacity. As we saw, the chapter ends with three main takeaways focused on *phronēsis*, not *epistēmē*: (i) *phronēsis* is accompanied by a good condition of the irrational part of the soul, (ii) nothing is stronger than *phronēsis*, and (iii) *phronēsis* is not *epistēmē*. We might understand the

135

second part of the chapter as exploring the role of *phronēsis*—what it does and does not require to function, and how best to categorize it. It is clear that Aristotle ends up categorizing it as something other than *epistēmē*, and if I am correct that his notion of *epistēmē* here is as a contrary-use capacity, then we might expect the argument to show why *phronēsis* is not a contrary-use capacity.

At the start of the second part, Aristotle remarks that if all the virtues are *epistēmai*, they will act as *epistēmē* does, and in particular one could "do injustice then from justice by doing unjust things, just as one also does ignorant things from *epistēmē*; but if this is impossible, it is evident that the virtues would not be *epistēmai*" (1246a37-b1).[148] If the virtues are *epistēmai*, then (RC) will also hold for them. Justice should be able to produce injustice, allowing one to misuse justice voluntarily. Just as one can use *epistēmē* to write incorrectly, one should be able to use justice to act incorrectly. Aristotle clarifies further that this ability concerns the contrary states of affairs produced by *epistēmē*. If *epistēmē* is able to make mistakes not through ignorance itself but through "doing the same things as from ignorance," then if justice is found not to be *epistēmē* in this regard, it will not be able to produce contrary things (1246b2-3).[149] Rather, it will produce one state of affairs, and the contrary state would be produced by injustice instead.

Aristotle then switches gears a bit, asserting that "since *phronēsis* would be *epistēmē* and something true, it will do the same thing; for it would be possible to act foolishly from *phronēsis*, and to err in the same things as the fool" (1246b4-7).[150] While Aristotle had previously been

---

[148] ἀδικήσει ἄρα ἀπὸ δικαιοσύνης τὰ ἄδικα πράττων, ὥσπερ καὶ τὰ ἀγνοητικὰ ἀπὸ ἐπιστήμης· εἰ δὲ τοῦτ᾽ ἀδύνατον, φανερὸν ὅτι οὐκ ἂν εἶεν ἐπιστῆμαι αἱ ἀρεταί

[149] καὶ τὰ αὐτὰ <ἃ> καὶ ἀπὸ ἀγνοίας ποιεῖν

[150] ἀλλ᾽ ἐπεὶ φρόνησις ἐπιστήμη καὶ ἀληθές τι, τὸ αὐτὸ ποιήσει κἀκείνη· ἐνδέχοιτο γὰρ ἂν ἀφρόνως ἀπὸ φρονήσεως, καὶ ἁμαρτάνειν ταὐτὰ ἅπερ ὁ ἄφρων, reading the manuscripts' ἐπεὶ instead of the OCT ἔτι εἰ

considering virtue generally, with a focus on justice, he now considers whether *phronēsis* is *epistēmē* and will continue to do so for the rest of the chapter. Why the switch? One theory would be that considering justice as a contrary-use capacity is a clear *reductio ad absurdum*: it is so obvious that justice cannot produce unjust actions that it is not even worth making explicit.[151] It is less than obvious, however, that Aristotle would find this to be a clear *reductio*. Even if Aristotle should consider it impossible to produce unjust actions from justice, we should keep in mind that Aristotle is setting himself up to argue against some of the parts of Plato's philosophy that we have studied earlier. Plato, as we saw, also will consider it impossible to do justice from injustice, but our examination of the *Hippias Minor* showed that this was by no means an obvious or easy conclusion. Rather, it involved taking SP1 on board and finding a highly specialized theory of *epistēmē*. If Aristotle wishes to reject SP1 or depart from Plato on these issues, it would be strange for him to consider the case closed at this point.

Aristotle's language gives us further reason not to think he has given us a *reductio*. His discussion of using justice as injustice twice employs the optative mood: first stating that it would be possible to use justice as injustice (εἴη ἂν, 1246a36), and second that if this were impossible then the virtues would not be *epistēmai* (οὐκ ἂν εἶεν, 1246b1). The use of the optative does not indicate any certainty about the virtues' not being *epistēmai*—indeed, without any further statements, Aristotle seems just as uncertain about the possibility of using justice as injustice.[152] In the absence of a clear conclusion that the other virtues are not *epistēmai*, it is more likely that Aristotle thinks he can most decisively disprove that the virtues are *epistēmai* by

---

[151] Woods' commentary on the *Eudemian Ethics*, for example, states that "Aristotle seems to regard even this as a *reductio ad absurdum* of the identification of virtue with knowledge" (170).

[152] Another point against considering this a *reductio*: Aristotle does call out later when we have reached a *reductio*, saying explicitly that some other conclusions are strange. He is, of course, under no obligation to flag every *reductio*, but we might hope that he would in this case, since elsewhere he is in the business of doing so.

137

focusing on *phronēsis* in particular. The discussion of justice and the other virtues serves to set

the stage for looking at *phronēsis*: if they are *epistēmai* and work in this way, then certainly

phronēsis, as the most intellectual of the lot, is as well.[153] This is one reason why I think it best

to preserve the manuscripts' ἐπεὶ at 1246b4. Aristotle has not concluded one argument about all

of the other virtues and now switched to discussing *phronēsis* in a disjointed shift. Rather, the

identification of the other virtues with *epistēmē* leads to the identification of *phronēsis* with

epistēmē, and disproving it for *phronēsis* is sufficient to do so for the other virtues.

   A key reason to shift the focus to *phronēsis* becomes clear a few lines later, when

Aristotle gives us some more explanation about *epistēmē*.   He notes that *epistēmē* itself is not

sufficient for misuse: "now for all the other *epistēmai* some other controlling factor makes the

turn [into error]" (1246b8-9).[154] The mention of a controlling factor should immediately recall

Metaphysics Θ.5 and its attribution of (CF) to *epistēmē*. We see the same considerations at work

here in the *Eudemian Ethics*; by once more connecting *epistēmē* to (CF), Aristotle is giving us

further evidence of what kind of *epistēmē* he has in mind.[155] Whereas the other *epistēmai* work

in this way, "what is the controlling factor of the very factor that controls them all?" (1246b9-

10).[156] Given that Aristotle has been discussing *phronēsis* in comparison to the other *epistēmai*,

---

[153] It is worth noting that the connection between *phronēsis* and what are described as "virtues of character" is stronger in the *Eudemian Ethics* than in the *Nicomachean Ethics*. *Eudemian Ethics* II, for example, lists *phronēsis* as one among the many virtues, rather than an intellectual virtue in a separate category of virtues. Therefore, using *phronēsis* as a demonstrative case for the other virtues would be an effective move in the *Eudemian Ethics*, more so than in the *Nicomachean Ethics*. On the character-like aspects of *phronēsis*, see Pearson (2007). On the strict unity of the virtues in the *Eudemian Ethics*, see Bonasio (2020).
[154] ἐπὶ μὲν οὖν ταῖς ἄλλαις ἐπιστήμαις ἄλλη κυρία ποιεῖ τὴν στροφήν
[155] Given the grammar of 1246b8-9, one might plausibly object that *allē kuria* in the feminine refers to *epistēmē*, and thus the controlling factor here is *epistēmē* rather than desire or deliberate choice. If this is so, then it is not something outside of *epistēmē* that controls it, and (CF) does not hold in VIII.1. I am inclined to agree with the first part of this objection but not the second. Even if *kuria* refers to *epistēmē*, it is another kind of *epistēmē* that controls the many *epistēmai*, and the controlling factor is outside of all of these contrary-use capacities. The *epistēmē* of grammar, for example, does not contain in itself whether to write correctly or incorrectly; even if what does determine this is also an *epistēmē*, it is not the particular *epistēmē* of this contrary-use capacity. We may also see the feminine as anticipating *phronēsis*, which Aristotle clearly has in mind in this sentence.
[156] αὐτῆς δὲ τῆς πασῶν κυρίας τίς <κυρία>

the indication is that *phronēsis* is the controlling factor in question. In order to be a contrary-use capacity itself, it must also satisfy (CF) and be guided by something outside of it. The question, then, is what can be a controlling factor for *phronēsis*.

Aristotle quickly provides a list of candidates that do not serve this function: "for it is not *epistēmē* or *nous*. But nor is it virtue. For it uses this. For the virtue of the ruling part uses that of the ruled part" (1246b10-12).[157] Aristotle states that *phronēsis* is the virtue of the ruling part of the soul and so is not controlled by what it rules. If it turns out that nothing controls *phronēsis*, this would be a significant step towards the conclusion that *phronēsis* is not *epistēmē*. Without satisfying (CF), it would be very difficult, if not impossible, to identify *phronēsis* as a contrary-use capacity.

Akrasia and Wise Foolishness

Aristotle recognizes, however, that not everyone will concede that nothing controls *phronēsis*. In particular, some consider *akrasia*, allegedly "an evil of the irrational part of the soul," to be this very phenomenon (1246b12-13).[158] As we saw in the *Protagoras*, *akrasia* is tricky to explain. There, Socrates argued that the Many's description of the phenomenon turned out to be ignorance rather than a case of knowledge being overpowered. Here it seems that Aristotle is dealing with the same issue: what precisely happens in a case of *akrasia* is disputed, so one must examine the alleged explanation in more detail. In addition, as with the *Protagoras*, this examination should allow him to draw further conclusions about virtue and *epistēmē*.

---

[157] οὐ γὰρ ἐπιστήμη γε ἢ νοῦς. ἀλλὰ μὴν οὐδ᾽ ἀρετή. χρῆται γὰρ αὐτῇ. ἡ γὰρ τοῦ ἄρχοντος ἀρετὴ τῇ τοῦ ἀρχομένου χρῆται, omitting ἔτι at 10

[158] Τίς οὖν ἐστίν; ἢ ὥσπερ λέγεται ἀκρασία κακία τοῦ ἀλόγου τῆς ψυχῆς

The theory for *akrasia* that Aristotle now considers involves *phronēsis* being guided by appetite, which causes it to produce an intemperate action and thus a contrary action to virtue. According to this theory,

> If the appetite is strong, it will turn it and the *phronēsis* of the
> akratic will reason the opposite things, it is clear that, even if there
> is virtue in this [rational] part, there is ignorance in the *logos*, by
> which they are changed. (1246b14-7)[159]

In this situation, a strong appetite can turn an akratic's rational part and make *phronēsis* reason contrary to what is correct. Appetite, then, would be a controlling factor for *phronēsis*, and (CF) would hold since something outside of *phronēsis* would be in charge. With this version of *akrasia*, one could plausibly consider *phronēsis* to be a contrary-use capacity. Appetite would use *phronēsis* to produce contrary states of affairs.

Admittedly, the reconstruction of the text in the angle brackets is beneficial to my reading by designating *phronēsis* as what is turned by appetite to reason the opposite things. While ἡ τοῦ ἀκρατοῦς φρόνησις is not a crazy emendation, it is far from being obviously correct; the manuscripts have only a mysterious σφι, and many other emendations have been proposed.[160] Although I am inclined to believe that something to do with *phronēsis* or *phronein* should be reconstructed here, that alone is not sufficient grounds to conclude that this version of *akrasia* shows how *phronēsis* could be a contrary-use capacity. What seems to be more telling, however, is the appearance of *logos* and its role in producing *akrasia*. According to the presented version of *akrasia*, while there is *phronēsis*, or virtue in the rational part of the soul, there is ignorance in the *logos*, which causes the akratic to perform an intemperate action. If we recall from the

---

[159] ἀλλ᾽ εἰ δή, ἂν ἰσχυρὰ ᾖ ἡ ἐπιθυμία, στρέψει καὶ λογιεῖται τἀναντία ἡ <τοῦ ἀκρατοῦ>ς φρ<όνησις>, δῆλον ὅτι, κἂν ἐν μὲν τούτῳ ἀρετή, ἐν δὲ τῷ λόγῳ ἄγνοια ᾖ, ἕτερα μεταποιοῦνται.
[160] Reconstructions of the phrase include ἢ ἀντιστρόφως by Dirlmeier, ἤ ἔστι by Jackson, τῇ σωφροσύνῃ by Allan, and ἢ ὁ νοῦς φρονεῖ by Fragstein.

Metaphysics, *logos* is what allows *epistēmē* to produce either of its contraries, since it makes

clear how to produce them both. It is when *logos* is gripped by ignorance, however, that one

inclines towards the intemperate thing instead of the temperate one. This is in keeping with the

point that *epistēmē* is intrinsically of only one of the contraries; some corruption has turned the

agent towards the other one. In this case, when ignorance affects the *logos*, it affects which

contraries are produced, and appetite can bring about one rather than the other.

  Logos reappears a few lines later, when Aristotle continues the argument by asserting that

if this kind of *akrasia* is possible, then so is the opposite case, i.e. a virtuous appetite turning a

bad rational part into a good one. Both cases would have to exist, "for it would be strange if

vice, when it is present, will indeed turn the virtue in the rational part through *logos* and make it

ignorant, but virtue in the irrational part, when ignorance is present, will not turn this and make it

judge wisely and the proper things" (1246b19-23).[161] Again, vice affects *logos* to turn *phronēsis*

and incline it to produce a bad action.[162] But if this version of *akrasia* is correct, then appetite, as

a controlling factor, can guide an ignorant rational part to act well if the irrational part itself is in

good condition. If the irrational part can control reason in this way, then "it will be possible [to

act] wisely from ignorance" (1246b25).[163]

---

[161] ἄτοπον γὰρ εἰ τὴν μὲν ἐν τῷ λογιστικῷ ἀρετὴν μοχθηρία ποτὲ ἐγγενομένη μέντοι λόγῳ στρέψει καὶ ποιήσει ἀγνοεῖν, ἡ δ᾽ ἀρετὴ ἐν τῷ ἀλόγῳ, ἀγνοίας ἐνούσης, οὐ στρέψει ταύτην καὶ ποιήσει φρονίμως κρίνειν καὶ τὰ δέοντα, omitting the angle brackets and emending the OCT's ἐν τῷ ἀλόγῳ to μέντοι λόγῳ in 20

[162] A careful observer will notice that I have emended the text to include *logos*; the OCT reads μοχθηρία ποτὲ ἐγγενομένη ἐν τῷ ἀλόγῳ, "whenever vice is present in the irrational part," eliminating the "by means of *logos*" that allows me to argue for *logos*' role as a starting-point here. Before I am accused of gratuitously tailoring the text to fit my needs, it is worth noting that my emendation is to an emendation itself, as the original manuscripts read "…ἐγγενομένη μέν τῷ λόγῳ." Changing μέν τῷ λόγῳ to μέντοι λόγῳ is plausible and more conservative than changing it to ἐν τῷ ἀλόγῳ, especially given the significant change of meaning that results from modifying λόγῳ to ἀλόγῳ. Kenny points out that μέντοι appears much more frequently in the *Eudemian Ethics* than in the *Nicomachean Ethics* (74), so using the term in this case would be in keeping with the language of the *Eudemian Ethics*. Perhaps some construction other than μέντοι can replace μέν τῷ satisfactorily; in any case, I believe that what is important here is preserving λόγῳ instead of making an unnecessary modification to ἀλόγῳ.

[163] ἔσται καὶ ἀπὸ ἀγνοίας φρονίμως

Intuitively, at least for Aristotle, these conclusions are strange, especially being able to act wisely from ignorance (1246b25-6).[164] One reason why this is so could just be intuition: good desires simply do not seem to be sufficient for getting us to act wisely. More importantly, Aristotle notes, "we do not see this at all for any of the other [*epistēmai*]" (1246b27).[165] In particular, "intemperance turns medicine or grammar, but not ignorance, if it [intemperance] is opposite, on account of there not being any excess in it" (1246b28-30).[166] Intemperance can control *epistēmai* and make them produce bad states of affairs: perhaps an unchecked desire for money causes a doctor to poison a wealthy client or not heal him as quickly in order to receive more payments for services. However, the opposite desire—presumably a temperate one— cannot use ignorance to produce either contrary. No matter how much I may selflessly desire to heal people, without *epistēmē* of medicine I am no closer to healing them than to poisoning them. This is because ignorance is an absence of *epistēmē*, not an excess or additional element; there is nothing in ignorance that can produce one contrary or the other. Both contraries—healing and poisoning, in the case of medicine—can result from *epistēmē*, hence its status as a contrary-use capacity. While *epistēmē* can be guided to one action or another, its absence lacks the *logos* that makes both contraries clear, and that absence does not provide the capacity to make either contrary.

---

[164] ἔστι δὲ ταῦτα ἄτοπα, ἄλλως τε καὶ ἀπὸ ἀγνοίας χρῆσθαι φρονίμως. Though Aristotle declares these results strange, they are not particularly foreign to the modern ethical landscape. Arpaly and Schroeder (1999) lay out the notion of "inverse *akrasia*," in which an agent acts rightly despite what she thinks is her better judgment. Their classic example is the case of Huck Finn, who cannot bring himself to turn in a runaway slave and admonishes himself as being weak-willed as a result. Arpaly and Schroeder also think that Aristotle raises a case of inverse *akrasia* in the third common book, where he considers the case of Neoptolemus from Sophocles' *Philoctetes*. Neoptolemus was persuaded by Odysseus to lie to Philoctetes but ended up telling the truth and not following through with his decision to lie. Aristotle gives this example as a potential case of "a sort of *akrasia* that is excellent" (1146a19). Later on, however, Aristotle denies that this is truly a case of *akrasia*. For a more in-depth discussion of the Neoptolemus case, see Cagnoli Fiecconi (2018).

[165] τοῦτο γὰρ ἐπὶ τῶν ἄλλων οὐδαμῶς ὁρῶμεν, reading the manuscripts' οὐδαμῶς instead of the OCT οὐδεμιᾶς

[166] τὴν ἰατρικὴν ἢ γραμματικὴν στρέφει ἀκολασία, ἀλλ᾽ οὖν οὐ τὴν ἄγνοιαν, ἐὰν ᾖ ἐναντία, διὰ τὸ μὴ ἐνεῖναι τὴν ὑπεροχήν, omitting the angle brackets

Virtue and vice, however, do not work in this way. First of all, since the proposed version of *akrasia* fails, nothing has been shown to be the controlling factor for *phronēsis*. Therefore, (CF) as it would apply to *phronēsis* is in jeopardy. Second, while ignorance is a lack that cannot be guided by temperance, "virtue in general deals more with vice when it is in this condition" (1246b30-1).[167] Virtue is not a special capacity that gives one a power to produce contrary actions, and vice is not an absence of this capacity. Instead, "the unjust person can do everything the just person can, and the incapacity is generally present in the capacity" (1246b31-2).[168] Although translators often render the above phrase as "the just person can do everything the unjust person can" (hence the <ἄ> before ὁ ἄδικος supplemented by the OCT), it is not only more natural to translate it the other way around but also brings to light how virtue is not a contrary-use capacity. Virtue does not give someone a special capacity to do certain actions the way that *epistēmē* of medicine gives someone the power to heal and to sicken. Rather, the person without virtue is capable of doing all the same actions as the virtuous one, even though virtuous and vicious people will not, in fact, produce the same actions.

We now turn to the three conclusions that finish the chapter. Aristotle concludes first that (i) *phronēsis* is accompanied by a good condition of the irrational part of the soul. *Phronēsis* has been shown not to abide by (CF), that something outside of it controls whether it produces a good state or a bad one. Rather, to be someone with *phronēsis* just is to be inclined towards the good: *phronēsis* guides the rest of one's soul to do good actions.[169] Because it does not admit of

---

[167] τὴν ἀρετὴν ὅλως μᾶλλον εἶναι πρὸς τὴν κακίαν οὕτως ἔχουσαν

[168] καὶ γὰρ ὁ ἄδικος πάντα ὁ δίκαιος δύναται, καὶ ὅλως ἔνεστιν ἐν τῇ δυνάμει ἡ ἀδυναμία, omitting the angle brackets

[169] Conclusion (i) seems to incline away from the view that *phronēsis* could be a controlling factor to an unwilling irrational part of the soul—in short, that there could be someone with both *phronēsis* and *enkrateia*. For this reason, I would hesitate to accept Callard's (2017) interpretation of VIII.1 as pointing to a picture of *enkrateia* that could co-exist with *phronēsis*. The brief discussion of *enkrateia* in the chapter comes within Aristotle's *reductio* of the version of *akrasia* mentioned above, and it seems more natural to read the discussion as painting an analogous

misuse in the way that a contrary-use capacity could, it can control a soul on its own. Aristotle thus agrees as well with (ii), the "Socratic point" that nothing is stronger than *phronēsis*.

More generally, (RC) does not hold for *phronēsis* and the virtues. Virtue does not give someone the ability to produce contrary states of affairs—for cases of temperance and intemperance, one is just as capable of doing the intemperate as the temperate thing. Virtue does not give us a *logos* that shows us how to do something of which we were previously ignorant, and *phronēsis*, for all its strength, does not give us the ability to produce contrary states. Therefore, it cannot be *epistēmē*, as (iii) submits. As *phronēsis* has failed the test, so have the other virtues: (iii) brings out this point especially clearly, connecting *phronēsis*'s and the virtues' difference from *epistēmē* to each other. What other kind of cognition *phronēsis* may be is left unresolved, but we know that it is not what we have been considering.

As shown above, Aristotle's three conclusions follow from the earlier parts of the chapter when we understand his argument as showing *phronēsis* not to be a contrary-use capacity. *Epistēmē* is such a capacity, and as such it can produce contrary states of affairs but must be controlled by something else to do so. Once this is established, Aristotle can show that *phronēsis* does not work in this way but is itself the highest controlling factor, inclining the soul to produce only one type of action. We can thus understand VIII.1 as demonstrating these points and supporting its clear conclusions.

The Common Books and Aristotle's Development of *Epistēmē*

The primary aim of *Eudemian Ethics* VIII.1 has been to argue for the difference between *epistēmē* and *phronēsis*. Taken on its own, this seems to be a perfectly fine topic for Aristotle to

---

version of *enkrateia* rather than Aristotle's endorsed version. Callard, for her part, admits that it is hard to argue for an enkratic *phronimos* on the basis of this passage alone.

consider near the end of the *Eudemian Ethics*. In this particular case, however, there is good

reason to wonder why Aristotle bothers to make such an argument at all. It just so happens that a

few books earlier, in Book V, Aristotle already gave us a relatively short and simple argument

for why *phronēsis* and *epistēmē* are not the same, leaving VIII.1 in a strange and seemingly

superfluous position. I believe that examining these arguments in tandem will, rather than

confirm VIII.1's superfluity, help us understand not only VIII.1 but also *epistēmē* itself as they

figure in Aristotle's thought.

     Eudemian Ethics V is the second of the three "common books" that appear in both the

Eudemian Ethics and the *Nicomachean Ethics* as we have them; *Eudemian Ethics* V is identical

to *Nicomachean Ethics* VI. We will consider the true "home" of the common books below; to

maintain neutrality at present, let us follow Kenny (1978) and call it Book B. In Book B,

Aristotle discusses, among other things, the various states of mind that grasp the truth, putting

epistēmē and *phronēsis* into that category, as well as *technē*, *nous*, and *sophia*. Aristotle

considers what *epistēmē* is in B.3, a topic that will become evident "if one is to speak in an exact

way and not be guided by mere similarities" (1139b18-9).[170] In this chapter, he emphasizes two

main features of *epistēmē*. First, an object of *epistēmē* "does not at all admit of being otherwise

[. . .] Hence what admits of being known by *epistēmē* is by necessity. Hence it is eternal"

(1139b20-3). The version of *epistēmē* presented here is one whose objects are eternal and

unchanging. The second main feature of *epistēmē* is that it "seems to be teachable, and what can

be known by *epistēmē* to be learnable" (1139b25-6). Gaining *epistēmē* involves learning

demonstrations, and Aristotle concludes that "*epistēmē* is a state affording demonstrations"

---

[170] All translations and citations of Book B are taken from Reeve's *Nicomachean Ethics* unless otherwise noted, with some modifications to bring out uses of key terms.

(1139b31-2). The objects of *epistēmē*, therefore, are (a) necessary and eternal, and (b) known

through demonstration.

In B.5, Aristotle turns to *phronēsis*. He characterizes someone with *phronēsis* as

someone who is "able to deliberate correctly about what is good and advantageous for himself"

(1140a25-7). To deliberate is to consider and calculate about things that are under one's control,

which means that "nobody deliberates about things that cannot be otherwise" (1140a31-2). This

does not bode well for categorizing *phronēsis* as *epistēmē*:

> So, since *epistēmē* involves demonstration, and the things whose
> starting-points admit of being otherwise cannot be demonstrated
> (for all of them also admit of being otherwise) and it is not possible
> to deliberate about what holds by necessity, *phronēsis* cannot be
> either *epistēmē* or *technē*—not *epistēmē* because what is doable in
> action admits of being otherwise, not *technē* because action and
> production differ in kind. (1140a33-b4)

Aristotle makes clear that both (a) and (b) fail for *phronēsis*: its objects are not necessary, nor

can they be known through demonstration. With *epistēmē* and *phronēsis* defined in these ways,

it is easy to see that the two are not the same. One may, in light of this, wonder about the status

of VIII.1's arguments: if there are further reasons that *phronēsis* is not *epistēmē*, why not

mention them in Book B? And given the efficacy of Book B's argument, why bring them up

them at all, not to mention in a much more convoluted way?

I submit that our earlier study of Plato will aid us in investigating this puzzle in Aristotle.

Plato started his epistemological project with a fairly ordinary notion of *epistēmē*, and as he

considered some issues surrounding virtue and wrongdoing, *epistēmē* transformed into

something incapable of voluntary misuse. At different places, Plato thus exhibits two different

notions of *epistēmē*. When we look more closely at Aristotle, we will see the same phenomenon

taking place. The clearest way to see the difference in the two conceptions of *epistēmē* is by

146

considering each one's relationship to *technē*. In the passage quoted above, Aristotle gives one

reason each that *phronēsis* is neither *epistēmē* nor *technē*: *epistēmē*, unlike *phronēsis*, does not

deal with what admits of being otherwise, and *technē*, unlike *phronēsis*, deals with production

instead of action. If *technē* and *epistēmē* were the same, Aristotle could give the same reason

that they are not *phronēsis*; their differences from *phronēsis*, however, bring to light the

differences between them. While *epistēmē* is a state concerning demonstrations and necessary

things, *technē* is "a productive state involving true reason" (1140a10). Furthermore, "things that

are or come to be by necessity are not the concern of *technē*" (1140a13-14). The notion of

epistēmē in Book B is as different from *technē* as it is from *phronēsis*.

By contrast, *epistēmē* in *Eudemian Ethics* VIII.1 and *Metaphysics* Θ is very similar, if not

identical, to *technē*. *Metaphysics* Θ.2 explicitly labels all of the productive *epistēmai* as *technai*.

While *technē* is not mentioned in VIII.1, Aristotle's examples of *epistēmai* strongly point us in

that direction. The examples of *epistēmai* given in VIII.1, medicine and grammar, are common

examples of *technai* in some of Aristotle's other discussions. In *Nicomachean Ethics* II.4, for

example, he compares *technē* to virtue by considering the case of someone with the *technē* of

grammar (1105a20-25), and in *Metaphysics* A he discusses how a doctor's *technē* of medicine

differs from mere experience (981a5-12). When Aristotle presents the argument in the

Eudemian Ethics, he is showing how *phronēsis* and the virtues do not fit with this *technē*-notion

of *epistēmē*, an argument which needs to be made separately from the one in Book B.

Epistēmē in VIII.1, therefore, is significantly different from *epistēmē* in Book B as well

as from other parts of Aristotle's philosophy (such as the *Posterior Analytics*, to which Book B's

discussion of *epistēmē* refers).[171] With regard to the charges of superfluity, however, we are not

---

[171] "Hence *epistēmē* is a state affording demonstrations and has the other features included in the definition we give in the *Analytics*" (1139b31). For more on *epistēmē* in the *Posterior Analytics*, see Bronstein (2016).

entirely off the hook. In particular, if *epistēmē* is to be understood along the same lines as *technē*, the problem has merely shifted in focus: Aristotle *also* argues against identifying *technē* and *phronēsis* in the same part of Book B, so VIII.1 once more appears out of place. These observations, I argue, put us in a position to make two final conclusions about Aristotle and his thought. The first is on a smaller scale and concerns the relationship between the *Eudemian* and *Nicomachean Ethics*. The second is more general and concerns the role of *epistēmē* and the Socratic paradoxes.

As mentioned before, Book B is the second of three books that appear in identical form in both the *Eudemian* and *Nicomachean Ethics*. There has been much speculation as to which treatise provides a more natural home for the common books (it is usually assumed that they were written in the context of one and "transplanted," either backwards or forwards, into the other). Kenny (1978) is largely responsible for not only launching more rigorous study of the *Eudemian Ethics* in its own right but also for arguing that the common books properly belong in the *Eudemian Ethics*, which has been the dominant view in the past few decades. Recently, this view has been challenged by Frede (2019), who contends that several themes in the common books fit more clearly into the *Nicomachean Ethics*, although both treatises must have had something like the common books.[172] One of these themes is the intellectual virtues in Book B; *phronēsis*, as well as *sophia*—theoretical wisdom, the combination of *epistēmē* and *nous*—are treated much differently in the *Eudemian Ethics*. Indeed, *sophia* is not mentioned outside of the common books at all, leading Frede to posit that Aristotle's picture of the intellectual virtues in Book B is absent or different from what may have appeared in the *Eudemian Ethics*.

---

[172] Frede also gives a nice summary of the development of scholarly positions on this issue (85-87).

Frede does not mention VIII.1, but I believe that what we have seen about its argument lends support to her theses. VIII.1, not to mention the rest of the undisputed *Eudemian Ethics*, gives no indication of a conception of *epistēmē* that matches what we see in Book B. Instead, we have a different conception of *epistēmē* in VIII.1 with no acknowledgement of said difference. The two arguments against identifying *phronēsis* and *epistēmē* point us to conclude that each one was from a separate work. Though Aristotle thought it necessary to distinguish *phronēsis* from *epistēmē* already when writing the *Eudemian Ethics*, he did not have the precise conception of *epistēmē* that would do this work most easily. We might therefore conclude that VIII.1 represents an early stage of grappling with *epistēmē* and the Socratic paradoxes, without a fully fleshed-out picture of the intellectual virtues.

I believe that we have further evidence of VIII.1's status as an early, pre-common book chapter through seeing an echo of its argument in Book B. As I mentioned above, *epistēmē* in VIII.1 is more akin to Aristotle's notion of *technē* than to *epistēmē* as it appears in Book B. When we look closer at Aristotle's treatment of *technē* and *phronēsis*, we are met with an interesting, almost throwaway passage about the difference between the two:

> Well, of *technē* there is certainly a virtue, whereas of *phronēsis*
> there is not one. And, in the case of *technē*, someone who makes
> errors voluntarily is preferable but with *phronēsis* he is less so, as
> is also the case with the virtues. It is clear, then, that it is some sort
> of virtue and not *technē*. (1140b21-24)

I would argue that this passage, more than anything else in Book B, displays the same line of thought as in VIII.1. Aristotle briefly indicates that *phronēsis* does not abide by (CF) while *technē* does—the latter needs a "virtue" (something that seems to play the role of a controlling factor) to guide it to do well, unlike the former. Both passages mention making errors voluntarily and the differences between *phronēsis* and *technē* in this regard. The conclusion of

this passage and VIII.1 are also similar in (admittedly brief) style; compare "it is clear, then, that it is some sort of virtue and not *technē*" (δῆλον οὖν ὅτι ἀρετή τις ἐστὶ καὶ οὐ τέχνη) to "for it is a virtue and not *epistēmē*" (ἀρετὴ γάρ ἐστι καὶ οὐκ ἐπιστήμη) at the end of VIII.1. The thrust of VIII.1 is not entirely absent from Book B but rather appears (1) as about *technē*, not *epistēmē*, and (2) as more of an afterthought than the main argument.

The situation we have examined above can be best understood with the following picture: the *Eudemian Ethics*, VIII.1 included, represents an early attempt by Aristotle to separate *phronēsis* from *epistēmē*. As he developed his ethics and epistemology, his notion of *epistēmē* developed as well, leaving the *epistēmē* of the *Eudemian Ethics* more similar to *technē*. By the time he wrote the common books and the rest of the *Nicomachean Ethics*, he had clear and concise reasons why *epistēmē*, *technē*, and *phronēsis* all differed from each other. One of the things separating *technē* from *phronēsis* maps onto the argument about *epistēmē* that he had laid out initially, and this argument finds its way into, though is not the primary focus of, Book B.

Let us, finally, consider what this picture means for understanding Aristotle's attitude towards *epistēmē* and the Socratic paradoxes. VIII.1 demonstrates an explicit attempt on Aristotle's part to move away from the "Socratic point" about identifying virtue and *epistēmē*. The argument Aristotle presents echoes both the *Hippias Minor*, in considering different domains of voluntary misuse, and the *Protagoras*, in examining alternative explanations for *akrasia*. When faced with challenges about voluntary wrongdoing in both dialogues, Plato's response was to develop *epistēmē* from something akin to *technē* into something robust and uniquely situated to guarantee virtue, of the sort that we saw in the *Republic*. Aristotle does not want to make this move. He does not see a way to connect *epistēmē* to virtuous action; likely he,

in accord with many scholars, is quite worried about the Problem of the Gap.[173] Instead he concludes that something else is necessary for living virtuously: *phronēsis*.

What we see with this contrast, I would argue, is not that Aristotle has rejected the Socratic paradoxes wholesale. Rather, he has found a new way to accommodate their force without embracing a connection between highest cognition and virtuous action. *Phronēsis*, as practical wisdom, takes the role that Plato had for *epistēmē*, as something that can guarantee virtue. We see its role in Aristotle's other conclusions from VIII.1: *phronēsis* is accompanied by virtue in the non-rational part of the soul, and nothing is stronger than *phronēsis*. Aristotle notes his debt to the "Socratic point" for this latter conclusion, acknowledging that something intellectual must rule in the soul that is not overcome by other motivations. The same problems of voluntary misuse that troubled Plato have led Aristotle to posit a unique and important role for *phronēsis*. Though in the *Eudemian Ethics* he has not yet constructed a full account of the intellectual virtues, he still finds the need for something *epistēmē*-like to guide right action and a good life.

My speculation would be that the *Eudemian Ethics* shows Aristotle at his most at odds with Plato on these points: he has found *epistēmē* of Forms very unsatisfying for ethics and epistemology, and he is separating his views from Plato with this picture in the forefront of his mind. Even at this stage, however, he recognizes that he is not rejecting the importance of *epistēmē* for virtue so much as shifting focus to another cognitive state. As he thinks further about his epistemology, he realizes that there is, in fact, a need for cognition that grasps eternal and unchanging things. Furthermore, this cognition is the best state of the best part of the soul—

---

[173] Aristotle may be the *first* scholar to worry about the Problem of the Gap. He is clearly uncomfortable with Plato's Forms and mentions arguing against them in many places. *Eudemian Ethics* I.8 displays this discomfort nicely: "even if it is the case that there are Forms and a Form of the Good, it is not useful in the least for a good life or good actions" (1217b23-25).

it is a virtue. It does not encompass *all* virtue; the Problem of the Gap never ceases to bother

Aristotle, and we must posit different intellectual virtues for theoretical thinking and practical

living. But with *phronēsis* replacing *epistēmē*, and *epistēmē* gaining its own role as part of the

virtue of theoretical thinking, we get Aristotle's own marriage of ethics and epistemology.

Conclusions

One standard assumption about the Socratic paradoxes is that they only hold force for

those who buy into the Socratic-Platonic strain of philosophy. Our examination of *Eudemian*

Ethics VIII.1 shows this to be only partially true. Aristotle concludes explicitly that *phronēsis*

and the virtues are not *epistēmai*, through showing *epistēmē* as able to be voluntarily misused in

the way that Plato's notion of *epistēmē* aims to preclude. If Aristotle had identified "Socratic

paradoxes" in Plato's thought, it looks like he would reject them. Nevertheless, I believe that our

examination of his argument in VIII.1 shows him to be up to a project that is more similar to

Plato's than one might think. Some intellectual state must be in control of the soul, unable to be

overcome or go wrong voluntarily. Plato and Aristotle have both argued that a *technē*-notion of

epistēmē cannot serve this role, so one must find it elsewhere and develop it accordingly. They

are motivated by similar problems about wrongdoing, and their solutions both involve aspects of

the Socratic paradoxes.

For Plato, the solution is to embrace a notion of *epistēmē* that is of only the highest

objects of cognition, one that will necessarily lead a person to right *doxa* and right action.

Aristotle has his doubts about this picture and instead finds a replacement for *epistēmē*'s role in

phronēsis. As he thinks further about the role of theoretical cognition, he finds a place for

epistēmē in his picture of virtue as well. The place of these intellectual virtues shows that for

Aristotle, virtue *is* knowledge, especially because *phronēsis* is also accompanied by the virtues

of character. On this picture, someone with *phronēsis* would not be akratic nor vicious: voluntary wrongdoing is alien to someone in such a condition. Even Aristotle, we can see, feels the pull of the Socratic paradoxes, and it is no surprise that his philosophy developed accordingly.

Conclusions and New Directions

In my dissertation, I have argued that the Socratic paradoxes are an essential part of Plato's epistemology. As we have seen, puzzles about voluntary wrongdoing have led Plato to conclude that no one does wrong voluntarily and to seek a conception of *epistēmē* that can explain this conclusion. The solution he comes to is that *epistēmē* involves measurement in such a way as to guarantee virtuous actions for its possessor. What *epistēmē* turns out to measure is truth—the one thing it has in common with *doxa* and its route to improving the imperfect *doxai* that we have about the world. When we are able to measure how much truth a *doxa* has, we can use our *epistēmē* to act in way that turns us towards the highest realities. In so doing, we lead a good life through having our souls in the best condition possible. This picture helps make sense of a number of peculiarities about Plato's philosophy, but its importance is not limited to Plato: as we have seen, Aristotle feels the force of the same puzzles and draws many similar epistemological conclusions, despite disagreeing with some aspects of the Socratic paradoxes.

One of my primary aims in this dissertation has been to illustrate the connection between the Socratic paradoxes and Plato's epistemology, and in doing so to show how understanding this connection enhances our grasp of each of them. In closing, I will briefly consider some avenues for further exploration of both of these features, prompted by the new picture we have uncovered. We have seen, first of all, that the Socratic paradoxes occupy a central place in Plato's thought—they are not some strange throwaway commitments but instead inform his philosophy as a whole. I have argued for their influence on *epistēmē* in particular, but we can already see other areas in which they are sure to have a central role. We have talked briefly, for instance, about Plato's account of voluntary action that we noticed in the *Hippias Minor*. With this account in mind, we would be able to examine whether Plato has a consistent account of

voluntary action throughout this corpus and how his commitment to SP1 and SP2 may shape his treatment of the voluntary. Another area that we have explored only briefly is political philosophy. If virtue is *epistēmē* but only the rulers of the city have *epistēmē*, we might wonder where this leaves Plato's account of virtue and its intersection with politics. The Socratic paradoxes may thus further shape not only the account of virtue in the individual soul but also that of virtue in the state.

Regarding Plato's epistemology, one primary discovery has been to understand *epistēmē* as a measurer of truth. I believe that an exciting path forward would be to try to get a more complete grasp of this picture and its importance. One promising way to do this would be to explore these themes in Plato's *Sophist* and *Statesman*. These two dialogues are usually thought to be late in Plato's career, after he wrote the *Republic*. One way they differ substantially from the *Republic* is in moving away from the account of Forms that figured so prominently in the latter. We might therefore expect that the epistemological picture we have laid out above would not have much to do with Plato's later thought. But the discussions in the *Sophist* and *Statesman* point us in a different direction: the *Sophist* focuses significantly on truth and the *Statesman* on measurement. Our investigation has given us the tools to understand Plato's epistemological development and its relation to other aspects of his thought. It would not be a surprise to find Plato continuing this development, even at a point where he backs off from his embrace of the Forms, and our work here may leave us uniquely positioned to understand it.

As I mentioned at the very start, the Socratic paradoxes and Plato's epistemology are noteworthy for their peculiarity. I also noted that it would be a misnomer to call his views paradoxical. But this assertion was overly hasty. We need not find a logical contradiction to call something a paradox: at its heart, the Greek *para doxan* means *against common opinion*. I do

not think that what I have argued above, nor the new directions that I mention, would render Plato unparadoxical in this sense. But that itself is no big worry: after all, the picture we have painted is one that blends together *epistēmē*, virtue, truth, measurement, and much more. It is a nearly breathtaking view, and it is no surprise that quite a bit about it *will* strike us as surprising. I do not aim to take away what is *para doxan* from Plato's philosophy. I merely hope instead to have improved our own *doxa* about it, in order to grasp it more fully and weigh it more clearly.

Bibliography

Allen, Reginald E. "The Socratic Paradox." *Journal of the History of Ideas* 21, no. 2 (1960): 256-65.

Allen, Reginald E. "The Argument from Opposites in *Republic* V." *Review of Metaphysics* 15 (1961): 325-35.

Annas, Julia. *An Introduction to Plato's* Republic. New York: Oxford University Press, 1981.

Annas, Julia. *Platonic Ethics, Old and New.* Ithaca, NY: Cornell University Press, 1999.

Annas, Julia, and Rowe, C. J., editors. *New Perspectives on Plato, Modern and Ancient*. Cambridge, MA: Harvard University Press, 2002.

Anscombe, G. E. M. "Modern Moral Philosophy." *Philosophy* 33, no. 124 (January 1958): 1-19.

Aristotle. *Ethica Eudemia*. Edited by R. R. Walzer and J. M. Mingay. Oxford: Oxford University Press, 1991.

Aristotle. *Ethica Nicomachea*. Edited by L. Bywater. Oxford: Oxford University Press, 1894.

Aristotle. *Eudemian Ethics Books I, II, and VIII*. Translated with commentary by Michael Woods. Oxford: Clarendon Press, 1992.

Aristotle. *Metaphysics*. Translated by C. D. C. Reeve. Indianapolis: Hackett, 2016.

Aristotle. *Nichomachean Ethics*. Translated by C. D. C. Reeve. Indianapolis: Hackett, 2014.

Arpaly, Nomy, and Timothy Schroeder. "Praise, Blame, and the Whole Self." *Philosophical Studies* 93, no. 2 (1999): 161-188.

Baima, Nicholas. "Philosopher Rulers and False Beliefs." *Ancient Philosophy* 37, no. 1 (2017): 19-37.

Baima, Nicholas. "*Republic* 382A-D: On the Dangers and Benefits of Falsehood." *Classical Philology* 112 (2017): 1-19.

Balaban, Oded. "The Moral Intellectualism of Plato's Socrates: The Case of the *Hippias Minor*." *Bochumer Philosophisches Jahrbuch Für Antike Und Mittelalter* 13 (2009): 1-14.

Baltzly, Dirk. "Knowledge and Belief in *Republic* V." *Archiv für Geschichte der Philosophie* 79 (1997): 239-72.

Bambrough, Renford. "Socratic Paradox." *The Philosophical Quarterly* 10, no. 41 (1960): 289-300.

Bizoń, Michał, and Kamil Sokołowski. "A Note to *Protagoras* 353de." *Phronesis* 57, no. 4 (2012): 319-31.

Bonasio, Giuilia. "*Kalokagathia* and the Unity of the Virtues in the *Eudemian Ethics*." *Apeiron* 53, no. 1 (2020): 27-57.

Brickhouse, Thomas, and Nicholas Smith. "Justice and Dishonesty in Plato's *Republic*." *Southern Journal of Philosophy* 21 (1983): 79-96.

Brickhouse, Thomas C., and Nicholas D. Smith. *Socratic Moral Psychology*. Cambridge: Cambridge University Press, 2010.

Brickhouse, Thomas C., and Nicholas D. Smith. "Socrates on Punishment and the Law: *Apology* 25c5-26b2." In *Soul and Mind in Greek Thought: Psychological Issues in Plato and Aristotle*, edited by Marcelo D. Boeri, Yasuhira Y. Kanayama, and Jorge Mittelman, 37-53. New York: Springer, 2018.

Bronstein, David. *Aristotle on Knowledge and Learning: The Posterior Analytics*. Oxford: Oxford University Press, 2016.

Buckels, Christopher. "Triangles, Tropes, and τὰ τοιαῦτα: A Platonic Trope Theory." *Plato Journal: The Journal of the International Plato Society* 18 (2018): 9-24.

Burnyeat, Myles. "Plato on Why Mathematics is Good for the Soul." In *Mathematics and Necessity: Essays in the History of Philosophy*, edited by T. Smiley, 1-81. Oxford: Oxford University Press, 2000.

Burnyeat, Myles. "*Epistēmē*." In *Epistēmē, etc.: Essays in Honour of Jonathan Barnes*, edited by Ben Morison and Katerina Ierodiakonou, 3-29. Oxford: Oxford University Press, 2011.

Butler, Travis. "Identity and Infallibility in Plato's Epistemology." *Apeiron* 39, no. 1 (2006): 1-25.

Büttner, Stefan. "The Tripartition of the Soul in Plato's *Republic*." In *New Essays on Plato: Language and Thought in Fourth-Century Greek Philosophy*, edited by Fritz-Gregor Herrmann, 75-94. Swansea: The Classical Press of Wales, 2006.

Cagnoli Fiecconi, Elena. "Aristotle on the Structure of Akratic Action." *Phronesis* 63, no. 3 (2018): 229-256.

Callard, Agnes. "Ignorance and Akrasia-Denial in the *Protagoras*." *Oxford Studies in Ancient Philosophy* 47 (2014): 31-80.

Callard, Agnes. "Akratics as Hedonists: *Protagoras* 352b-355a." *Ancient Philosophy* 36, no. 1 (2016): 47-64.

Callard, Agnes. "*Enkratēs Phronimos*." *Archiv für Geschichte der Philosophie* 99, no. 1 (2017): 31-63.

Carelli, Paul. "Power and Character in Plato's *Hippias Minor*." *Ancient Philosophy* 36, no. 1 (2016): 65-79.

Clark, Justin. "The Strength of Knowledge in Plato's *Protagoras*." *Ancient Philosophy* 32, no. 2 (2012): 237-55.

Cooper, Neil. "Between Knowledge and Ignorance." *Phronesis* 31, no. 3 (1986): 229-42.

Crombie, I. M. *An Examination of Plato's Doctrines: II. Plato on Knowledge and Reality*. Abingdon, UK: Routledge, 1963.

Cross, R. C., and A. D. Woozley. *Plato's* Republic*: A Philosophical Commentary*. London: Macmillan, 1964.

Dimas, Panos. "Good and Pleasure in the *Protagoras*." *Ancient Philosophy* 28, no. 2 (2008): 253-84.

Duncan, Roger. "Courage in Plato's *Protagoras*." *Phronesis* 23, no. 3 (1978): 216-28.

Dyson, M. "Knowledge and Hedonism in Plato's *Protagoras*." *Journal of Hellenic Studies* 96 (1976): 32-45.

Fine, Gail. "Knowledge and Belief in *Republic* V." *Archiv für Geschichte der Philosophie* 60 (1978): 121-39.

Fine, Gail. "Knowledge and Belief in *Republic* V-VII." In *Companions to Ancient Thought*, I. *Epistemology*, edited by S. Everson, 85-115. Cambridge: Cambridge University Press, 1990.

Frede, Dorothea. "The Impossibility of Perfection: Socrates' Criticism of Simonides' Poem in the *Protagoras*." *The Review of Metaphysics* 39, no. 4 (1986): 729-53.

Frede, Dorothea. "On the So-Called Common Books of the *Eudemian* and the *Nicomachean Ethics*." *Phronesis* 64, no. 1 (2019): 84-116.

Frede, Michael. *A Free Will: Origins of the Notion in Ancient Thought*. Berkeley: University of California Press, 2011.

Fricker, Miranda. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford: Oxford University Press, 2007.

Gentzler, Jyl. "How to Know the Good: The Moral Epistemology of Plato's *Republic*." *The Philosophical Review* 114, no. 4 (2005): 469-96.

Gerson, Lloyd P. "Plato on Virtue, Knowledge, and the Unity of Goodness." In *Essays in Ancient Greek Philosophy III: Plato*, edited by John P. Anton and Anthony Preus, 85-100. Albany, NY: State University of New York Press, 1989.

Gerson, Lloyd P. *Knowing Persons: A Study in Plato*. Oxford: Oxford University Press, 2003.

Gettier, Edmund. "Is Justified True Belief Knowledge?" *Analysis* 23, no. 6 (1963): 121-3.

Gill, Christopher. "Plato on Falsehood—Not Fiction." In *Lies and Fiction in the Ancient World*, edited by Christopher Gill and T. P. Wiseman, 39-87. Exeter, UK: Exeter University Press, 1993.

Goldberg, Larry. *A Commentary on Plato's* Protagoras. New York: P. Lang, 1983.

Gonzalez, F. "Propositions or Objects? A Critique of Gail Fine on Knowledge and Belief in *Republic* V." *Phronesis* 41 (1996): 245-75.

Gosling, J. C. B. "*Republic* Book V: τὰ πολλὰ καλά etc." *Phronesis* 5, no. 2 (1960): 116-28.

Gosling, J. C. B. "Δόξα and Δύναμις in Plato's *Republic*." *Phronesis* 13, no. 2 (1968): 119-30.

Gosling, J. C. B., and C. C. W. Taylor. *The Greeks on Pleasure*. Oxford: Oxford University Press, 1982.

Hackforth, R. "Hedonism in Plato's *Protagoras*." *Classical Quarterly* 22, no. 1 (1928): 39-42.

Haden, James. "On the *Hippias Minor*: Achilles, Odysseus, Socrates." In *Plato's Dialogues - The Dialogical Approach*, edited by Richard Hart and Victorino Tejera, 143-68. Vol. 46. Studies in the History of Philosophy. New York: Edwin Mellen Press, 1997.

Hartman, Margaret. "How the Inadequate Models for Virtue in the *Protagoras* Illuminate Socrates' View of the Unity of the Virtues." *Apeiron* 18, no. 2 (1984): 110-17.

Hemmenway, Scott R. "Sophistry Exposed: Socrates on the Unity of Virtue in the *Protagoras*." *Ancient Philosophy* 16, no. 1 (1996): 1-23.

Hoerber, Robert G. "Plato's Lesser Hippias." *Phronesis* 7, no. 1 (1962): 121-31.

Hulme Kozey, Emily. "The Good-Directedness of Τέχνη and the Status of Rhetoric in the Platonic Dialogues." *Apeiron* 52, no. 3 (2019): 223-44.

Irwin, Terence. *Plato's Ethics*. Oxford: Oxford University Press, 1995.

Jackson, H. "*Eudemian Ethics, Θ* i, ii, 1246a26-1248b7." *Journal of Philology* 32 (1913): 170-221.

Jaeger, Werner Wilhelm. *Aristotle: Fundamentals of the History of His Development*. Oxford: Clarendon Press, 1934.

Jones, Russell E., and Ravi Sharma. "The Wandering Hero of the *Hippias Minor*: Socrates on Virtue and Craft." *Classical Philology* 112, no. 2 (2017): 113-37.

Kahn, Charles H. *Plato and the Socratic Dialogue: The Philosophical Use of a Literary Form*. Cambridge: Cambridge University Press, 1996.

Kahn, Charles H. "Socrates and Hedonism." In *Plato's* Protagoras*: Proceedings of the Third Symposium Platonicum Pragense*, edited by Aleš Havlíček and Filip Karfík, 165-75. Prague: OIKOYMENH, 2003.

Kenny, Anthony. *The Aristotelian Ethics*. Oxford: Oxford University Press, 1978.

Ketchum, Richard J. "Plato on the Unknowability of the Sensible World." *History of Philosophy Quarterly* 4, no. 3 (1987): 291-305.

Lampert, Laurence. "Socrates' Defense of Polytropic Odysseus: Lying and Wrong-doing in Plato's *Lesser Hippias*." *The Review of Politics* 64, no. 2 (2002): 231-59.

Lévystone, David. "La Figure d'Ulysse Chez Les Socratiques : Socrate *polutropos*." *Phronesis* 50, no. 3 (2005): 181-214.

Lorenz, Hendrick. *The Brute Within: Appetitive Desire in Plato and Aristotle*. Oxford: Oxford University Press, 2006.

Mann, Wolfgang-Rainer, and Vanessa de Harven. "Pleasure, Pain, and the Unity of the Soul in Plato's *Protagoras*." In *Pleasure and Pain in Classical Times*, edited by W. V. Harris, 111-38. Leiden: Brill, 2018.

McCoy, Marina Berzins. "Protagoras on Human Nature, Wisdom, and the Good: The Great Speech and the Hedonism of Plato's Protagoras." *Ancient Philosophy* 18, no. 1 (1998): 21-39.

McTighe, Kevin. "Socrates on Desire for the Good and the Involuntariness of Wrongdoing: *Gorgias* 466a-468e." *Phronesis* 29, no. 3 (1984): 193-236.

Mingay, J. M. "Some Controversial Passages in the *Eudemian Ethics*." In *Untersuchungen zur* Eudemischen Ethik*, Akten des 5. Symposium Aristotelicum*, edited by P. Moraux and D. Harlfinger, 51-62. Berlin: De Gruyter, 1971.

Moraux, P. "Das Fragment VIII 1: Text und Interpretation." In *Untersuchungen zur* Eudemischen Ethik*, Akten des 5. Symposium Aristotelicum*, edited by P. Moraux and D. Harlfinger, 253-84. Berlin: De Gruyter, 1971.

Morris, Michael. "Akrasia in the *Protagoras* and the *Republic*." *Phronesis* 51, no. 3 (2006): 195-229.

Moss, Jessica. "Appearances and Calculations: Plato's Division of the Soul." *Oxford Studies in Ancient Philosophy* 34 (2008): 35-68.

Moss, Jessica. "Hedonism and the Divided Soul in Plato's *Protagoras*." *Archiv für Geschichte der Philosophy* 96, no. 3 (2014): 285-319.

Moss, Jessica. "Right Reason in Plato and Aristotle: On the Meaning of *Logos*." *Phronesis* 59, no. 3 (2014): 181-230.

Moss, Jessica, and Whitney Schwab. "The Birth of Belief." *Journal of the History of Philosophy* 57, no. 1 (2019): 1-32.

Mueller, Ian. "Mathematical Method and Philosophical Truth." In *The Cambridge Companion to Plato*, edited by Richard Kraut, 170-99. Cambridge, UK: Cambridge University Press, 1992.

Mulhern, J. J. "TROPOS and POLUTROPIA in Plato's *Hippias Minor*." *Phoenix* 22, no. 4 (1968): 283-88.

Naddaff, Ramona. "Reckless Speech: On Plato's Theory of Lies in the *Hippias Minor*." *Arethusa* 50, no. 2 (2017): 193-219.

Nussbaum, Martha. *The Fragility of Goodness: Luck and Ethics in Greek Tragedy and Philosophy*. Cambridge, UK: Cambridge University Press, 1986.

O'Brien, Michael J. *The Socratic Paradoxes and the Greek Mind*. Chapel Hill, NC: University of North Carolina Press, 1967.

Page, Carl. "The Truth about Lies in Plato's *Republic*." *Ancient Philosophy* 11 (1981): 1-33.

Papapetrou, Peter D. "The philosopher Socrates had exophthalmos (a term coined by Plato) and probably Graves' disease." *Hormones* 14, no. 1 (2015): 167-71.

Pasnau, Robert. "Where Socratic Akrasia Meets the Platonic Good." *Journal of the History of Philosophy* 59, no. 1 (2021): 1-21.

Pearson, Giles. "*Phronēsis* as a Mean in the *Eudemian Ethics*." *Oxford Studies in Ancient Philosophy* 32 (2007): 273-95.

Penner, Terry. "Desire and Power in Socrates: The Argument of *Gorgias* 466A-468E that Orators and Tyrants Have No Power in the City." *Apeiron* 24, no. 3 (1991): 147-202.

Penner, Terry. "Socrates on the Strength of Knowledge: *Protagoras* 351B-357E." *Archiv für Geschichte der Philosophie* 79, no. 2 (1997): 117-49.

Penner, Terry. "The Unity of Virtue." *Philosophical Review* 82, no. 1 (1973): 35-68.

Plato. *Opera Volume I: Euthyphro, Apologia Socratis, Crito, Phaedo, Cratylus, Sophista, Politicus, Theaetetus*. Edited by E. A. Duke, W. F. Hicken, W. S. M. Nicoll, D. B. Robinson, and J. C. G. Strachan. Oxford: Oxford University Press, 1995.

Plato. *Opera Volume III: Theages, Charmides, Laches, Lysis, Euthydemus, Protagoras, Gorgias, Meno, Hippias Maior, Hippias Minor, Io, Menexenus*. Edited by J. Burnet. Oxford: Oxford University Press, 1922.

Plato. *Rempublicam*. Edited by S. R. Slings. Oxford: Oxford University Press, 2003.

Plato. *Complete Works*. Edited by John M. Cooper. Indianapolis: Hackett, 1997.

Plato. *Protagoras*. Introduction, notes, and appendices by J. Adam and A. M. Adam. Cambridge: Cambridge University Press, 1905.

Plato. *Protagoras*. Translated with notes by C. C. W. Taylor. Oxford: Clarendon, 1976 (rev. 1991).

Plato. *Republic*. Translated by C. D. C. Reeve. Indianapolis 2004.

Richardson, Henry S. "Measurement, Pleasure, and Practical Science in Plato's *Protagoras*." *Journal of the History of Philosophy* 28, no. 1 (1990): 7-32.

Roberts, Robert C., and W. Jay Wood. *Intellectual Virtues: An Essay in Regulative Epistemology*. Oxford: Oxford University Press, 2007.

Rosen, Gideon. "Culpability and Ignorance." *Proceedings of the Aristotelian Society* 103, no. 1 (2003): 61-84.

Rosen, Gideon. "Skepticism about Moral Responsibility." *Philosophical Perspectives* 18, no. 1 (2004): 295-313.

Rowe, Christopher J. "Hedonism in the *Protagoras* again: *Protagoras*, 351b ff." In *Plato's* Protagoras*: Proceedings of the Third Symposium Platonicum Pragense*, edited by Aleš Havlíček and Filip Karfík, 133-47. Prague: OIKOYMENH, 2003.

Rowett, Catherine. "Why the Philosopher-Kings Will Believe the Noble Lie." *Oxford Studies in Ancient Philosophy* 50 (2016), 67-100.

Rudebusch, George. *Socrates, Pleasure, and Value.* Oxford: Oxford University Press, 1999.

Russell, Daniel. "Protagoras and Socrates on Courage and Pleasure." *Ancient Philosophy* 20, no. 2 (2000): 311-38.

Russell, Daniel. *Plato on Pleasure and the Good Life*. Oxford: Oxford University Press, 2005.

Santas, Gerasimos. "The Socratic Paradoxes." *Philosophical Reviewi* 73, no. 2 (1964): 147-64.

Santas, Gerasimos. "Plato's *Protagoras* and Explanations of Weakness." *Philosophical Review* 75, no. 1 (1966): 3-33.

Schaps, David M. "Socrates and the Socratics: When Wealth Became a Problem." *The Classical World* 96, no. 2 (2003): 131-57.

Schwab, Whitney. "Explanation in the Epistemology of the *Meno*." *Oxford Studies in Ancient Philosophy* 48 (2015): 1-36.

Schwab, Whitney. "Understanding *Epistēmē* in Plato's *Republic*." *Oxford Studies in Ancient Philosophy* 51 (2016): 41-85.

Sedley, David. "Philosophy, the Forms, and the Art of Ruling." In *The Cambridge Companion to Plato's* Republic, edited by G. R. F. Ferrari, 256-83. Cambridge, UK: Cambridge University Press, 2007.

Segvic, Heda. *From Protagoras to Aristotle: Essays in Ancient Moral Philosophy*. Edited by Myles Burnyeat. Princeton, NJ: Princeton University Press, 2009.

Sesonske, Alexander. "Hedonism in the *Protagoras*." *Journal of the History of Philosophy* 1, no. 1 (1963): 73-79.

Shaw, J. Clerk. *Plato's Anti-Hedonism and the* Protagoras. Cambridge: Cambridge University Press, 2015.

Simpson, David. "Administrative Lies and Philosopher-Kings." *Philosophical Inquiry* 18 (1996): 45-65.

Simpson, David. "Truth, Truthfulness and Philosophy in Plato and Nietzsche." *British Journal for the History of Philosophy* 15 (2007): 339-60.

Smyth, Herbert Weir. *Greek Grammar*. Cambridge, MA: Harvard University Press, 1920.

Sprague, Rosamond Kent. *Plato's Use of Fallacy*. Oxford: Routledge, 1962.

Stefou, Konstantinos. "Does the Good Man Do Injustice Voluntarily? In Defense of Plato's *Lesser Hippias*." *Letras Clássicas*, no. 16 (2012): 35-60.

Stocks, J. L. "The Argument of Plato, *Protagoras*, 351b-356c." *The Classical Quarterly* 7, no. 2 (1913): 100-104.

Stokes, Michael. "Plato and the Sightlovers of the *Republic*." *Apeiron* 25 (1992): 103-32.

Sullivan, J. P. "The Hedonism in Plato's *Protagoras*." *Phronesis* 6, no. 1 (1961): 10-28.

Szaif, Jan. "*Doxa* and *Epistêmê* as Modes of Acquaintance in *Republic* V." *Études Platoniciennes* 4 (2007): 253-72.

Taylor, Christopher C. W. "The Hedonism of the *Protagoras* Reconsidered." In *Plato's* Protagoras: *Proceedings of the Third Symposium Platonicum Pragense*, edited by Aleš Havlíček and Filip Karfík, 148-64. Prague: OIKOYMENH, 2003.

Taylor, Christopher C. W. "Plato's Epistemology." In *The Oxford Handbook of Plato*, edited by Gail Fine, 165-90. Oxford: Oxford University Press, 2008.

Vasiliou, Iakovos. *Aiming at Virtue in Plato*. Cambridge: Cambridge University Press, 2008.

Velleman, David. "The Guise of the Good." *Noûs* 26, no. 1 (1992): 3-26.

Vlastos, Gregory. *Socratic Studies*. Cambridge: Cambridge University Press, 1994.

Vlastos, Gregory. "Degrees of Reality in Plato." In *New Essays on Plato and Aristotle, Volume 3*, edited by Renford Bambrough, 1-20. London: Routledge, 1965.

Vlastos, Gregory. "Socrates on Acrasia." *Phoenix* 23, no. 1 (1969): 71-88.

Vlastos, Gregory. "The Unity of the Virtues in the *Protagoras*." *Review of Metaphysics* 25, no. 3 (1972): 415-58.

Vogt, Katja. *Belief and Truth: A Skeptic Reading of Plato*. Oxford 2012.

Vogt, Katja Maria. *Desiring the Good: Ancient Proposals and Contemporary Theory*. Oxford: Oxford University Press, 2017.

Weiss, Roslyn. "Ὁ Ἀγαθός As Ὁ Δυνατός in the *Hippias Minor*." *The Classical Quarterly* 31, no. 02 (1981): 287-304.

Weiss, Roslyn. "Courage, Confidence, and Wisdom in the *Protagoras*." *Ancient Philosophy* 5, no. 1 (1985): 11-24.

Weiss, Roslyn. "Ignorance, Involuntariness, and Innocence: A Reply to McTighe." *Phronesis* 30, no. 3 (1985): 314-22.

Weiss, Roslyn. "The Hedonic Calculus in the *Protagoras* and the *Phaedo*." *Journal of the History of Philosophy* 27, no. 4 (1989): 511-29.

Weiss, Roslyn. "Hedonism in the *Protagoras* and the Sophist's Guarantee." *Ancient Philosophy* 10, no. 1 (1990): 17-39.

Weiss, Roslyn. *The Socratic Paradox and Its Enemies*. Chicago: University of Chicago Press, 2006.

Wilburn, Joshua. "Akrasia and the Rule of Appetite in Plato's *Protagoras* and *Republic*." *Journal of Ancient Philosophy* 8, no. 2 (2014): 57-91.

Wilburn, Joshua. "Plato's Protagoras the Hedonist." *Classical Philology* 113, no. 3 (2016): 224-44.

Williamson, Timothy. *Knowledge and Its Limits*. Oxford: Oxford University Press, 2000.

Wolfsdorf, David. "The Ridiculousness of Being Overcome by Pleasure: *Protagoras* 352b1-358d4." *Oxford Studies in Ancient Philosophy* 31 (2006): 113-36.

Wolt, Daniel. "The (In)Voluntary in the *Timaeus* and the *Eudemian Ethics*." *Apeiron* 52, no. 3 (2019): 245-72.

Woolf, Raphael. "Truth as a Value in Plato's *Republic*." *Phronesis* 54 (2009): 9-59.

Woolf, Raphael. "Plato and the Norms of Thought." *Mind* 122 (2013): 171-216.

Zeyl, Donald J. "Socrates and Hedonism: *Protagoras* 351b-358d." *Phronesis* 25, no. 3 (1980): 250-69.