Chapter I: The Intellectualist Hypothesis

I. Aristotle and the Problem of 'Socratic Intellectualism'

II. The Method of Hypothesis

III. Hypothesis, Elenchus, Rhetoric

IV. The Hypothesis and its Consequences

I. Aristotle and the Problem of 'Socratic Intellectualism'

Interpretation of the intellectualist hypothesis has long been under the sway of Aristotle, and his reports on the ethical views of Socrates. According to Aristotle, Socrates “thought that all the virtues are kinds of wisdom [phronēseis]” (Nicomachean Ethics VI.13, 1144b19-20) or “kinds of knowledge [epistēmai]” (VI.13, 1144b28-30). The same goes for the particular virtues, or parts of virtue, such as courage and justice; thus “Socrates thought courage to be knowledge [epistêmê]” (III.8, 1116b1). Indeed he thought “that to know justice and to be just came simultaneously; for the moment that we have learned geometry or building we are builders and geometers. Therefore he inquired what virtue is, not how or from what it arises” (Eudemian Ethics I.5, 1216b2-10). This claim that virtue is knowledge -- the intellectualist hypothesis, or intellectualism -- goes, according to Aristotle, with a whole pack of paradoxical commitments, often known as 'the Socratic paradoxes'. These include the claim that all desire is for the good, that all wrongdoing is involuntary, and that akrasia, weakness of will, is impossible: "for it would be strange -- so Socrates thought -- if when knowledge was in us, something else could master it and drag it around like a slave. For Socrates used to completely reject the idea, on the grounds that there was no such thing as akrasia: for
no one acts contrary to what is best while grasping that he is doing so, but only because of ignorance" (N.E. VII.2, 1145b23-7).

There's room for debate as to whether Aristotle has in mind here primarily the historical Socrates or the character of Plato's early, 'Socratic' dialogues.¹ (My concern here, as throughout, is with the latter; and I use 'Socrates', unless otherwise specified, for Plato's literary protagonist – subject to the caveats in the Introduction.) Most scholars do take Aristotle to be making a claim true of the Socrates depicted in the early dialogues; for one thing, every one of the passages I've cited can be read as referring to Plato's Protagoras. Most also take Socrates to speak for his author here, so that the early dialogues are taken to propound a moral theory, ‘Socratic intellectualism’, which is largely Plato's doctrinal inheritance from the historical Socrates.

A fuller account of Socratic intellectualism is given in the Magna Moralia, a work included in Aristotelian corpus but apparently written by a later (and not very bright) follower. Aristotle only offers scattered remarks about Socrates, usually as a foil for his own views; it’s in the MM that these remarks are bundled together into a full account. That account is aggressively critical and unsympathetic. Socratic intellectualism is a simplistic, bizarrely optimistic theory, which Plato rightly repudiates when he introduces the tripartite theory of the soul in Republic Book IV:²

"But neither did he [Socrates] speak correctly... by making the virtues forms of knowledge, he does away with the irrational part of the soul. And in doing this, he does away with both emotions and moral character. This is why he does not treat the virtues correctly. But afterwards Plato divided the soul correctly into its rational and non-rational parts and assigned to each its appropriate virtues." (Magna Moralia 1182a17-23)

¹ According to 'Fitzgerald's canon', references by Aristotle to Socrates the character are marked by use of the article ho; of the passages I have cited, only the third, on courage, is marked in this way.
² Cf.... Pace Cooper
The denial of *akrasia*, in particular, is taken to be so obviously false that, apparently, no further refutation is required:

Owing to such considerations he did not think that there was *akrasia*. But there he was wrong. For it is absurd for us to, trusting in this argument, do away with what convincingly occurs. For people are akratic, and knowing things to be bad do them nonetheless." (*Magna Moralia* 1200b25-33)

Taken as a whole, the *Magna Moralia* reading offers the following account of Socratic intellectualism, which remains more or less standard today:

1. In Plato’s early dialogues, Socrates holds that (a) virtue is a kind of knowledge, and (b) becoming virtuous is a purely cognitive process.
2. This is because Socrates recognises (a) only desires oriented to the good and (b) only a rational soul in human beings.
3. Socrates thus denies the importance of emotional, affective and non-rational factors in virtue and moral education.
4. Socrates thus commits himself to a number of claims in flat contradiction with the phenomena: there are no good-independent desires [=2a], irrational passions never cause action [implicit in [3]], and *akrasia* never occurs.
5. Plato rejects this Socratic moral psychology in *Republic* IV, when he acknowledges irrational desires and *akrasia*, divides the soul into rational and irrational parts, and gives a non-intellectualist account of the virtues.

Now this reading should give us pause. Socrates is not known to have led a particularly sheltered life. And yet he is supposed to just insist, with no attempt to save the

---

3 Of course once an interpretation comes to be recognised as 'standard', the end is nigh. A wave of recent attacks and alternatives already includes Rowe, Brickhouse and Smith.... The present attempt is distinctive in trying to change not only the answer but the question -- from 'what exactly does Socrates believe?' to 'what exactly is Plato trying to do?'.
phenomena, that no emotional or irrational forces ever determine our agency, that akratic or weak-willed behaviour never occurs, and that philosophical discussion magically makes people good. And yet this general line of interpretation remains mainstream, even dominant today. So for instance in Plato’s Ethics, Terence Irwin says, “the doctrine we have found in the early dialogues corresponds very closely with the views that Aristotle ascribes to the historical Socrates”. Charles Kahn, after a heroic attempt to make sense of intellectualism so understood, ends up describing it as essentially 'normative' or 'protreptic' rather than as genuinely descriptive. Less sophisticated readers tend to end up suggesting that Socrates was somehow misled by the workings of his own psyche. For instance the Loeb edition of the Magna Moralia says in a footnote: "Probably his extraordinary power of self-control made him under-rate the force of passion in others". This is profoundly unsatisfying – it’s a bit like explaining Cartesian dualism by suggesting it may have been true for Descartes himself, if no one else.

But anyone turning from these ancient or modern discussions to Plato’s actual texts is in for a surprise. For there is remarkably little textual evidence in Plato’s own works for many of the key propositions of ‘Socratic intellectualism’. The early dialogues simply do not propound any view of the nature of the soul or the emotions, let alone a startlingly radical and implausible theory of them. This raises what we might call the evidentiary qualm about the Magna Moralia reading: there is no real evidence for [1b], [2b], or [3] above. Whatever Socrates and Plato are up to in the early dialogues, their agenda can hardly be to deny the importance of affect and the irrational in human psychology and morality, in favour of some kind of rationalism: for the contrast and the issue scarcely come up. There is also room to dispute both [4] and [5], and scholars

---

4 p. 242
5 Plato and the Socratic Dialogue, pp. 226-33, cf. pp. 241-75. Note that Kahn’s unitarian, and more specifically proleptic reading of the early dialogues does not really have room for a distinct ‘Socratic’ ethical theory. But what he does have to say about the determinate views expressed in the early dialogues, such as they are, is not so far from the Magna Moralia reading after all.
7 p. 242
have recently done so.\textsuperscript{8} In particular, Daniel Devereux and Thomas Brickhouse and Nicholas Smith have argued powerfully for a sophisticated conception of 'Socratic intellectualism' which gives desires, emotions and irrational motivations their due.\textsuperscript{9} Claim [5] is also suspect because of a further, philosophical qualm or reason for doubt. When Plato introduces the new moral psychology of Republic IV, it is the basis of banal and obvious observations about irrational desires, mental conflict, and akrasia. For instance, Socrates tells the memorable little story of Leontius, who had a strong (perhaps sexual) desire to look at some corpses, but was ashamed to do so. It is hardly credible that Plato was simply unaware of such shameful desires and the resultant mental conflicts until (probably in early middle age) he set about writing the Republic. If the early dialogues did propound some version of 'intellectualism', it must have been on the basis of a sophisticated understanding of such basic phenomena, and some tacit understanding of why someone like Leontius is no counterexample. But what was that explanation, and why would Plato later find it inadequate and change his mind? The Magna Moralia reading has nothing to say about any of this.\textsuperscript{10}

Now some elements of 'Socratic intellectualism' are discussed in the early dialogues. In particular, Socrates is obsessed with the core intellectualist hypothesis [1a], the proposition that virtue is knowledge. But here too, anyone turning from Aristotle (or most modern accounts) to Plato’s own texts is in for a bit of a surprise. For one thing, since the knowledge in question is understood in terms of technê, craft or skill, the hypothesis is not particularly ‘intellectualist’ in the sense of excluding all non-rational factors. Habituation, discipline and natural aptitude are, after all, standard preconditions for acquiring any technê of importance. A fortiori nothing in the early dialogues commits Plato to the absurd view that simply identifying the correct definition of a virtue would be sufficient for possessing that virtue. We are told that Protagoras believed that learning anything required both natural ability and practice (Anec. Par. I 171, 31), and we have every reason to assume that the same would be true

\textsuperscript{8} For [4], see Devereux, Bobonich, and Brickhouse and Smith; for [5], see Carone.

\textsuperscript{9}

\textsuperscript{10}
for the moral technē considered in the early dialogues -- after all, the only account of moral education on offer in the early dialogues is actually given by Protagoras.

So there is plenty of room for interpreters to construct a sophisticated intellectualism on Plato or Socrates’ behalf – that is, one which has answers to the MM's critique -- and interpreters have recently devised several versions.\footnote{Devereux, Brickhouse and Smith, and Segvic} Two possible strategies stand out. One is to read Socrates as a proto-Stoic (he was undoubtedly an inspiration and important source for Stoic moral psychology) -- acknowledging the reality of emotions, non-rational motivations, and the other phenomena he is usually alleged to deny, but insisting that these things are properly understood as cognitive, in terms of beliefs.\footnote{Cf. Segvic.} Alternatively, we might take him to insist that these non-rational psychological forces matter for behaviour (and, in turn, character) precisely by influencing or determining beliefs.\footnote{cf. Brickhouse and Smith, Bobonich on the Protagoras} (As Brickhouse and Smith point out, Plato's account of punishment does seem to presuppose that disordered appetites and passions distort our judgements about value.) Both are philosophically intriguing options, and either is preferable to simply assuming that Socrates has never met anyone like Leontius. But the texts do not give us much help in working either reading out, let alone choosing between them. 'Socratic moral psychology' (if that is the right phrase for it at all), as Plato presents it, is simply not a complete or fully determined theory. And to assume that it must be able to respond to all the problems we (or for that matter the Magna Moralia) are interested in, such as the relation of reason to the appetites, is to start treating 'Socrates' as a real philosopher to whose ideas Plato has only given us frustratingly defective access. This makes no sense unless you are prepared to commit to a very implausible reading of the early dialogues as mere commemorations, and not very informative ones, of the historical Socrates. And it risks obscuring the question of what problems Plato as author is trying to solve.

So sophisticated intellectualist readings do not really rid us of the problems created by the Magna Moralia reading. For they accept the deepest respect in which it distorts the texts, which is in treating them as expositions of theory in the first place.

---

\footnote{Devereux, Brickhouse and Smith, and Segvic.}
\footnote{Cf. Segvic.}
\footnote{cf. Brickhouse and Smith, Bobonich on the Protagoras.
Again, if we turn from the interpreters to Plato’s own texts, we will find that the spirit of his discussions of intellectualism is not at all what we would expect. In fact there is not a single dialogue in which Socrates explicitly and unequivocally expounds and defends the view that virtue is knowledge. The *Protagoras*, clearly Aristotle’s principal source, does include arguments for a related thesis, the unity of the virtues. Specifically, Socrates argues for the identity of piety and justice, temperance and wisdom, and (apparently, though this argument is cut off almost at the very beginning) justice and temperance; and, later in the dialogue, he argues that courage is wisdom about what is and is not to be feared. Taken together, these arguments would give the result that all the virtues are identical with wisdom. But the arguments are left incomplete and the conclusion is not asserted. In the latter part of the dialogue, Socrates also proposes that ‘our salvation in life’ (aka virtue, presumably) consists in an ‘art of measurement’ -- that is, a kind of knowledge. But this belongs to an argument which is distanced from Socrates himself in a number of ways, explicitly hypothetical (based on the hypothesis of hedonism, which Socrates elsewhere rejects) (358c), and presented as ad hominem. Moreover, the debate opens with Socrates presenting an argument against the claim that virtue can be taught, which he takes to entail that it isn’t knowledge. Most readers of the *Protagoras* feel that Protagoras’ ‘Great Speech’ rebuts that argument effectively, but there is no clear evidence that Socrates and Plato agree -- especially as the arguments against the teaching of virtue are recycled, and left unrefuted, in the *Meno*. Yet another difficulty is that ‘virtue’ in these early stages of the dialogue seems to denote something very different from the art of measurement introduced later on. Socrates’ own conclusion in the *Protagoras* is thoroughly aporetic: if the discussion had a voice of its own, he says, it would say that both he and Protagoras were ridiculous for changing sides, and that further investigation is needed (361b-c).

The *Meno* comes closer to fitting Aristotle’s testimony, for it does include a clear and complete argument to the effect that virtue is knowledge (86e-9a). However, Socrates then proceeds to argue against that claim, on the grounds that virtue is not taught -- here we have the recycling of the arguments from the *Protagoras*. And in the *Meno* those arguments seem to stand, and Socrates concludes that the virtue of the
Athenian statesmen must be a matter, not of knowledge, but of divine luck and right opinion.

So neither the *Meno* nor the *Protagoras* is naturally read as propounding the doctrine that virtue is knowledge. Rather, both present arguments -- in each case, inconclusive and questionable arguments -- on both sides of the question. And this is a recognised ancient genre: both dialogues are exercises in sophistic *antilogikê*, a form of argument associated with Protagoras -- and the one *certain* purpose of the *Protagoras* is to depict Socrates as doing it better than the master.14

Finally, in one dialogue, the *Euthydemus*, Socrates does argue, with no counter-argument in sight, that wisdom is the cause of happiness, for "knowledge seems to provide men not only with good fortune but also with well-doing, in every case of possession or action" (281b). But the all-important conclusion that wisdom must therefore be identical to the whole of virtue, *aretê*, is never explicitly drawn -- even though Socrates does group wisdom and virtue together in some allusions just before and after the argument (278d, 283a-b, cf. 273d-5a). (In the *Meno*, a similar argument from the necessity of wisdom for successful action is presented as a reason to identify virtue with knowledge. But then this identification is overturned by the argument that true belief is also beneficial while it lasts.) This vagueness and incompleteness is understandable, for the argument is explicitly offered as an exercise in protreptic -- that is, as an advertisement for philosophy rather than a full exposition of doctrine.

The dialogues which discuss the particular virtues are even more underwhelming as evidence for ‘Socratic intellectualism’. The relation of virtue to knowledge is certainly a central topic. The *Laches* discusses a definition of courage as wisdom; the *Charmides* discusses several definitions of moderation [*sôphrosunê*] in terms of knowledge; one definition of piety in the *Euthyphro* is as knowledge of how to sacrifice and pray. But all of these definitions are rejected for various reasons; and all these dialogues end in *aporia*, with no definition affirmed as satisfactory.

14 Plato also practices *antilogikê* in a number of other dialogues, such as the *Phaedrus*, which argues both against and in favour of *erôs*, the *Cratylus*, which argues both that there is and is not a natural ‘correctness of names’, and, as Vasilis Politis has argued, the *Charmides*. Like the intellectualist hypothesis itself, this practice is part of a debt to Protagoras far greater than is usually recognised.
As I argued in the Introduction, the aporetic form doesn’t imply that no positive results can be extracted from these dialogues. The point for now is just the preliminary one that in these cases too, we are not presented with anything that looks much like the exposition and defence of doctrine. Were it not for Aristotle, and perhaps for certain modern assumptions about what philosophical writing must do, it would strike us as natural to speak of ‘Virtue is knowledge’ not as a doctrine being defended, but as a fascinating possibility subjected by Plato to repeated and apparently inconclusive debate. Intellectualism doesn’t function in Plato like empiricism in Epicurus or materialism in Hobbes, as a dogma to be hammered home; it’s more like forgiveness in Mozart and marriage in Wagner -- a thematic obsession, a possibility endlessly returned to with different results every time. That isn't to deny, of course, that it's the object of argument in Plato. The problem is to see exactly what these complex, incomplete, overlapping, and often conflicting arguments are intended to do.

II: The Method of Hypothesis

I will try to show that we can do better than consign ‘virtue is knowledge’ to the vague status of theme or obsession. For it is discussed by Plato as a hypothesis [hypothesis] -- in fact, as the hypothesis around which the early dialogues are collectively organized. What this means can be gathered from the two dialogues in which Socrates discusses hypothesis as a philosophical method, the Meno and Phaedo.15 (Hypothesis is also prominent in the Divided Line in the Republic and in the Parmenides, but these later uses are very complicated, and raise questions I cannot address here.) As background to these, it is worth noting that in the early dialogues, a hypothesis can be quite generally

15 Much of what I have to say about hypothesis here follows the very helpful and sensible Robinson (1953); but cf. n. 11. Cf. also Benson and Wolfsdorf. See also Wolfsdorf, pp. 38-9, n. 9 for a full listing of occurrences, including the later dialogues -- though only of hypothesis, not the verb forms. Scholars seem to assume that the noun form is much more 'marked' and tied to a special Platonic sense, but I am not so sure. Cf. also Huffman for a very useful survey of pre-Platonic usage, confirming the close ties between hypothesis and archê.
any proposition set down for examination, and is mostly used to refer to definitions: Euthyphro's definitions of piety are *hupotheseis* (11c5) and set down (*hupothemenos*, 9d8); so is Gorgias' definition of rhetoric (454c4) and one of Hippias' definitions of the Beautiful (*Hipp. Maj.* 302e12).\(^{16}\) Undefended, dialectically useful truisms (e.g., 'temperance is fine', *Charm.* 160d) can also be set down or hypothesized, as can disputed or difficult points granted for the sake of argument [*ex archês hupotithemetha, Charm.* 171d3, cf. 169d3-4). All this accords with pre-Platonic uses of *hypothesis* and *hupotithemai*, on which to hypothesise is to set something down as a topic, first-principle or starting-point.\(^{17}\)

The *Meno* discussion of hypothesis arises from an impasse: Meno repeatedly asks Socrates to pronounce on whether virtue can be taught, while Socrates tries to shift the discussion to the more basic question of what virtue *is*. Socrates finally agrees to investigate Meno's question "by means of a hypothesis", the way that geometers do (86e3-4).\(^{18}\) The geometrical procedure alluded to is the one later discussed and deployed as (and probably already known as) analysis.\(^{19}\) Socrates even gives a geometrical illustration, in terms of inscribing a triangle in a given area, though this is so confusingly expressed that it only makes matters more obscure, and has precipitated an enormous scholarly literature.\(^{20}\) I will not add to that, but will just note that,
regardless of the details of the example, Plato thus presents Socrates as bringing to bear some of the rigour and sophistication of geometrical reasoning to questions of ethics. And the basic idea of investigation 'from a hypothesis', as he explains it, is an intuitive one, and emerges clearly enough. Hypothesis is a method for reducing an obscure problem to one which I know how to solve. If I want to know whether X has some property or predicate Y, I can adopt a hypothesis to the effect that if and only if it has some other property Z, it has Y. Then I can answer my question by determining whether or not X has Z. In the case at hand, Socrates proposes, all and only kinds of knowledge are teachable (87b6-c6). So if I want to know whether virtue is teachable, I can answer my question by discovering whether or not virtue is knowledge.\textsuperscript{21}

It is unfortunately left a bit hazy – and is a matter of scholarly controversy -- just which step of Socrates' procedure constitutes the hypothesis itself. In fact, there are clearly multiple hypotheses in play. The thesis that virtue is good, which Socrates uses as a starting-point to argue that virtue is knowledge, is explicitly referred to as a hypothesis at 87d2. But Meno is almost certainly referring to something else as a hypothesis at 89c3, when he says, "clearly, on our hypothesis [\textit{kata tên hupothesin}], if virtue is knowledge, it can be taught." Scholars differ, though, over whether this other hypothesis in play is 'Virtue is knowledge' or the biconditional: 'Something (and virtue in particular) is teachable iff it is knowledge'.\textsuperscript{22} The latter is probably the minority view, reinforce the allure of the promised method for potential students of the Academy, while mathematically literate Platonists would be familiar with the method already.

\textsuperscript{21} Socrates thus manages rather cunningly to get Meno to accept a shift to the question he has all along insisted is prior: What is virtue? In the dialectical context, his introduction of the method of hypothesis is a bait and switch (cf. Sharples pp. 162-3).

\textsuperscript{22} 'Virtue is knowledge' is opted for by Robinson in his second edition (1953) (under the influence of Cherniss and Friedländer), and has been followed by many interpreters since (including Scott, Benson, and Menn). Since I want to argue that this is the guiding hypothesis of Plato’s project throughout the early dialogues, it may seem perverse of me to deny that it has that status here. But the (to my mind insuperable) problem is that 'virtue is knowledge' is here \textit{deduced} as the conclusion of a complex chain of reasoning at 87d-9a; it is not something hypothesized, 'set down', at all. Robinson is thus forced to admit that his reading is "not very like Platonic hypothetical method in general as I described it in the previous chapter" (121). It might be objected that hypotheses are, as such, not supported by argument, then I cannot say that 'Virtue is knowledge' functions as a hypothesis for the early dialogues as a whole -- for Plato does
but seems to me by far the more plausible reading, for four reasons. First, to hypothesise something is literally to set it down – to posit it, as we would say, for future argument. That is why, as David Wolfsdorf has shown, hypothesis can mean something like archê, a foundational first-principle or starting-point, in pre-Platonic texts like On Ancient Medicine; and also why it is an appropriate term for the definitions thrown up and batted about in dialogues like the Charmides and Euthyphro. So there is an implicit contrast between claims which are hypothesised and those which have been deduced. And in the Meno passage, 'virtue is knowledge' has been deduced, starting from the hypothesis that virtue is good. Second, the geometrical example offered as a model is also structured around a biconditional (or, as scholars often put it, whether or not a 'limiting condition' obtains) -- indeed that is practically the only feature of it that registers to the non-geometrical reader, so one would expect Plato to put it to some use (87a2-3). Third, this reading fits much better with Meno's phrasing, noted above, at 89c3: "clearly, on our hypothesis [kata tên hupothesin], if virtue is knowledge, it can be taught." If the hypothesis were simply identical with the antecedant here, we would expect something like "clearly, if our hypothesis is true and virtue is knowledge, it can be taught". The kata seems rather to allude to the relation between the antecedant and the consequence, which is indeed governed by the biconditional relation (teachable iff knowledge). Fourth, Socrates' phrasing shortly after recalls the idea of a hypothesis once again: "That virtue is teachable if it is knowledge, I do not take back [anatithemai] as not well said; but consider whether it is reasonable of me to doubt whether it is knowledge"(). As Robinson points out, it is natural to understand the imagined 'taking back' here [anatithemai] as the retraction of a hypothesis (cf. the cognate hupotithemai), which suggests that the hypothesis must be what Socrates has just recalled: "that virtue

present arguments for it, in the Meno, implicitly in the Euthydemus, and indirectly in the Protagoras, Laches, etc. To this I can answer only that there is already a certain amount of abstraction and decontextualization involved in speaking of the hypothesis as playing that role, not always explicitly, across multiple dialogues. It is sufficient for my purposes that the arguments in favour of it are not presented as conclusive: they are reasons to take the hypothesis seriously and investigate its consequences just as Plato does, rather than final demonstrations.

23 At 87a3-7, the long geometrical conditional seems to be giving the content of the hypothesis tis which the geometer claims to have at a2.
is teachable if it is knowledge" – again, the biconditional, in effect, as applied to the problem at hand.

So we can already set out the first and most obvious feature of the Platonic ‘method of hypothesis’:

(1) A hypothesis is a thesis literally set down or posited \( \textit{tithenai, hupotithenai} \), as a heuristic starting-point – rather than being presented as, say, a conclusion, a free-standing conviction, a known first principle, or part of an induction. It is not, qua hypothesis, proven or deduced. In saying 'I hypothesise \( P \)' (‘postulate’ would be just as apt a translation), I bracket for the time being all questions about how well supported \( P \) might be -- and, for that matter, how deeply I myself am committed to it. That is why the definitions scrutinized in dialogues like the \textit{Euthyphro} can be described as hypotheses: they are dialectical proposals, put forward without any particular argument or justification, as a basis for discussion. (Note that it is perfectly acceptable for a Euthyphro or a Charmides to put forward one definition after another as they fall flat; Socrates insists that his interlocutor do his best, and say what he honestly thinks, but clearly there can be no assumption that each of these hypotheses is a matter of deep commitment, let alone something \textit{known}.) Thus the identity of some thesis as a hypothesis is situational; and hypotheses are not restricted to any particular kind of content. Obviously some claims are more suited to this role of dialectical starting-point than others, though. Existential propositions positing some basic entity -- an \textit{archê}, or some cause \( \textit{aitia} \) whose effects can then be traced -- are natural candidates (as we will see in the \textit{Phaedo}). At the other extreme, one would be unlikely to 'hypothesise' something trivial or known to be false; and a starting point which was self-evident to all would hardly need to be flagged as a hypothesis. In the \textit{Phaedo}, as to the selection of hypotheses, Socrates will say only that we select whatever thesis seems 'strongest': presumably that involves both explanatory power, sufficient to ground the consequences we are interested in, and some independent plausibility.

To see what other features can be added we need to turn to this more complex use of hypothesis in the \textit{Phaedo}. This follows the famous account of Socrates’ youthful enthusiasm for natural science, and his eventual disillusion. In describing his own
second-best alternative, the so-called ‘second sailing’, he refers to it as proceeding by hypotheses. The passage is controversial, and puzzling on many points of detail; I will note some points of interest and puzzlement en passant in the notes:

I thought I must take refuge in discussions and investigate the truth of things by means of words.... I started in this manner: taking as my hypothesis [hupothemenos] in each case the account [logos]²⁴ that seemed to me the strongest [errômenestatos], I would consider as true, about cause and everything else, whatever harmonized [sumphônein]²⁵ with this, and as untrue whatever did not so agree... (99e4-100a5)

Socrates observes that he has left his audience behind, and offers to explain:

This, he said, is what I mean. It is nothing new, but what I have never stopped talking about, both elsewhere²⁶ and in the earlier part of our conversation.²⁷ I am going to try to show you the kind of cause [tês aittias to eidos] with which I have concerned myself. I turn back to these oft-mentioned things and start [archomai] from them.²⁸ I set down [hupothemenos] the existence of a Beautiful, itself by itself, of a Good and

²⁴ So what one hypothesises is a proposition, a logos: that a certain Form exists, not just the Form itself. This is confirmed at 100b5, where Socrates hypothesises that the Forms are.
²⁵ sumphônein means 'harmonize', and is often used literally in musical context; in a logical context, it is thus a rather vague term for Plato to select here, suggesting something more positive than compossibility or coherence but weaker than entailment. Cf. on 101d5 below.
²⁶ So the method to be introduced (and the kind of entities to be hypothesised) are not new -- they are familiar to Socrates’ audience, and I take it that Plato only makes this sort of thing explicit when the point is to tell us, the reader (with a delicate breach of the fourth wall), that they should be familiar from other works.
²⁷ Presumably this reference is to the recent refutation of the claim that the soul is a harmony (labelled a hupothesis at 94b1). This was shown to conflict with the theory of recollection (also a hupothesis, 92d6), which depends on the soul pre-existing its body. Socrates repeatedly puns on the inharmonious results of the harmonia theory at 92c-d, using terms loosely reminiscent of sumphônia (sunaisetai c3, sunôidôi c5, sunôidos c8).
²⁸ Again, a hupothesis is archê-like – it is something to start from.
a Great and all the rest. If you grant me these and agree that they exist, I hope to show you the cause as a result, and to find the soul to be immortal. (100b5-9).

The ultimate demonstrandum here is the immortality of the soul: so in a sense that is the counterpart here to the *Meno*’s problem about geometry or the teachability of virtue, but it is a desired result rather than an open question to be resolved. In another departure from the *Meno*, Socrates also goes on to sketch a pair of further procedures in which the hypothesis itself, it seems, becomes the object of explanation and argument:

But afraid, as they say, of your own shadow and your inexperience, would cling to the safety of your own hypothesis [*hypothesis*] and give that answer. If someone then attacked your hypothesis itself, you would ignore him and would not answer until you had examined whether the consequences that proceed [*hormēthenta*] from it harmonize or disagree with one another [allēlois sumphônei ê diaphônei]. And when you must give an account of your hypothesis itself you will proceed in the same way: you will assume another hypothesis, the one which seems to you best of the higher ones until you come to something sufficient [*hikanon*], but you will not jumble the two as the contradictors

---

29 Use of the Forms is here being contrasted with causal explanation through material factors etc. (‘it is beautiful because of the bright colour’). How to generalize from this kind of case to other kinds of hypothesis (and other rival explanatory strategies) is a difficult question.

30 I take it that these 'consequences' (literally just 'things which have proceeded') are the things earlier set down as harmonizing with the hypothesis. The verb *hormaô* means to be put in motion from some starting point (again, the hypothesis is thus represented as an *archê*). Once again Plato’s terminology is deliberately blurry, even by pre-Aristotelian standards, as to the precise logical relations here.

31 Are harmony and disagreement meant to be exhaustive alternatives? If so, it is tempting to think that *diaphônía* must be logical incompatibility, and *sumphônía* compatibility. But that seems to give too weak a sense to *sumphônei* in the earlier occurrence (cf. n. 25 above).
[antilogikoi] do, by discussing the hypothesis and its consequences [hòrmêmêna] at the same time, if you wish to discover some one of the things that are. This they do not discuss at all nor give any thought to, but their wisdom enables them to mix everything up and yet to be pleased with themselves; but if you are a philosopher I think you will do as I say. (101c9-2a1)

This passage certainly raises more questions than it answers: what is it for something to 'follow' from a hypothesis? What makes a hypothesis sufficient? Who are these 'contradictors', and what exactly are they doing wrong? It is also unclear how far this

32 Antilogikê, the art of arguing on both sides of a question, was associated with Protagoras; and we see it practised by Euthydemus and his brother in the Euthydemus. So we would expect the criticism here to be aimed at some of the kinds of sophistic argument we see depicted in the early dialogues, picking up on the critique of 'those who spend their time on contradictory arguments' [hoi peri tous antilogikous logous diatripsantes] at Phaedo 90c1. The problem is to see how the procedure criticised here actually corresponds to anything sophists are represented as doing in our texts (Platonic or otherwise).

33 Again, how and where do the antilogikoi do that? What is Socrates envisaging? It depends very much on what sort of 'consequences' he has in mind, which is far from clear.

34 Grube* has "discover the truth", but perhaps Socrates' point is that we are using the hypothesis of (certain) Forms to identify the nature of another reality, the soul. (If the idea of a Form of soul sounds odd, say instead: what soul is by nature, as such, or essentially.) The antilogikoi are not interested in such projects of definition: their mode of argument depends on exploiting latent ambiguities rather than clarifying them.

35 This objection to 'jumbling together' hypothesis and consequence is particularly puzzling and intriguing. The idea that consideration of the hypothesis itself should be distinct from (and come after) discussion of the consequences fits well with the current argument in the Phaedo (where Socrates asks to simply be granted the existence of Forms), but not so well with cases (envisaged here, and I will argue, practiced throughout the corpus) in which the hypothesis itself is being investigated through its consequences. The most prominent hypothesis-consequence pairs in the ethical dialogues are various definitions of the virtues and the further propositions which follow (e.g., that virtue is teachable, or the virtues are one) -- and it is typically Socrates who moves back and forth between these. Conceivably there is actually a criticism here of Plato's own earlier method in some dialogues, and we are to think of cases like the pair of conflicting arguments [antilogikê] in the latter part of the Meno. These 'discuss the hypothesis and its consequences at the same time' in the sense that the propositions (1) virtue is knowledge and (2) virtue is teachable are taken to be up for
is the same procedure as that depicted in the Meno; the Phaedo seems to present both a new level of methodological reflectiveness and to some extent a new method. But there are also clear continuities with the Meno, and with what Plato has been doing all along -- as Socrates himself seems here to assure us (). Understood along these broad lines, and incorporating the picture of the Meno as far as possible, Plato's (pre-Republic) hypothetical method would seem to have five further salient features:\(^{36}\)

(2) The hypothesis is used as a starting-point for the deduction of consequences -- literally, 'things that follow' [hormêthenta]. This is, in the Phaedo, initially phrased in terms of setting down what ‘harmonizes’ or ‘agrees’ [sumphônein] with the hypothesis as true, and whatever doesn’t so ‘agree’ as not true (100a). But mere compatibility, if that is what ‘agreement’ amounts to, clearly would include too much; and the language of 'following' [hormêthenta and hôrmêmêna] suggests a closer relation. Socrates' terminology is consistently vague and non-technical, however; and this is perhaps deliberate, marking the fact that the relation must fall somewhat short of strict logical entailment.\(^{37}\) For in most cases little or nothing will be entailed by a hypothesis alone -- certainly nothing interesting follows from Socrates' biconditional in the Meno without some further premises about what things can be known or taught. So use of a hypothesis must involve the addition of auxiliary premises. Insofar as these are uncontroversial, or supported by strong independent arguments, the hypothesis can still be taken to bear full responsibility for the consequences that ‘follow’, and to stand or fall by their acceptability. So I will follow Plato in speaking of consequences as

grabs simultaneously: Socrates infers (2) from (1) and then, arguing on empirical grounds against (2), infers the denial of (1). If this suggestion is right, then risk in discussing a hypothesis and a consequence at the same time is that one might be taking some key term in different senses in each, as Socrates does in these arguments. The same sort of thing happens when, for instance, Protagoras protests that the virtues are not a unity, since one can have courage without any of the others; this is true of virtue conventionally understood, but not of the virtue which consists in knowledge.

\(^{36}\) cf. Robinson

\(^{37}\) Socrates does not seem to use sumbainein for the hypothesis-consequence relation, which is his preferred term for entailment.*
‘following’ from the hypotheses alone, when the relation is really what we might call 'not-quite-entailment'.

(3) This procedure of deduction from a postulate may be used for at least three different kinds of inquiry: to settle an open question, to establish some demonstrandum, or to investigate the hypothesis itself. This is where the *Meno* and *Phaedo* present different pictures. In the *Meno*, hypothetical reasoning takes the form of deducing the solution to a problem – answering the question, can virtue be taught? or, can the triangle be inscribed? In the *Phaedo* itself, the point is the deduction of a desired result -- the immortality of the soul. But Socrates’ discussion, as we have seen, alludes to further adaptations of the method. He depicts the hypothesis itself becoming the object of investigation and testing in its own right: we are to deduce multiple consequences from the hypothesis, and test it by investigating whether they agree with or contradict each other.38

(4) The *Phaedo* emphasises that in the third sort of operation, we are to check the consequences for some kind of internal coherence or the lack of it -- whether they 'agree' or 'disagree' with each other. How to take this is not quite clear. It is tempting to take it as a reference to *consistency*, so that the claim is that we search for *contradictions* among the consequences, and, presumably, reject the hypothesis if we find one. This is a very appealing picture for at least two reasons. First, it makes the method of hypothesis look very much like the Socratic *elenchus*; given Socrates' encouragement in the *Phaedo* to view the method as familiar, we might at this point wonder whether we are really dealing with two distinct methods at all. (I will discuss this in the next section.) Second, this reading gives a sharp contrast with the rival contemporary methods of argument, making it much easier to see why Plato would

38 Robinson has a good discussion of these different conceptions of the method. He concludes that the *Meno* version, in which the hypothesis is a tool for the deduction of consequences, is prior. The *Phaedo* version, in which the hypothesis itself is the demonstrandum or refutandum, evolves from the other through the practice of the Socratic elenchus. Cf. Robinson (1953), pp. 110-3.
find it so valuable (see section III). For it entails that we cannot reject a hypothesis simply on the basis of considerations about *to eikos* -- what is 'plausible' or 'likely' or 'reasonable to suppose' -- as the lawyers and rhetoricians do, but are to look to strictly logical relations -- entailment, possibility and impossibility.

At the same time, it is hard to accept that contradiction-hunting is the whole point of the exercise. The problem is not just that this is only one of several operations for which a hypothesis might be used; even when we confine ourselves to the third form of the method, in which multiple consequences are adduced and the hypothesis itself comes under scrutiny, this conception seems too narrow. For one thing, as noted already, Plato's terminology is here apparently deliberately vague and broad. As already noted, the 'consequences' of a hypothesis are not restricted to those which are logically entailed by the hypothesis alone; perhaps, as Hugh Benson has suggested, they extend to "all those beliefs, assumptions, or common opinions (*endoxa*) appropriately associated with ... <the>... component concepts of the hypothesis" -- everything we will believe on the relevant question if we add the hypothesis to the set of our prior beliefs. (Cf., perhaps, the uncontroversial 'auxiliary premises' in the case of the

---

39 It might be objected that this contrast is misleading; after all, if we can add in auxiliary premises at will, then whenever the hypothesis seems to sit ill with some intuition, we can add in that intuition as a premise to get the required contradiction. This is still different, though, from just assessing a thesis on the basis of general plausibility. For the method will at least force us to get clear about exactly what our hypothesis does and does not entail, and exactly which of our other commitments excludes it.

40 The contrast between what is deduced from a hypothesis and the merely plausible is explicit at *Phaedo* 92c-d: cf. also on the *Laches* in section III. It is worth pondering the fact that the crucial argument in the *Meno* for the unteachability of virtue is merely an argument in terms of *to eikos*. For it shows only that we might 'reasonably' [*eikotôs*] doubt that virtue is knowledge (*89d6*), and that we "would conjecture rightly in conjecturing" that if there are no teachers of it, it cannot be taught [*kalôs eikazontes eikazoimen*, an odd and emphatic phrase also involving cognates of *eikôs*] (*89e2*). Moreover, the line of argument is one familiar from rhetorical texts (for instance *Gorgias*' *Defense of Palamedes* or *Laches*' speech in the *Laches*): Socrates' claim is that the effects we would reasonably expect to hold if the hypothesis were true cannot in fact be observed. If virtue can be taught, there *should* be agreement that it can be taught, there should be some generally recognised teachers of it, and those who are best able to teach it should be seen to do so -- none of which is the case.

And if ‘harmonizing’ is something in between consistency and entailment, there might be kinds of ‘disharmony’ weaker than contradiction that would serve to scuttle a hypothesis.

My suggestion, then, is that we should construe the ‘method of hypothesis’ broadly, as a practice of making explicit and testing inferential and conceptual connections, so as to make concepts and their interrelations clear. Consider the case of the *Meno* use of hypothesis, where the hypothesis that virtue is knowledge iff it is teachable does apparently entail a contradiction. For it is used to argue first that virtue (since it must be knowledge) is teachable; and then that (since there are no teachers of it) it is not knowledge. Even here contradiction seems not to be decisive or an end in itself, for Socrates shows no inclination to conclude that we should reject the hypothesis. Instead (to be dogmatic about some points I will consider in more detail later on), we are offered a different sort of solution -- or rather two solutions, one explicit and one implicit. The explicit solution is that the latter line of argument wins: we should reject the claim that virtue is knowledge, since the argument which supported it has been shown to be faulty. Knowledge is not the only psychological state which leads to reliably good results: true belief is just as good while it lasts. But implicitly, Socrates' concluding remarks on a Teiresias-like ideal of virtue (*Meno* 99e-100a) gesture towards the prospect of a virtue which *would* consist in knowledge. And that hint in turn should prod us to remember that it has not really been shown that virtue *cannot* be taught -- only that it does not seem to be 'taught' by conventional methods. That in turn should prompt us to recall that real learning is recollection; which means that real teaching is the kind of dialectical education Socrates has just acted out with the slave. And all bets are off as to whether and how virtue might be taught, or at least teachable, in that way. The bottom line is that 'virtue' and 'teaching' have both turned out to be ambiguous between a higher and a lower sense: and the exceptionally rare, higher virtue of a 'Teiresias', which does consist in knowledge, may perfectly well be teachable -- we have even had a glimpse of how it might be taught.
So if it is fair to take this part of the *Meno* as a self-conscious case study in the method of hypothesis, then Plato is showing us that the real value of the method is in the clarification of concepts. The eliciting of a contradiction is here primarily a means to a higher end -- a provocation and a prompt to disambiguation, rather than a decisive result in itself. And it is easy to see how other means to that end might work just as well: after all, if we elicit a raft of consequences and no contradiction is found among them, we will still learn a great deal about what the hypothesis must have meant all along.

So I will in what follows use 'the method of hypothesis' in a broad sense, to include all investigations in which Plato's aim seems to be to tease out through deductive argument the meanings of the central terms of the hypothesis set down through the investigation of their consequences, whether they involve (or even seem to aim at) the discovery of contradictions or not.

(5) A hypothesis can be supported by being derived from a higher hypothesis. The *Phaedo* is again rather vague about how this works, saying only that we should identify something 'adequate' or 'sufficient' for that role; but the idea seems to be that hypotheses can be strung together in deductive chains -- each somehow explained and supported by the hypothesis above and tested through the consequences or lower hypotheses below. Like being a hypothesis, the status of a higher hypothesis is apparently -- at least to some extent -- situational. That is, a higher hypothesis could be any thesis which, like the initial hypothesis, recommends itself as plausible, and could be used to derive the hypothesis which follows, where this again must mean a relation of not-quite-entailment. An example would perhaps be the *Protagoras*, where 'Knowledge is sovereign' (which has the look of a hypothesis, being asserted by Protagoras' consent, without argument) is supposed to be grounded by the thesis that we all pursue pleasure as the good. So the higher hypothesis here would be

42 Note too that the *Phaedo* never actually says that the discovery of diaphônia among the consequences is a reason to discard one's hypothesis. It may be hard to imagine another upshot, or reason to look for such 'disagreements', but this is nowhere explicit; it seems a natural supposition only because we are reading in something like Socrates’ use of the elenchus in the early dialogues.
psychological hedonism -- and perhaps, 'above' it again, that all desire is for the good (whatever it might be). But we might suspect that a higher hypothesis must also be in some way prior to the initial hypothesis -- say, if the two deal with a cause and its effects respectively, or with explanation and explanandum, or with the essence of a thing and its other properties (the 'priority of definition'). When the Phaedo talks about looking for a second hypothesis 'among those higher', it seems to be alluding to an antecedent metaphysical or epistemic hierarchy.\(^{43}\)

In the Divided Line of the Republic, Socrates speaks of philosophical dialectic as proceeding all the way up to an 'unhypothetical first principle' (510b-11e): the passage is notoriously puzzling and open to interpretation, but it at least usefully confirms that to use a hypothesis as a hypothesis is to treat it as a provisional starting-point (511b). The Phaedo's exhortation to move 'upwards' until you come to something 'sufficient' points vaguely in the Republic's direction. But there is no hint there that any higher hypothesis could be 'sufficient' in the sense of being itself unhypothetical. So the Divided Line brings out by contrast the essential limitation of the method in the earlier dialogues, where there is no such deus ex machina: hypothetical reasoning, because its starting-points are unproven, can only be provisional and defeasible (as Plato argues emphatically at Cratylus 436c-e). This seems worth listing as a final feature:

(6) Both a hypothesis and the consequences which follow from it can only be provisional and defeasible. The method of hypothesis can demonstrate nothing, at least nothing positive. To really prove a hypothesis, one would have to somehow establish the truth of one's highest hypothesis, thus rendering it something other than a hypothesis; and even that could only be sufficient to prove the initial hypothesis if one could also sharpen up the not-quite-entailment relation into the real thing. In the absence of all this, hypothesis is only a method for clarifying the menu of philosophical options, by showing which positions follow from sets of other propositions. At most, it can exclude some sets -- some combinations of options on the philosophical menu -- by

\(^{43}\) But then what of the way in which the Meno argues from 'virtue is good' to 'virtue is knowledge'? This line of argument violates the 'priority of definition', so cannot be a path to knowledge; does that mean that 'virtue is good' cannot count as a genuine higher hypothesis here? (It is referred to as a hypothesis at 87d3.)
showing that they generate contradictions. Hypothetical reasoning is thus not a method of demonstration (as Aristotle would put it), but a vehicle for tracing networks of entailments and conceptual connections, and detecting options which can be ruled out.

III. Hypothesis, Elenchus, Rhetoric

At this point the method may well sound both modest and rather amorphous. To see what is distinctive about it, and how it relates to Plato's actual argumentative practice, we need to bring it into connection with two other concepts. The first of these is the so-called 'Socratic elenchus' or method of refutation.

The method of the elenchus is standardly described by scholars (it is never defined by Socrates, who indeed does not even seem to have a word for it) more or less as follows. Socrates elicits a target thesis $P$ (usually a proposed definition) from his interlocutor; he adds some auxiliary premises ($Q, R, S...$), presumed to be uncontroversial, with the interlocutor's assent; and, eliciting their joint consequences, shows that they lead to a contradiction:

$$
\begin{align*}
  P \\
  Q \\
  R \\
  S \\
  (Q \land R \land S) \rightarrow \neg P \\
  \therefore P \land \neg P
\end{align*}
$$

Since a contradiction cannot be true, something has gone wrong here: and the target thesis $P$ is invariably rejected as the culprit -- though all the method can prove strictly speaking is that (if all the inferences are valid) some of the premises must be false.
This procedure of refutation should look very familiar. For it is virtually identical to the *Phaedo* proposal to test a hypothesis by checking whether its consequences ‘disagree’ with each other. (Note that both involve the same oversimplification in attributing the contradiction to the target thesis alone, ignoring the possibility that the necessary auxiliary premises might be at fault.) This at least shows that the procedure I described above -- the examination of a hypothesis by scrutinizing its consequences for coherence -- is not such a weak test as it might seem. It is in fact what we see Socrates doing repeatedly in the early dialogues, with very powerful negative results.

This might raise the question of whether there is anything much to be gained from discussing Socrates’ procedure in the early dialogues in terms of one ‘method’ rather than the other. One argument for preferring to talk of hypothesis would be the fact that Socrates does himself speak of this as a method, while the ‘method of *elenchus*’ is strictly a scholarly construct. (A later development, perhaps, but he at least refers to various definitions in the early dialogues themselves as being hypothesized.) More importantly, talk of Socratic method in terms of ‘the elenchus’ has imported a pair of assumptions which have proved dangerously distorting in the interpretation of the early dialogues. One is that the early dialogues are occupied only with refutation; the other, that those refutations are all alike in form, conforming to the template of the *elenchus* sketched above.

On any reasonably open-minded reading of the early dialogues, both these claims are false. For one thing, Socrates’ investigations clearly have a constructive side. Some of the hypotheses discussed are put forward by Socrates himself, after all, and at least some of these some of these seem to survive investigation. Socrates’ definition of piety as a part of justice in the *Euthyphro* is a good example. At least one dialogue is *primarily* devoted to constructive argument rather than refutation: namely

---

44 Cf. Robinson (1953), pp. 106-7; Benson, "it must be admitted that the difference between the method of hypothesis and *elenchos* may appear rather slight" (p. 100).
45 There is a growing rebellion against these assumptions, coming from writers with very different viewpoints: cf. Schwartz, Wolfsdorf, Scott...
46 Cf. Matt Schwartz, Wolfsdorf
the *Crito*, which is bizarrely ignored in most recent studies of Socratic method. In most of the other early dialogues Socrates is interrogating some bogus claimant to wisdom, as described in the *Apology*, so it is no surprise if they reach only negative results; but the *Apology* also makes it clear that we should examine our own lives and commitments, and the *Crito*, set for once on the ‘inside’, seems designed to show what that examined life is like. As for the second assumption, even when Socrates is arguing critically and generating refutations, many of his arguments do not actually fit the narrow template of the *elenchus* above. In the *Euthyphro*, approximately none of the refutations follow this pattern: instead, one definition is rejected as too narrow, another as giving an accidental rather than an essential property of the definiendum, another because it cannot be explained with sufficient clarity, and the last because it seems to lead back to the second. No doubt all of these *could* be reframed in elenctic form, with enough auxiliary premises added in; but neither Socrates nor Plato can be bothered to do so. Or, to take an almost random example from the *Charmides*, Charmides' second definition of temperance [*sôphrosunê*] is as modesty (160e-61b):

(P) Temperance is modesty.
(Q) Since temperance is a fine thing, temperate men are good.
(R) Temperance makes men good. (from Q, sort of)
(S) Only something good can make men good.
(T) Therefore temperance is a good thing. (R, S)
(U) In some situations, modesty is not good, so that modesty as such "both is and is not a good" and "is no more good than bad".
∴ (~P) Therefore, temperance is not modesty. (T, U,...?)

The proof is intuitive enough, but it is not properly speaking an *elenchus*: *P* plays no role in eliciting the conclusion, which comes directly from *T* and *U*. Or at any rate it is

---

47 Moreover, the critical side of Socratic dialectic is to the fore elsewhere because of Plato’s obsession with the hypothesis that virtue is knowledge, a hypothesis Socrates investigates with claimants to the various kinds of virtue and knowledge in question. This critical slant is not imposed by the method itself.
entailed by $T$ and $U$ if we also assume the intuitive principle that if $x$ and $y$ differ in their properties, $x$ and $y$ are not identical; and if we further stipulate that being 'no more good than bad' is different from and excludes being 'a good thing'. So there is more going on in the argument than meets the eye: but none of it makes any use of the fact that $P$ and $\neg P$ could be juxtaposed in the form of a contradiction. That contradiction is lost sight of, subsumed under the simpler finding that $P$ is wrong. Scholars sometimes broaden the relevance of *elenchus*-talk by including this kind of argument as a subspecies, but this seems to me to muddy the waters to no purpose. The *elenchus* as represented in the schema above *does* deserve to be distinguished as an important contemporary form of argument, shared with the sophists (as 'eristic'), and with a special utility for Socrates' investigations into expertise. I will try to reserve the term *elenchus* to refer strictly to this form of argument. For there is no way to broaden the sense of the term sufficiently to make 'Socratic *elenchus*' equivalent to 'Socratic argument' or 'Socratic philosophical method', and no good reason to try. Talk of Socratic method in terms of the *elenchus* can only ever lead to confusion, by conflating Socrates’ rich and varied dialectical practice with one particular form of negative argument.

So if the 'method of hypothesis' were simply 'the Socratic *elenchus*' under a change of name, it too would be a Procrustean and inaccurate a term for Socrates' method. But nothing in the term *hypothesis* implies an exclusive orientation to critical argument, let alone to contradiction-hunting in particular. And as I argued above, we should understand the method of hypothesis broadly, as the procedure of teasing out conceptual clarifications through deductive arguments of all shapes and sizes. So part of my hope here is that a change of terminology can provide an occasion to look at

---

49 It is perhaps also worth noting that the *elenchus* properly speaking seems to have been the discovery of Protagoras (). So another reason to prefer speaking of the dialectic of the early dialogues in terms of hypothesis is that it picks out the distinctive Platonic contribution to the method, which is *not* the search for contradictions as such but the assembling of more complicated strings of hypotheses and consequences, recognised as provisional and investigated to clarify their conceptual connections -- in other words, (1), (3), (4) and (5) in the analysis of hypothesis above rather than the essentially sophistic (2).
Socrates’ actual dialectical practice afresh, and help us do a better job of grasping its diversity.

This emphasis on the potential diversity of the method of hypothesis may threaten to water it down to no method in particular. To see what’s distinctive and exciting about the method, we need to ask ourselves what Plato is rejecting. And here, I think, a clear answer emerges from the contrast between his preferred methods and the alternatives of the day. The rivals of hypothesis are the argumentative methods of the Athenian courtroom, which held such sway over the fifth- and fourth-century thought that we find them reproduced in genres as diverse as history, medicine, sophistic and political argument, and Euripidean tragedy.\textsuperscript{50} Roughly speaking, and to oversimplify a bit, these arguments tend to take the form of speeches (often, in genres like history and tragedy, matched speeches pro and con, again showing the dominance of the courtroom paradigm), appealing either to testimony or to to eikos -- the probable, plausible or likely, particularly in relation to patterns of cause and effect.

Plato dramatizes this contrast in methods at the start of the Laches. Here Laches and Nicias give opposed speeches on whether young men should study fighting in armor. Nicias appeals to to eikos, giving a long list of beneficial consequences which one would reasonably expect to follow from the training. Laches’ speech refutes him with eyewitness testimony, telling a long unfunny story about a teacher of the training who made a fool of himself in actual combat -- which suffices to show that the likely benefits do not necessarily follow. When Socrates enters the fray, his strategy is very different from both. He proceeds by identifying a single point of principle which would be sufficient to settle the question: if this training makes men courageous, it should be practiced. What follows is a regress of hypotheses to determine whether the antecedent holds (a good case of the 'dialectical spiral' discussed in the Introduction). To know whether studying this subject does make men courageous, we need to know whether courage can be taught at all.\textsuperscript{51} And to know whether courage can be taught we

\textsuperscript{50} Note that the ‘higher hypothesis’ seems to be conceived as giving sufficient (not necessarily necessary) conditions for the lower one to hold.
need to know what courage is. Specifically, we need to know whether it consists of knowledge or something else, and so definitions of both kinds are examined.

The *Laches* thus displays the basic proposal involved in the hypothetical method -- and displays it as a correction to the methods of rhetoric, a way out of the impasse that speeches like those of Laches and Nicias are likely to lead us to. (Rhetorical arguments are ways of evading the 'spiral'.) When faced with a difficult practical question, rather than appealing to authoritative witnesses or piling up arguments about what is likely, we should zero in on a single claim which would be sufficient to settle the question deductively, and identify higher hypotheses which would suffice to establish or exclude it in turn. We are to work upwards to higher hypotheses and downwards to consequences in a systematic way, rather than assessing a proposed thesis in isolation; and we are to check for entailments and contradictions instead of merely gauging levels of intuitive probability. For a thesis which generates a contradiction is *impossible*; and as Sherlock Holmes said, “when you have eliminated the impossible, whatever remains, however improbable, must be the truth” (*Sign of Four*, ch. 6). 52

IV. The Hypothesis and its Consequences

Now my claim, again, is not just that the method of hypothesis is what Socrates practices in the early dialogues, but that the early dialogues are organized around one hypothesis in particular, namely the intellectualist hypothesis that virtue is knowledge. There is a lot to be said about just what that means and why Plato thinks it is so important -- so much, in fact, that I will have to treat the relevant conceptions of ‘knowledge’ and ‘virtue’ separately, in the next chapter, and later chapters will look at the consequences as they emerge in each of the dialogues. But in order to get a sense of the overall picture, it is worth at least listing some of the principal consequences of the

52 Cf. Holmes’ praise of ‘reasoning backwards’, his description of his own method as hypothetical, etc. -- the kinships between Holmes and Socrates are a large topic worth exploring. Direct influence certainly cannot be ruled out -- Watson’s comments on Holmes’ character at the end of *The Final Problem* seem to echo the end of the *Phaedo*. 
hypothesis. For the sake of convenience I will count as consequences the claims, worked through in different dialogues, that the particular virtues such as courage, temperance, and piety are kinds of knowledge: sometimes Plato seems to think of them that way, but they are also supported by independent arguments in each case (*). Of the 'across-the-board' consequences, regarding virtue in general, the most central is the one I will call *Teaching*: if virtue is knowledge, then virtue can be learned and taught. As the *Protagoras, Meno, Euthydemus* and *Laches* all strongly suggest, this crucial claim must have been the springboard for Plato's inquiry as a whole. It is because of the problem of education – and more specifically because sophists such as Protagoras and Euthydemus claim to be teachers of virtue, while others deny that claim (and even deny that virtue is the sort of thing to be taught at all) – that the question whether virtue is knowledge enters the agenda for philosophical investigation. The verdict of the dialogues on *Teaching* seems to be, as I have already suggested for the *Meno*, yes and no: it is true that virtue can be taught only if we take 'virtue' to be something quite different from what Pericles, for instance, had and failed to pass on.

Another major consequence is *Content*: if virtue consists in knowledge – that is, in an *epistêmê* or *technê* (cf. Ch. II) – then it must have some determinate subject matter distinct from that of the other crafts and sciences. (In fact Plato seems to assume that the subjects of real *technai* may not even overlap, cf. *Gorgias*, *Charmides*.) Plato addresses *Content* in the *Protagoras, Laches, Charmides* and *Euthydemus*, with the proposal that virtue consists in knowledge of good and evil -- or, less misleadingly, of good and bad things, i.e. value in general. This presents potential problems: for Socrates also seems to hold that virtue itself is the human good. And one might worry that the two claims together produce a vacuous regress or circularity, in which the good is defined as knowledge of the good. This worry, *Circularity*, is raised in the *Euthydemus* and *Republic* -- and from another angle in the *Charmides*, which worries about whether there can be a knowledge which knows only itself. The problem isn't explicitly presented as a contradiction, but it easily could be, and I think presents the kind of consequence which might well be taken to sink the hypothesis. But *Republic* VI resolves the worry, showing that *Circularity* can be sidestepped. Our good is indeed to know the Good, but that
means the Form of the Good: the two uses of ‘good’ are each substantive and conceptually distinct.

Content and Teaching are two of the four consequences which seem to particularly occupy Plato, recurring in more than one dialogue. Another is Unity: i.e., as I understand it, that the virtues are not only mutually entailing but identical, consisting in a single psychological power [dunamis]. This doesn't follow immediately from the hypothesis since in principle each virtue might be a different kind or area of knowledge. Rather, it follows from Content, on which all the virtues are knowledge of good and evil. The Laches brings out that Unity is in contradiction with the more natural assumption that virtues such as courage are each parts of virtue, distinct from the whole and from each other. That assumption is there introduced as an auxiliary premise, used to derive an elenchus which refutes the definition of courage as knowledge. If it were self-evident or inescapable, an indispensable part of our conception of virtue, then once again the hypothesis would be sunk. But the strong suggestion at the end of the Laches is that the assumption can be discarded: the virtues might indeed turn out to be all the same thing.

Another mediated consequence, following from both Content and Unity (and very likely Teaching as well), is what we might call Difficulty. The knowledge which is virtue -- a science of living constituted by a comprehensive knowledge of what is good and what is bad -- will be almost superhumanly difficult to obtain. Socrates himself disavows any such thing, and he certainly never encounters anyone who has it. It is more an ideal or educational programme than an achieved reality. This implies that the people ordinarily considered virtuous have only some inferior simulacrum, a distinct corollary I will refer to as Levels. The idea that there are different levels of virtue, some more the real thing than others, appears in various forms in the Meno, Protagoras, and later dialogues such as the Phaedo and Republic. It is one of the most load-bearing of all the consequences, for it entails that a lot of intuitive objections to the hypothesis, and to its other consequences, miss their mark. For instance, the unity of the virtues strikes interlocutors like Protagoras as obviously empirically false: but that is because they are

---

53 Penner vs Vlastos etc.
thinking of what passes for virtue by low everyday standards, not of the vanishingly rare virtue which is knowledge. It's no surprise, and no objection to the hypothesis, that the dispositions we commonly accept as virtuous are found separately – that’s part of what makes them inferior to the real thing. So Difficulty and Levels dovetail helpfully with the other consequences -- and not only with Unity but also Teaching. Real virtue must be taught if it is knowledge; but how the second-rate variety is acquired is another question.  

Content, Unity, Difficulty, Levels and Teaching are the major consequences of the intellectualist hypothesis, discussed repeatedly from different angles in a wide range of dialogues. And Plato’s investigations show that, far from contradicting each other, these consequences work together – they ‘harmonize’ in a positive way, interlocking to support and clarify each other and to disarm potential objections.

Three other consequences look more problematic, and will have to be discussed more fully later on. One is the denial of akrasia in the Protagoras -- the position the Magna Moralia took as a reductio of the whole intellectualist programme. A full treatment will have to wait for the discussion of the Protagoras; but I would like to prepare the ground by suggesting that we should once again be skeptical of the MM’s reading. First, if a denial of akrasia is a consequence of the intellectualist hypothesis at all, it seems not to be a particularly important one in Plato’s own view. The point comes up only once, in the Protagoras, and seems connected to none of the other consequences. Second, what the crucial Protagoras argument actually shows is far from obvious; and it is unlikely that its aim is to deny the possibility of akrasia across the board. For Socrates is supposed to be arguing in support of a thesis endorsed by himself and Protagoras both, namely that knowledge is sovereign. And a proof that akrasia is

---

54 In the Meno, real learning turns out to consist in recollection: so 'teaching' and 'learning' too come in higher and lower varieties. So 'virtue can be taught' does not mean that old-fashioned schooling will turn any student into a pillar of Hesiodic rectitude. Nor does it mean that Protagoras can make any one of us into a Pericles. For Plato, as the Republic eventually makes clear, it means that a successful education in Socratic dialectic could make the right person into a philosopher-king. (Not that this solution could be deduced from the early dialogues: but the end of the Meno points in its general direction.)
equally impossible in the case of the person who knows and in the case of the person who believes, does nothing to enhance the ‘sovereignty’ of the former.\footnote{Cf. Terry Penner} \footnote{Moreover, the thesis that knowledge is sovereign seems to be a version of the intellectualist hypothesis – a particularly ‘Protagorean’ version of it, tantamount to the claim that there is an art of successful deliberation and action which guarantees right action. That means that in context, Socrates’ argument about akrasia is presented as support for the intellectualist hypothesis, not as a consequence deduced from it. So the argument should perhaps be classed with Socrates’ argument at \textit{Meno} 77b–78b that, correctly understood, no one really wants what is bad; or his argument in the \textit{Gorgias} that wicked tyrants do not do what they want, based on his distinction between our actions and their desired ends. In each case, Socrates’ concern is to support the principle that all desire is for the good by showing how potential counterexamples (respectively, akratic action, action in apparent pursuit of bad things, and wicked actions) can be reclassified. These are perhaps ‘dialectical’ arguments in the Aristotelian sense, supporting a basic principle for which no demonstration can be given. Plato never does give a demonstration that desire must be aimed at the good, though he proclaims it often enough, so perhaps it is a first principle of this indemonstrable. How this mode of argument fits in with the hypothetical method is another question.}  

A second particularly controversial or paradoxical consequence follows both directly from the hypothesis and from the ‘higher hypothesis’ (as I will argue it to be) that desire is for the good. This is \textit{Involuntariness}: the claim that all wrongdoing is due to some kind of ignorance. (The \textit{Protagoras} argument about akrasia can be seen as supporting this claim for akratic wrongdoing in particular.) This too is not as prominent in Plato’s own texts as one would expect from his interpreters, but it is certainly controversial; and it not only strikes Socrates’ interlocutor Polus as wildly implausible, it is hard to square with Socrates’ own quite conventional insistence on rewards and punishments for good and bad deeds ().\footnote{But note that here tenet [5] of the \textit{Magna Moralia} reading breaks down: for Plato goes on insisting on the involuntariness of wrongdoing through his very latest writings, albeit with some variation in the reasoning he adduces and what he seems to mean by the thesis.} The third consequence we might expect to pose a particularly strong threat to the hypothesis is the one brought out by the argument of the \textit{Hippias Minor, Bivalence} or Bipolarity. If virtue, and justice in particular, really is a craft, then it must be open to being used for opposite effects. The doctor is the most effective poisoner, and only the person who knows the truth can tell
lies reliably. In contexts of craft, it is the person who errs voluntarily who is better than the one who does so involuntarily, for he is the one who knows what is doing. Socrates concludes, with some distaste:

"So the one who voluntarily errs and does what is shameful and unjust, Hippias -- that is, if there is such a person -- would be no other than the good man." (376b)

This clearly cannot be right, and unlike the other consequences which have roused interpretive alarm over the centuries, it is here flagged as unacceptable by Plato himself. The threat to the hypothesis is powerful and direct, since (as I will argue in Chapter II) it is predicated on just this conception of moral knowledge as a craft. Whether Plato thinks this problem has a solution, within the framework of intellectualism itself, will have to await our discussion of the Hippias Minor.

So much by way of sketching the consequences and their investigation. I hope it is at least sufficient to show that a reading of the dialogues as hypothetical doesn’t just boil down to the point that Plato only asserts ‘Virtue is knowledge’ in a provisional or tentative way. That’s part of it, but more important is the claim that the dialogues are organized as they are to uncover, clarify and scrutinize the implications of that hypothesis. And with that goes a claim about what kind of writing the early dialogues are. They aren’t expositions of doctrine, presenting the results of Platonic philosophizing. They aren’t teasing advertisements for doctrines to come in the Republic. Least of all are they commemorations of the historical Socrates, the man and his theory. They’re philosophical experiments: a series of dialectical tests and exercises in which Plato is doing his best to work through a complicated tangle of problems right before our eyes.