

*A re-reading of the Lysis: some preliminaries*I SOME METHODOLOGICAL PROLEGOMENA; AND A MAJOR
OBJECTION FROM PROPONENTS OF THE
'ANALYTICAL-ELENCTIC' APPROACH

In this second part of the book, we propose a second trip through the *Lysis* – this time with some philosophically more adventurous, and so more controversial, explanations (which will also be more pointed and more single focus) of the course of its argument as a whole. These explanations are more adventurous, first, by virtue of the extent to which they elaborate on *what the claims are* that we¹ regard as clear enough allusions, *in the text of the Lysis*, to Socratic claims² made explicitly only in other dialogues of

¹ Throughout this chapter and the two that follow, 'we' stands for Penner and Rowe, unless otherwise indicated. Rowe came later to some of the philosophical views that will be recommended, and indeed in a number of cases came to them only after long resistance. To that extent Rowe's title to the ownership of such views is less clear than is Penner's; but in no case where 'we' claim to believe something is Rowe anything less, now, than a willing participant in the claim.

² To be clear, when we appear to be distinguishing *Socratic* claims from *Platonic* claims, we *often are not doing so*. In fact we regard *fundamental* Socratic and Platonic positions as nearly identical save on one point – and on the implications of that one point (which may, however, be fairly extensive; see §4 below). Socrates – the historical Socrates as Penner thinks, the Socrates of a certain fairly well-marked part of Plato's stylometrically early dialogues as Rowe is inclined to think, though he is close to moving to Penner's view – is fundamentally at odds with Plato on the implications of only one question: a question about psychology of action. This is the question whether it is possible for any actions in that standard group of actions which Aristotle would later call 'voluntary actions' to be the direct result merely of irrational desires taken together with certain beliefs. In the Platonic 'parts of the soul' doctrine, actions of the sort Aristotle – though as we shall see (§4 below), not Plato – would call 'voluntary' or 'willing' can be brought about by brute or nearly-brute irrational desires, in accordance with the following sort of desire/belief explanation-schema:

I am thirsty (= I desire some water);
I believe there is water to be drunk from this glass here;
So
I take the glass and drink from it.

(See Aristotle, e.g. *Nicomachean Ethics* vii (= *Eudemian Ethics* vi).3; Plato, e.g. *Republic* iv, 435C–439D.) For Socrates, by contrast, no such explanation-schema is ever applicable to what Aristotle would call a voluntary action. On the contrary, for Socrates, the only desire that can ever function as

Plato – as explicitly, at any rate, as Plato's Socrates ever claims anything. Then, beyond that, in a second degree of adventurousness, we also make use of other Socratic claims – claims *not* alluded to in the *Lysis*, though required in order to see just how we are to take those claims that *are* alluded to.³ What this gives us is an entire web of interlocking claims about knowledge, desire, love and the good. All of these claims – we propose – are involved in the argument of the dialogue, and if we are fully to understand

the desire-half of a desire/belief explanation-schema is desire for a single ultimate good consisting in the agent's own maximum available happiness (which we take to be the maximum available good), given those circumstances of the agent's present life from which the agent now begins. (More detail on Socrates vs Plato on the psychology of action in §4 below.)

³ A word of clarification may be desirable concerning our talk of 'claims' Socrates and his interlocutors make. We engage in such talk because of our reluctance to go along with the usual way in which analytical philosophers identify what Socrates and his interlocutors are saying – in terms of the *propositions* they assert. Perhaps we might agree that a Socratic argument is a matter of Socrates' examining *what his interlocutors are saying* in using a given sentence. But most analytical philosophers engaging in analysis of Socratic arguments go on to identify what an *interlocutor* (including Socrates) is saying (in using a given sentence) with what the *sentence* in question says; and they identify what the *sentence* says (once due allowance has been made for indexicals and ambiguities) with the *proposition* expressed by the sentence (no matter who the interlocutor may be). The proposition expressed by the sentence is in turn identified with what Ryle 1945 calls the sentence's 'logical powers' – though the idea goes back to Frege 1879: 2–3. Briefly, but we hope adequately for our purposes, we may say that – according to the view in question – two sentences, e.g.

a. Piety is what is loved by the gods (*Euthyphro* 7A ff.)

and

b. Piety is what is loved by such beings as Zeus

differ in logical powers, because neither sentence can be inferred from the other by purely logical laws (or even – for those who believe in meanings – from logical laws plus a rule allowing substitution of synonym for synonym). A further premiss is needed if we are to get either from the other, namely 'The gods are such beings as Zeus.' Such a construal of what Euthyphro might be saying by means of such sentences as (a) and (b) would of course facilitate the examination of Socrates' arguments, since they could then be put into propositional logic and first-order quantificational logic and then deftly examined by contemporary standards (which is of course just what tends to happen when analytical philosophers take to examining Socratic arguments). We reject this account of what Socrates and his interlocutors are saying because of the excessively narrow (and ultimately falsifying) account of the identity of what someone is saying. By contrast, in the example we have just given, we maintain that what Euthyphro is saying by means of sentence (a) using the expression 'the gods', he could just as easily have said by means of sentence (b) using the expression 'such beings as Zeus'. For what an interlocutor claims, we think, is (in the simplest sorts of cases) a matter of what they 'have in mind' to refer to (what they intend to refer to), and, in addition, of what they intend to say about the object they have in mind or intend to refer to. 'What they intend to refer to' involves everything they believe about that object no matter how expressed – and perhaps rather more (if they think that the thing they have in mind to talk about must inevitably have properties they don't know about, and even properties that actually contradict things they believe about the object: but we leave this further matter aside till we come to – what we are calling – the 'principle of real reference', in §2 below). So if what Euthyphro *has in mind* to refer to, or *intends* to refer to, when he uses the words 'the gods' is such beings as Zeus, we have our point. Whatever other speakers might have in mind – and whatever the logical powers doctrine may tell us – we hold that when Euthyphro uses the words 'the gods' in (a), *he* has in mind to refer to, and *intends* to refer to, such beings as Zeus.

that argument, we need to take cognizance of all of them. It is true that the *Lysis* itself presents just one perspective on this complex web, placing some of its aspects (claims, beliefs) more to the fore, others more in the background. Other perspectives on this same web of belief will show up in other dialogues, depending on the differing focuses of attention in those dialogues. But, to repeat the essential point, as we see it *all* the Socratic claims in question – and all the aspects of his doctrines that are explicit only elsewhere, even aspects well in the background here⁴ – are involved in one way or other in the *Lysis*.⁵ Third, we believe that a correct judgement on what claims Socrates employs as (what we call) premisses of his arguments requires correct judgement as to whether the claims in question are true or false,⁶ as well as a (sufficiently) correct view of what those things and

⁴ To make this a little clearer, we take it that when an aspect of something *x* (courage, knowledge, the gods, Jocasta) is presented, Socrates' attention is not curtailed off from *x* in such a way as to exclude attention to unnoticed aspects of *x*, or even attention to *x* itself. (Contrast, in the example in the preceding note, the aspect of the gods which introduces them as 'the gods' and the aspect which introduces them as 'such beings as Zeus'. Thus while these two expressions have the same reference, they present that reference from two different aspects.) It will turn out that, for those who believe in them, propositions and meanings are certain sorts of standard aspects. (For such people, 'the gods' and 'such beings as Zeus', even if they refer to the same thing, have different meanings, and make for different propositions when combined with 'Piety is what is loved by . . .'. We of course insist, on the contrary, that Euthyphro's claim that piety is what is loved by the gods *is* his claim that piety is what is loved by such beings as Zeus.)

⁵ Our appeal to Socratic claims in other dialogues is in apparent violation of the maxim that one should try to understand a single Socratic dialogue entirely on its own terms and without aid from claims seemingly endorsed by Socrates in other dialogues (Gill 2002). No doubt that is a reasonable starting-point for interpreting a Socratic dialogue. But perhaps our view is not strictly inconsistent with the maxim in question. After all, the allusions that occur within the dialogue (at what we have called our first stage of adventurousness) *do* occur within the dialogue. Does the maxim say we must ignore the fact that an allusion is made *within* the dialogue? Does it say we must not ask whether Socrates (or Plato) *intends* us, *within* the dialogue, to consider the doctrine alluded to, in accordance with how he would want us to understand that doctrine further – for example, in the way the doctrine is understood in other Socratic passages? Does it say we are not to try to understand the allusion in a Socratic way – even if to understand what that is, we need to look to other dialogues? It is true that we do not endorse all of Gill's claims (for example, about the recollection passages in the *Meno* and the *Phaedo*). But we spend so much time here on Gill's maxim because we find ourselves in sympathy with his general animus against what he calls 'cross-dialogue' interpretations. There is a real temptation for analytical philosophers to atomize dialogues into 'arguments' or 'elenchi', then to atomize the arguments and 'elenchi' into *propositions* – these propositions then being taken to represent doctrines, those doctrines then being found in other arguments in other dialogues; in all of this ignoring how one bit in one dialogue needs to be understood in terms of another bit of the same dialogue. The crucial objection to cross-dialogue interpretations, we think, is that they abandon too soon the effort to understand one 'argument' or 'elenchus' in a dialogue in terms of what the dialogue as a whole is trying to do. We hope we have not done that. If we observe this caution, we see no defect to the present treatment of allusions to other dialogues that occur within the *Lysis*. In sum, we don't see that any of the moves we propose to make here are excluded by Gill's maxim.

⁶ This is not the case with propositions, as will be clearer from the 'analytical-elenctic' objection to our procedure which we consider immediately below. (Part of the idea of the proposition is that central

attributes are (if any) in the real world which the claims single out.⁷ (And, by the way, we accept the implication that all accounts of just what claims are being made, our own account included, must therefore remain tentative.) What is more, in a fourth degree of adventurousness which we recognize many will find frankly extraordinary, we shall argue that most if not all of these claims are actually *true* – and that they are broadly defensible. We argue this without in the slightest implying that we – or even Socrates or Plato – has any kind of full grasp on what the whole truth is about any such claims. It is one thing to believe something firmly, quite another to claim to know it.⁸

Now we are well aware that these exegetical methods put us on a collision course with the practice and impulses of those of our colleagues who approach the study of Plato by way of analytical philosophy.⁹ Perhaps their opposition to our approach may be crystallized in the following kind of

idea of logic from the time of Aristotle to the present that has it that sentences *say the same thing* whether they are true or false, and regardless of what the reference is of terms occurring in them.) Our contrasting position – on the need to know what the reference is of terms referred to by the sentence used if we are to know *what is being said* by means of that sentence – is indicated in n. 3 above. (Incidentally, it should be clear that these considerations concerning the identity of *what is said* will apply to many analytical philosophers who eschew the word 'proposition'. If it is taken that 'Piety is what is loved by such beings as Zeus' says something different from 'Piety is what is loved by the gods,' on the grounds that an extra premiss is needed to get one from the other, our objection remains whether or not the word 'proposition' is used. The issue is solely the issue of the identity criteria of *what someone is saying*, whatever word we use to bring out the relevant identity conditions.)

⁷ Suppose, in the simple case, that Euthyphro is using a sentence to attribute some property to some particular object picked out with the referring expression 'the gods'. Then we shall take it that if, on another occasion, Euthyphro is attributing the same property to some particular object picked out by the referring expression 'such beings as Zeus', then he is *saying the same thing* on both occasions. This simple case allows us to reduce the harder question of *what someone is saying by means of a given sentence* to the easier question *what someone is referring to by means of a given referring expression*. (It parallels the way in which Frege and his followers reduce the harder question of the reference of a *sentence* to the easier question of the reference of a *referring expression*.)

⁸ Nor do we think it will help to resort to talk of *partial* knowledge of this truth about any such claims, or – what comes to the same thing – knowledge of a part of this truth. (We do not see any other way of making sense of the idea of partial knowledge.) For without the whole truth, any restricted or retrenched claim bids fair to come into conflict with unknown parts of the whole truth that do *not* show up in the more restricted claim. (Of course, the positing of propositions is a standard way of attempting to generate entities of which one *can* have knowledge.)

⁹ We acknowledge that we ourselves came to the interpretation of Plato from the tradition of analytical philosophy, and indeed that we believe that a very great deal of the best work on the philosophy of Plato done in the past half-century came from analytical philosophers – beginning with Vlastos, but followed by several generations of extremely able workers in the same tradition, such as Santas, Irwin and Kraut (to name just three). This has come about in part because of an entirely laudable willingness (which we ourselves heartily endorse) to engage with Plato by making his arguments confront modern assumptions in terms of which it is natural to try to understand those arguments. (See Penner 1987: xiv–xvi, on why this attempt is a necessary feature of good interpretation.)

objection (which we shall call 'AN-ELENCH' as representing what we shall be calling the 'analytical–elenctic' approach to the analysis of argument):

AN-ELENCH: 'You two (Penner and Rowe) are surely going to end up playing fast and loose with the interpretation of the text if you start talking about things only explicit in passages outside the passage you are analysing. Why do you not restrict yourselves to what is *said* – and in the *explicit* premisses of the argument?¹⁰ Why do you not restrict yourselves – certainly in the first instance – to the propositions actually expressed in the argument? Why not formulate explicit premisses and conclusion, and then (making due exception for inductive arguments, e.g. arguments from analogy) assess the argument for validity and soundness before turning to other passages to shed light on the passage in question? At that point, it *might* be appropriate – if it should prove absolutely necessary to our comprehension of the passage – to turn to propositions explicit only outside the passage. But surely you should be starting from the text of the actual argument before you, and assuming that it is meant to be self-sufficient – at any rate till such time as we despair of seeing how it *could* be self-sufficient. Stick to the text!

'But we have another difficulty with what you say – this time going beyond mere questions of interpretation to more purely philosophical matters. Surely what you are proposing is going to take you very far from all logic and reason. Indeed, it is surely going to make logic impossible. Let us explain. You say that a correct judgement on what those sentences *say* which Socrates and his interlocutors employ as (what we call) premisses of his arguments requires correct judgement as to whether the sentences in question are true or false, as well as a (sufficiently) correct view of what the terms employed in that sentence refer to in the real world. Now surely this is a preposterous view of *what sentences say*. Surely we can know what a sentence says without knowing whether it is true or false? (Haven't you two heard of such a thing as *knowing the truth-conditions* of a sentence?) Indeed, all logical reasoning would be impossible, unless we can know at least what a sentence *says* without knowing whether it is true or false. For the whole idea of logic is that we can sometimes *come* to know whether

¹⁰ The premisses here will probably be sentences. They will not be claims or beliefs, unless the claims or beliefs are identified with what the sentences say. (We have been careful in nn. 3, 6–8 above to make clear that we make no such identification.) On the *explicit*, see the remarks preparatory to the account of the so-called 'Socratic Elenchus' in Robinson 1953: 1, 3, 5. Few who remember their first reading of Robinson will deny the exhilaration they felt at the thought of Robinson's breaking of exegetical windows, and the pouring in of fresh light on the issues. What we are suggesting, however, is that the time has come to repudiate Robinson's approach, much though, through Vlastos and those who follow Vlastos on the 'elenchus', it has captured the field.

a given proposition is true or false only by first examining its relations to other propositions. For example, before one knows whether it is true or false that p , one may surely engage in some reasoning, and come to the view that p , q and r together entail *not- p* – or what indeed amounts to the same thing, some other contradiction (or logical inconsistency) – and consequently come to the view that at least one of these three propositions must be false. If one is also confident of the truth of q and r , one may then conclude that p is false. (And this is of course just what we say happens in standard uses of the Socratic elenchus.)¹¹ But now notice that this entire – surely unexceptionable – argument determining that p is false proceeded on the basis that the logic of this argument is exactly the same (or would have been exactly the same) if p is true (or had been true). In classical logic, our sentence-letters stand for the same proposition whether the sentence is true or false. What an asserted proposition *says* is that the truth-conditions *do* obtain in the world; and to deny that a proposition is true is to hold that *those same truth-conditions* do *not* obtain in the world. Whether what a sentence says is *true* – that is a matter of how things are in the world. But what the sentence *says* – the proposition it expresses – must be the same whether the sentence is true or false, since otherwise, we could not employ logic in hopes of discovering whether it is true or false. (In an important sense, what a sentence refers to or is primarily about is its truth-conditions. That way a sentence can be about the same thing whether it is true or false.)¹² That is a foundation-stone of classical logic, and indeed of analytical philosophy.¹³

¹¹ See Vlastos 1994: 11 ff., who, however, uses 'standard' only for cases where the explicit contradiction involved is ' p and *not- p* '.

¹² Contrary to what we say in describing our third degree of adventurousness in the opening paragraph of this chapter. The analytical-elenctic approach to what a speaker is saying filters it, first, through the speaker's sentence (due allowance being made for indexicals and ambiguities), then through antecedently stipulated truth-conditions – those being just another (semantical) version of what we have referred to in n. 3 above as the sentence's 'logical powers'. In this approach, *logic* becomes a theory of logical *language*. Logic is no longer about things and attributes, conjunctions, alternation, existence and so forth, but about interpretations given by us (antecedently) to names, predicates, the conjunction sign, the alternation sign, the existential quantifier and so forth. (We believe this is a variant of what Bergmann, Rorty, Davidson and Dummett refer to as 'the linguistic turn'.) Of course a great many analytical philosophers will insist that what things and attributes, conjunction, alternation, existence are is *given by* our interpretations of these symbols. That is, on this approach, what things and attributes are *reduces to* what our interpretations make of the appropriate symbols (cf. nn. 3, 10 above). On the other hand, surely neither Socrates nor Plato would have gone along with the idea that things and attributes are given by how we use names and predicates (as Socrates says to Critias in the *Charmides*, 'As for me, I'll allow you, so far as concerns names, to assign (*tithesthai*) each in whatever way you like; just point out (*déloun*) what it is that you're applying whatever name it is to': *Charmides* 163D5–7). Compare the discussion in §2 below of the *real* nature of cutting; and compare the complaint in Brouwer 1928 that Hilbert was confusing mathematics with metamathematics. On this point, and probably this point alone, Socrates and Plato would have been in agreement with Brouwer.

¹³ This foundation stone of analytical philosophy – that you must know the (antecedently stipulated) truth-conditions of a sentence before you can even raise the question of its truth-value, that before

If you two don't accept that point, you might as well give up systematic philosophy!

This is a formidable objection. We wish we could reply to it in detail, beyond giving the kinds of hints we have given, or will give, in nn. 3–7, 9, 12–15 in the present chapter. Alas, our effort to do so began to take over far too much of the present treatment of the *Lysis*. We have decided that the best we can do in the circumstances is to try to make the difference between our approach and the usual one as clear as we can make it in brief compass. So for now we have just the following to say of the 'analytical–elenctic' objection. The objection makes the philosophical assumption that the best way to analyse an argument is to begin by doing it the honour of supposing that it is intended to be a logically valid, and sound, argument, consisting in propositions of the sort just characterized together with inferability relations (based on purely logical rules of inference alone) between such propositions. (Allowance is of course to be made if the argument involves inductive steps, e.g. arguments from analogy, as sometimes happens in Socratic argument. But on the 'analytical–elenctic' approach, these remain second-best arguments.)¹⁴ Ideally, then, one will attempt to formulate the argument in deductive form so as to assess it for validity and

you can even *look for* the answer to a question, you must know what the *question* is – is visible in perhaps its clearest form in the absolutely beautiful long opening paragraph of Frege 1918b, though it can also be found in Frege 1884: sec. 47 *ad finem*. Penner discusses this principle in detail (and rejects it) in 'Platonic justice and what we mean by "Justice"' = Unpub A, the main point of which article is briefly characterized in n. 23 below.

¹⁴ Hidden here is the idea – also elsewhere rejected by Penner: see his *Plato and the Philosophers of Language* (= Unpub B) – that there are two philosophically distinct ways of assessing the goodness of an argument, the first being the superior way, in terms of a deducibility relation that generates logical validity, the second being merely inductive or abductive or explanatory inference. (Indispensably connected with this dualism is the 'logical powers' doctrine of what sentences say.) Penner argues (*ibid.*) that this is an untenable dualism. Deduction, like proof, is not a way of attaining to knowledge. For there *are* no self-evident axioms or rules of inference, axioms and rules merely being *postulated* or *laid down*. And if not, then when one lays down axioms and rules, one does so, surely, on substantial (non-deductive, unproved) grounds that are as reasonable as one can find. If the issue, then, is how to argue for the truth of something, why is there a philosophical difference – a difference *in principle* – between

A. arguing for the truth of something by employing a deduction that appeals to 'axioms' or rules for which one has substantial (non-deductive, unproved) grounds that are as reasonable as one can find

and

B. arguing for the truth in question on substantial (non-deductive, unproved) abductive grounds that are as reasonable as one can find?

What could the grounds be for the deductive system being employed (or for the axioms postulated) but themselves substantial and explanatory? How then could there be a *philosophical* difference involved? Each attempt to argue for a position has to be assessed on its substantial (non-deductive, non-postulated) merits. (Here, some may concede that we do not come to know the theorems of an axiomatic system by means of proofs, but insist nevertheless, as Russell once did, that we do come to know this: that *if* the axioms and rules are correct, *then* the theorems are true. But this is an error.

soundness. Certain propositions are identified as premisses, and another proposition as the conclusion of that argument, which proposition is to follow from the others by means of purely logical rules of inference. The philosophical idea here is that the gold standard for interpreting certain sorts of philosophical arguments, especially those in Socratic dialectic, is rigorously deductive argument with explicit premisses, in accordance with the canons of first-order quantification-theory (and perhaps also the canons of some of its higher-order extensions) – the logic of *if . . . then . . . ; . . . or . . . ; . . . and . . . ; not . . . ; every; and some*, as elaborated in terms of now standard [metalinguistic] doctrines of validity, soundness, entailment, deductive inconsistency, semantic inconsistency¹⁵ and the like. And the idea

To suppose this would be to suppose that the axioms and rules of logic that justify the hypothetical claim *if the axioms then the theorems* are in better case than other axioms and rules. And there is no reason to think this is so – and many reasons for being rather more doubtful of these supposed axioms and rules than of many others, namely, the paradoxes.) We add here – what we have already suggested – that the absence of any absolute status for deductive truth also undercuts the ‘logical powers’ doctrine embodied in the theory of the proposition.

- ¹⁵ We take the (metalinguistic, or metalogical) distinction between deductive (proof theoretic, or syntactic) inconsistency and semantic inconsistency in the usual way. We also repudiate the idea that it is possible just to speak of validity in a more commonsense ‘baby logic’ way, as what obtains when, if the premisses are true, the conclusion *must* be true. For the modality of the word ‘must’ is the heart of the matter here, and needs to be explained; and we hazard the opinion that no one has made much progress without something like this metalinguistic machinery (bringing with it, we add, all the highly restrictive assumptions required, at almost every turn, to avoid paradox). Analytical philosophers should not delude themselves with the idea that they do not stand on the shoulders of Frege, Hilbert and Tarski.

To see how these metalinguistic and semantical notions too involve the ‘linguistic turn’ mentioned in n. 12 above, consider this. The premisses (some number of formulae, constructed from sentences by replacing all non-logical constants by symbols of appropriate sorts) of an argument lead to *deductive* (or proof-theoretic or syntactic) inconsistency if, via the axioms and rules of inference of logic, they together yield the conjunction of one formula with a second formula formed from the negation sign plus the first formula. The premisses of an argument lead to *semantic* inconsistency if there is no interpretation of the non-logical constants employed in the argument that would allow the premisses all to be true. Deductive inconsistency is a *linguistic* matter – a matter of what sentences of what forms (what formulae) can be deduced by the rules of inference (which are themselves rules about relations between kinds of formulae). The ideal is for such deductive consistency to be testable mechanically, without one’s even knowing how the formulae are to be interpreted – without knowing what the sentences from which the formulae were constructed affirm or are about. Semantic inconsistency, on the other hand, is (supposed to be) very much a matter of what formulae can be interpreted as being *about* – of what could be true and what could not be true, or, more precisely, of what the singular terms and predicates of the formulae can be interpreted as referring to or otherwise designating. It is sometimes thought that in semantics we are getting back to talking about real things and real attributes. But here one should not fail to notice that even when we are talking about those things in the world which singular terms and predicates refer to or otherwise designate – even when we are doing semantics – we are *never doing this except* via the linguistic entities which are the singular terms and predicates involved. We are still talking in the first instance about the language – the language with which we *hope* to describe these realities, and whose meanings or interpretations will determine these realities. On this semantical approach, nothing can be thought about except via what the language takes that thing to be (= what the meaning of the relevant

is that there is no substantial defect in supposing that Plato and Socrates would themselves have accepted the substantial accuracy of such analyses of Socratic dialectic.

These analytical–elenctic ideas are present for all to see in the doctrine of the so-called ‘Socratic elenchus’ which regiments each of the four or more supposed main arguments (‘elenchi’) of most of the stylometrically early dialogues of Plato into deductions of the following sort.

SOC-ELENCH: Socratic arguments typically consist, first, in Socrates’ asking of some primary question (sometimes, but not always, of the ‘What is *X*?’ form, but always a central question concerning human goodness or the human good). This primary question elicits from the interlocutor a ‘primary answer’, expressing the proposition that *p*, which answer Socrates then proceeds to attempt to refute. He does so by next asking several secondary questions – often of a trivial-looking nature, and indeed even of a quite irrelevant-looking nature¹⁶ – which then elicit certain ‘secondary’ answers, expressing the propositions that *q* and *r*. Socrates then proceeds to show that *p*, *q* and *r* together entail either a logical inconsistency or some other proposition the interlocutor finds repugnant. At this point Socrates – at

expression determines the reference to be). A metalinguistic approach to reality of such a kind represents, once more, ‘the linguistic turn’. But it is a matter of faith whether we can thus recover the realities. (The article of faith in question is equivalent to the claim that meaning determines reference.) This faith amongst proponents of the linguistic turn is not unlike the faith sense-datum theorists used to have that one could somehow recover reality by translating everything into talk of sense-data. We have already referred (n. 12 above) to the faith Hilbert had that he could translate talk of numbers and functions into talk of numerals and function-symbols (and indeed meaningless marks) and then recover newly validated numbers and functions – a faith Brouwer thought ended in the merest confusion of mathematics with metamathematics.

¹⁶ For example, few first-time readers of the *Republic* fail to feel such irrelevance when they contemplate Polemarchus’ primary answer that *justice is telling the truth and returning what one owes* (together with the secondary answer that *one owes good to friends and harm to enemies*) and Socrates’ subsequent question: who would be best able to help friends and harm enemies in matters of health and disease – the doctor or the just man? (331E–332D) ‘Why does Socrates think *ability* is relevant to questions of justice?’, the first-time reader asks. ‘And why should Polemarchus be forced to accept that it is?’ It is as if Socrates was urging on readers of his dialectic what Poe had to say about argument in the court system (‘The mystery of Marie Rogêt’ in Poe 1952: 382):

It is the malpractice of the courts to confine evidence and discussion to the bounds of apparent relevancy. Yet experience has shown, and a true philosophy will always show, that a vast, perhaps the larger, portion of truth arises from the seemingly irrelevant.

These remarks seem to us to support the approach we advocate in the opening paragraph of the present chapter, and to show the profound defects of analyses of what people are saying that employ the extraordinarily narrow conception of relevance built into the identity conditions of propositions. (See below, ch. 12, n. 11.)

least in many cases – appears to conclude that he has shown that the primary proposition *p* is false.¹⁷

To repeat: we unfortunately cannot undertake to discuss in any detail in this book the large issues involved here – whether in the analytical-elenctic approach to philosophical argument in general, or in the theory of the so-called ‘Socratic elenchus’. (We report with regret that an attempt by Penner to include such a discussion proved likely to take the book over and make it twice as long as it already is, as well as delaying its publication by many months.)¹⁸ So, again, just a few remarks.

We believe that the ‘analytical–elenctic’ approach to Platonic dialogues, as illustrated in *SOC-ELENCH*, falsifies them in numerous ways, of which we shall here mention two. First, it falsifies them in the way that it regiments what interlocutors are claiming in their arguments into propositions. In so doing, it gives us criteria of identity for what Socrates and his interlocutors are saying and arguing for that are *far too narrow*, so falsifying what they – Socrates and his interlocutors – are claiming. Thus to take the example introduced in n. 3 above, we say that when Euthyphro uses the sentence ‘Piety is what is loved by the gods’ he is saying the same thing as he would have been saying had he used the sentence ‘Piety is what is loved

¹⁷ Notice the words ‘proposition’, ‘entail’, ‘deduction’, ‘logical inconsistency’ in this characterization of the so-called ‘Socratic elenchus’, as well as the expressions ‘validity’ and ‘soundness’ in the objection (*AN-ELENCH*) above. These modern metalinguistic words are crucial to the idea of the ‘Socratic elenchus’, as it is to be found since Robinson 1953. Robinson 1953: 7, 15, 22 speaks of the primary premiss *p* together with the secondary premisses *q* and *r* *entailing* an inconsistency. So too Vlastos 1994: 11, 20, 21, 23, 25, Brickhouse and Smith 2000: 93, 83, cf. 79–80, Benson 2000: 33, 48, 62–4, 65, nn. 26, 95. For *validity* and *soundness*, cf. Robinson 1953: 15; Santas 1979: 136, 138, 166, 178–9; Vlastos 1994: 20, nn. 40, 41; Irwin 1995: 18, 20 with 40; also Benson 2000: 45–6, 49, 69 n. 47. The reference to *propositions* is of course ubiquitous – though, as we have remarked above (n. 6), what is crucial here is not the word ‘proposition’ but the ‘logical powers’ doctrine. We ourselves think that the idea of the ‘Socratic elenchus’ that is commonplace within the community of students of Greek philosophy is so inextricably entwined with these metalinguistic notions of deduction, validity, entailment, propositions and the like that we ourselves cede the expression ‘the Socratic elenchus’ to proponents of that deduction / entailment conception of Socrates’ methods of conversing with and cross-examining his interlocutors. When speaking in our own persons we shall restrict ourselves to such expressions as ‘converse with’, ‘cross-examine’, and ‘Socratic dialectic’. We should also mention here that we endorse *one* view in Davidson 1985 – the view that these latter expressions are quite as appropriate to Plato’s philosophical methods in middle and later dialogues as in the early dialogues – and without change of ‘sense’. Cf. also Chapter 4, n. 25, on Vlastos on the *Lysis* as not involving the so-called ‘elenchus’.

¹⁸ Some of this material will be found in Penner Unpub D. Penner is grateful to George Anagnostopoulos and Jerry Santas and members of their seminar at the University of California, Irvine and the University of California, San Diego, for giving him the chance to present some of the ideas here in Chapter 10 at two meetings of the seminar, and to discover just how much more work needed to be done to make clear the larger issues involved here. (For the same reason, Penner acknowledges a very large debt also to Antonio Chu.)

by such beings as Zeus,' while the propositional analysis makes it the case that Euthyphro would be saying something quite different by means of the second sentence, since (without the premiss 'the gods are such beings as Zeus') neither sentence is inferable from the other by purely logical means. It should be readily apparent that this makes the criteria of identity for what someone is saying by means of a given sentence very narrow indeed.

Second, the counting up of 'elenchi' in a dialogue – Robinson (1953: 24) counts 39 in the early dialogues, Benson (2000: 58–80) counts 5 in the *Euthyphro*, 4 in the *Laches*, 8 in the *Charmides* – tends to lead to the breaking up of the dialogue into fragments of elenchus length held together only by a literary stitching that is largely irrelevant from a philosophical point of view. For such analyses make it far easier for one to conclude that what is being discussed in one so-called 'elenchus' is not, and is not part of, what is being discussed in another. (Again, the point is about the narrowness of the criteria of identity for what is being discussed: an elenchus involving the sentence 'Piety is what is loved by the gods' would so far be logically independent of an elenchus involving the sentence 'Piety is what such beings as Zeus love.') On our reading Plato would take it that such so-called 'elenchi' are intimately inter-dependent. At any rate, we are convinced that this breaking up of Plato's text into elenchus-length fragments tends to deprive the dialogues of such unity as they may have in Socrates' mind or in the mind of Plato the author. We believe this happens all the time among modern readers of Plato, and that it is the saddest defect of all too many (otherwise extraordinarily able) interpretations of most of Plato's stylometrically early dialogues, regularly undermining the extraordinary unity that these beautiful dialogues often have – and that we strongly believe the *Lysis* has. All we can do here is to hope that we have been as clear as we can be in brief compass about our own assumptions, by contrast with those which we oppose, and hope that the reader will keep an open mind on the question whether the unity we have been able to find in the *Lysis* is not at least some evidence favouring our approach here.

2 'THE PRINCIPLE OF REAL REFERENCE'

We have one further clarification to make of the philosophical assumptions we attribute to Socrates and Plato. It has to do with what we have several times in Chapters 1–9 referred to as the 'principle of real reference'. It makes its first appearance in the *Lysis* in some otherwise largely inconsequential remarks Socrates makes to Hippothales in the introductory section of the

dialogue, where Socrates says of what Hippothales refers to as his praise of Lysis that it is actually praise of Hippothales himself. Hippothales replies:

‘But it’s not to myself, Socrates, that I’m composing or singing.’

‘You certainly don’t *think* so,’ I said.

‘But how’s that?’ he said.

‘It’s to you most of all,’ I said, ‘that these songs of yours refer (*eis se teinousin*) . . .’
(205D5–E1)

This shows that what Socrates is claiming here is that what Hippothales would be saying using the kind of sentence he *does* use, namely,

The songs I sing are in praise of Lysis,

is the same as he would be saying had he used the sentence

The songs I sing are in praise of myself.

To put the point in another way, the reference of ‘this praise of Lysis’ is ‘this praise of Hippothales’.¹⁹ How can this be?

The point here is based upon ‘the principle of real reference’. The idea is that while Hippothales doesn’t *think* he has in mind or intends to refer to praise of himself, nevertheless that *is* what he has in mind to refer to or intends to refer to. How so? Doesn’t Hippothales have some kind of first-person authority over what he has in mind or intends to refer to? We think not, and we think Socrates and Plato also thought not, since in our view Socrates and Plato would have rejected the (Protagorean) idea of first-person authority. Nor is our point merely that what Hippothales has in mind to refer to or intends to refer to *does in fact* – not in Hippothales’ own mind: rather entirely *outside* of the realm of what Hippothales has in mind or intends to refer to – designate something other than what he thinks. (For those who know the relevant literature, it will be apparent that we are discussing here, on the one hand, the traditional opaque, or better *oblique* readings of such psychological expressions as ‘intends to refer to’, and on the other hand *transparent* readings of such contexts – or in other terminology, *de dicto* versus *de re* readings of such contexts. These are the sorts of readings that in our 1994 we called inside/outside theories of psychological contexts.)²⁰

¹⁹ Cf. n. 7 above.

²⁰ See Penner (Unpub B), where it is argued that there *are* no so-called transparent senses, since they would require that there be a relation that is in some sense psychological between the subject of a psychological state and the supposed object of the psychological state – where, however, the subject need have no attitude whatever towards the object! If there are *opaque* (or, better, *oblique*) senses of psychological verbs, then the so-called transparent senses are best understood as existential

What our point is – given that it cannot be expressed in either of the traditional ways – will be best brought out by considering what Socrates says about the real nature of cutting at *Cratylus* 387A with 385D–396A. Here Socrates says that when I want to cut, I don't want to cut in accordance with what people *believe* about cutting, or indeed in accordance with our conventions for the use of the word 'cutting', but in accordance with the real nature of cutting. We take the implication to be that this *desire to refer to* the real nature of cutting rather than to what people's beliefs about cutting pick out from the world – or what our conventions for the use of the word 'cutting' pick out from the world – is a *desire to refer to* that real nature *even if that real nature is different from what we (or the conventions of our language) take it to be*. The result is that in standard cases we can be intending to refer to something and not know what that thing is. (This is what happens in cancer research applications. Some of these will be applications for funds to find out what cancer is in the distinct belief that the process of discovery will throw up in answer something not currently designated by any of our beliefs or any meanings in our dictionaries – something, indeed, whose nature may falsify both those beliefs and those meanings. So the cancer researcher doesn't know what the thing is that he is referring to, other than having a few ideas which, if he thinks them *more* on the right lines than the usual ideas, he can still be fairly confident are not in the end going to prove correct. There is a similar phenomenon to be observed in what happens when we refer to those we love. We have no desire to speak, or to be taken as speaking, of those determined by the totality of our beliefs about them, or by the conventions of our language for the words involved in expressions by means of which we refer to them. For we have no desire to think about those beings that our beliefs are exactly true of. Given human fallibility, such beings are not to be found in *this* world, and indeed are to be found at best in some *other* world – some dream world of ours. But it is the actual people in question, the ones in the real world, as they are, with all of their properties known and unknown that we want to think about, speak about and to be taken to be speaking about. One more example, which Penner has long used in his classes: Edith comes home, and Archie says to her, 'The Reverend Felcher called, Edith.' She says, 'The Reverend Fletcher, Archie.'

generalizations of the corresponding sentence with the psychological verb taken obliquely. (Cf. the overt existential generalization in the closely related sentence 'There is a description ["The Queen of Thebes who is *not* my mother"] under which Oedipus wants to marry this woman who is in fact his own mother'.) In addition, since we also deny that there *is* any first-person authority, we also deny there are any such things as oblique or *de dicto* readings of psychological statements. A whole new approach, we think, more in line with what we find in Socrates and Plato, is necessary:

He replies 'Whatever!' Archie's point, unarticulated here – and in this case, surely well taken – is: 'You know who I am talking about; I know who I am talking about; who gives a damn what his name is. If I've got something wrong, you fix it up!' That is, we are all clear who it is that Archie *has in mind*, or *intends to refer to*. It is the particular – at best incompletely known²¹ – referend in question, with all of his properties known and unknown, that Archie intends to single out.)

We believe this to be a familiar phenomenon in Socrates and Plato. Let us illustrate that briefly. First, there is a passage, *Gorgias* 466A–468E, which we shall be discussing in more detail in §4 below – in connection with the idea of all desire being desire for the real good. This says that what people are aiming at in their desire is – *from the inside* – the *real* (and at least partly unknown) good, even if that real good is different from what they believe it to be. They don't know what the real good *is*, but that is what they want – not merely what they *think* is good (the *apparent* good). 'If I have it wrong, you fix it up!' Second, this idea is also to be found at *Gorgias* 474B, where Socrates is explaining to Polus his view that while he is no good at persuading crowds, he can bring as witness for his position one person – the person he is conversing with at the moment:

SOC.: . . . For it is *my* opinion that I and you – and everyone else – hold that doing injustice is worse than suffering it, and also that not being punished is worse than being punished.

POL.: But *I* don't think that either I or anyone else would prefer suffering injustice to doing it – and would even you?

SOC.: Yes, and so would you and everyone else!

POL.: Far from it – either in my case or in anyone else's. (474B2–10)

As in the case of Hippothales, Socrates is telling Polus that he believes something Polus himself denies he believes. How does Socrates get to this? Once again, the issue is what Polus *has in mind* to refer to, or *intends* to refer to, when he uses the expression for doing injustice. Is what he wants to talk about what is determined by the totality of his *beliefs* about doing injustice, or about what our linguistic conventions for the word *adikein* determine in the world? Even if what those beliefs or conventions determine could only exist in some dream world of Polus' (assuming, given human fallibility, that a number of his beliefs and a number of these conventions are actually

²¹ 'But the person we are talking about can't be *completely* unknown to us!' We do not deny that if someone is too badly wrong in the reference they make, this may destroy the reference altogether, or that it may leave the speaker uncertain *what* to say he is referring to. The question is whether we need to be right in every respect – and whether we need to know in advance in what respects our view of the thing we are referring to is correct and in what respects our view is incorrect. Cf. also the note after next.

incorrect about what doing injustice is)? Or does he have in mind to refer to, or intend to refer to, doing injustice as it really is in *this* world? ('If I have it wrong in various ways, you fix it up!') But then – should Socrates be right, and should we 'fix up' Polus' error – this doing injustice will be a doing of injustice which is worse than suffering injustice. That being so, in thinking doing injustice is whatever it really is, Polus would be thinking – unbeknownst to himself, of course – that doing injustice is worse than suffering it. Thus does the principle of real reference generate Socrates' claim that Polus agrees with Socrates²² – even though, of course, there is no obstacle to Socrates' granting that Polus also *disagrees* with him.

This said, we should note that we have also seen the principle of real reference elsewhere in the *Lysis*. First, there is the passage where, in arguing that a person cannot love an object that is bad because what is bad will harm one, and no one wants what harms them, Socrates appears to require the assumption that the object that is bad is not even desired – even if one in ignorance believes that the object in question is good (Chapter 3, §(c); Chapter 4, §1, para. 1). We suggest that this only makes sense on the assumption that the object we want has to be something we want *as it really is* and even if it differs from how we think of it. Put in another way, since we do not want the harm this supposedly desired object will lead to, we do not desire the object, even if we (mistakenly) think we do. Once again, the idea is that what we desire (even from the inside – as *we* see what we desire) will be something incompletely known to us. The point here is closely related to the point just made about *Gorgias* 466A–468E – if it is not the very same point. Second, there is the passage about the poets 'riddling' us, in (as it were) pretending that they think that 'like loves like' covers cases of the bad loving the bad, while in fact their 'riddle' is to get us to see that what they really believe – given that the bad are not even even like themselves, let alone each other – is that 'like loves like' applies only in the case of likes that are *good* (Chapter 4 above, esp. nn. 28, 35).

To return briefly to the Hippothales passage: what we find there is that if Hippothales wants to be talking about his praise of Lysis as that praise

²² Our treatment of this passage may be contrasted with that of Vlastos (1994), which involves the postulating of an ambiguity to 'belief' (compare the oblique and transparent senses of 'believes that' discussed in the preceding note). According to Vlastos, there are *two* senses of belief: the ordinary 'overt' sense in which Polus believes just what he *thinks* he believes, namely, that doing injustice is *not* worse than suffering injustice, and what Vlastos (1994) admits is a 'marginal' ('covert') sense of 'belief' – we would have called it an *ad hoc* sense of 'belief' – in which one believes everything entailed by what one believes in the original sense. This marginal sense of 'belief' seems to us as plainly artificial as the supposed 'transparent' sense of belief. We think that underlying the mistake, here and in n. 20, is the suggestion that belief (and therefore intention to refer to something) must have first-person authority: you must know what you believe and you must know what it is you intend to refer to. It is this belief that is challenged by the principle of real reference.

is in *this* world, and not as it is in some other dream-world of his own, then the praise he intends to be talking about is the praise which a prudent Hippothales would see was a form of self-praise that would disadvantage him. Unless Hippothales is willing to represent himself as one who chooses *imprudently*, then he will grant that the praise of his darling *is* the praise of himself. If one takes Hippothales to desire to live in the real world (and not some dream-world), then one will take it that what Hippothales wants to refer to here is his songs of praise as they actually *are* in the real world – even if the way they are differs from how he thinks of them.

That is at any rate all we can say here on ‘the principle of real reference’ – the principle that the object a speaker has in mind to refer to, or intends to refer to, is the object as it is in the real world, even if it differs (as it almost always will) from the ways in which he thinks of it. We take this principle to be one of the most central features of Platonism.^{23, 24}

²³ Penner (Unpub A) also makes use of the principle of real reference in arguing, contra Sachs 1963 that the Socrates in the *Republic* (who says, in effect, ‘justice is a certain sort of psychological well-adjustment’) can disagree with Thrasymachus (who says ‘justice is the weak following the rules laid down in the interest of the stronger’) – in spite of the appearance that they are talking about two different things. For Thrasymachus’ intention is correctly represented not by

I am referring to the weak following the rules,

or even by

I am referring to the weak following the rules laid down in the interest of the stronger whether or not that that should turn out to have anything to do with the real nature of justice,

let alone by

I am referring to the weak following the rules laid down in the interest of the stronger even though I grant that this has nothing to do with the real nature of justice,

but rather by

I am referring to the weak following the rules laid down in the interest of the stronger, i.e. the real nature of justice even if how it is with it is different from how I think it is.

The result is – since both Socrates and Thrasymachus, when they use an expression of the sort ‘the real nature of justice’, will intend to refer to that real nature, even if, in various ways, they are mistaken about that real nature – that they can disagree with each other over the truth about justice; despite the fact that believers in meanings would say they each ‘mean something different’ by ‘justice’. Notice that none of this is to deny that when Thrasymachus discovers that the ‘i.e.’ above doesn’t work, he may come to the view that he no longer knows *what* to say (though, in fact, in the *Republic*, Thrasymachus does appear to give up and admit that he was wrong about the thing he intended to refer to all along). The discovery that one no longer knows what to say when the ‘i.e.’ one was employing doesn’t work represents a difference, but not a decisive difference, from the more usual case where perception that the ‘i.e.’ doesn’t work leads to a straightforward and pretty well automatic retrenchment. (‘The next president of the United States, Hubert Horatio Hornblower, . . . er, sorry, Humphrey.’ Here there isn’t any difficulty in seeing that it was Senator Humphrey to whom Jimmy Carter all along intended to refer. The mistake the ‘i.e.’ constructs is over-ridden without difficulty by who the person intended actually is.)

²⁴ The idea that the reference of ‘cutting’ is determined neither by our beliefs about cutting, nor by our linguistic conventions for the use of ‘cutting’, would of course lead Plato to deny the modern

We turn now to a statement of the principal conclusions to be argued for in the next chapter, along with a preliminary airing of a problem about self-interest.

3 PRINCIPAL CONCLUSIONS ABOUT THE *LYSIS* TO BE ARGUED FOR IN THE NEXT CHAPTER; AND A PROBLEM ABOUT SELF-INTEREST

Here we begin with a simple list of the conclusions at which we shall arrive in our re-reading of the *Lysis* in the next chapter.

1. In elaborating on our identification of the mysterious 'first friend' of the ostensible main conclusion of the dialogue with *wisdom* or *knowledge* (knowledge of the good, as it happens), we will show Socrates affirming the correctness of this identification of the 'first friend' with knowledge or wisdom *without having him reject* two common alternative identifications of the 'first friend' (as *happiness*, and as what Plato in the *Republic* would call '*the Form of the Good*').
11. We will put forward a proposal to show how *philia*, *erōs* and desire (desire for good) are so related to each other that, while each is distinct

assumption that meaning determines reference. We understand the meaning-reference distinction in the following absolutely standard – Fregean – way (Frege 1892): The *reference* of the expression 'The Morning Star' is the object in the world for which that expression stands. That object is, as it happens, the very same object, the planet Venus, that the quite different expression 'The Evening Star' stands for. Two expressions, one and the same reference. The *meaning* of the expression 'The Morning Star' is *not* the reference of that expression, however. To see why this is so, think of the meaning of the expression 'The Morning Star' as a set of instructions embodied in a dictionary, or simply in our knowledge of English, for taking us from the expression 'The Morning Star' to the thing in the world that it stands for. Take the meaning of the expression 'The Evening Star' as a similar set of instructions. Then the instructions assigned to the referring expression 'The Morning Star', namely,

look for the last bright heavenly body beside the moon in the morning

and the instructions assigned to the referring expression 'The Evening Star', namely,

look for the first bright heavenly body beside the moon in the evening

are plainly different sets of instructions. So, we say, the expressions 'The Morning Star' and 'The Evening Star' have different *meanings* – in spite of their having the same reference. Two expressions, two meanings, and just one reference.

Now for those who believe in meanings at all, meanings *determine* reference in the following way. Given the way the world is, the reference (if any) of a referring expression – some causal theorists make proper names an (unexplained) exception here – is the thing in the world which the relevant set of instructions takes one to. That is, the reference is the thing satisfying the conditions laid down in the meaning. (Even with indexicals – words like 'I', 'now', 'here', 'there' and so forth – meaning is still taken to determine reference given a specification of *context*.)

It is our view that once one abandons the doctrine that meaning determines reference, there is no remaining motive for any notion of meaning at all within philosophy. For more on our reasons for denying that meaning determines reference, see also the preceding note.

from the other, nevertheless *philia* and *erōs* each have as their underlying structure the desire for good. That is, we will be proposing that *philia* and *erōs* are particular forms of, or particular species of, the desire for good.

- III. We will suggest (a) that Socrates holds that anyone who has this desire for good has, as her or his ultimate desire, generative of all of his or her so-called 'voluntary' actions – actions being thought of as means to the ends desired in those actions – desires for *his or her own* good (that is, his or her own maximum possible good given his or her circumstances). We will also show that (b) the sort of desire for good just introduced generates, on the basis of Socratic passages in *other* dialogues, what we will call a teleological – and indeed *hierarchically* teleological – account of desire for good. (That teleological, hierarchical account of desire for good – prefigured in claim (a) by the representation of so-called 'voluntary' actions as means to further ends – is introduced in §4 below.) And we will show that (c) while this teleological, hierarchical conception of *desire for good* is hardly explicit in the *Lysis*, there is nevertheless excellent evidence for the presence of precisely such a teleological, hierarchical conception of *philia* (and indeed of *erōs*) in the dialogue. The presence of these latter conceptions – together with what we shall call the 'near-interchangeability' of *philia*, *erōs* and desire in the last part of the dialogue – will provide confirmation of the truth of each of the claims (a) and (b).

But there is another implication of claim (IIIIa) above that requires our attention for a moment. For this reference to all desire for good being desire for the agent's *own* good will create, for many, the apparent difficulty that the Socratic accounts of *philia* are, at base (so to speak), self-interested. To many modern interpreters of ancient philosophy (who think that what Socrates, Plato, Aristotle and the Stoics are all about is *morality*), this succeeds in making the Socratic account of *philia* either a contradiction in terms, or the next thing to it. For the main point about morality – the main attraction morality has for its proponents – is its intrinsic over-riding of self-interest in the cases (taken always to be possible) of conflict between morality and self-interest. This intrinsic connection to morality is evidently what is presupposed in conceptual analyses of love and/or friendship after the manner of the single most influential article on the *Lysis* over the past half-century or so. We refer to Vlastos 1969, which offers a particular conceptual analysis of love as a baseline for assessing the account of *philia* in the *Lysis*. Vlastos' suggested analysis, derived from some remarks in Aristotle (though

given a special twist by Vlastos, as he himself seems to admit)²⁵ may be expressed as follows:

VI.1. For x to love y is for x to desire the good of y purely for y 's sake and quite independently of any regard whatever for x 's own good.

Evidently, if one's love for someone is based upon, or has as part of its underlying structure, one's desire for one's own good, one would not be desiring the good of the other person independently of any regard for one's own good. So on this conception, love requires morality, or at least pure altruism – we might call it 'morally pure altruism'. The account of love based on self-interest which we find in the *Lysis* would not, then, after all, by Vlastos' analysis, be an account of love at all; it would be a complete non-starter.²⁶ That is indeed what Vlastos concludes. But, we think, he does so on the basis of largely unquestioned (though widely shared) philosophical

²⁵ Vlastos 1969: 3 gets the locution concerning desiring the good of one's friend 'for the friend's own sake and not for one's own' from Aristotle (*Rhetoric* III.4, 1380b35–1381a1), and uses that to support his own reading of 'for its own sake' in the *Nicomachean Ethics* in terms of total independence from one's own good (the modern 'intrinsic good', closely connected to the moral good). But Vlastos then admits (5–6) that for Aristotle true friendship requires that one get one's own good, as well as the good of one's friend, from the relationship (*Nicomachean Ethics* VIII.3, 1156b7ff. and VIII.4, 1157b1–1158a1, esp. 1157a33: 'and loving the friend they love what is good for themselves'). Instead of concluding that he has not correctly understood Aristotle's use of 'for the other person's own sake' in the *Nicomachean Ethics*, Vlastos concludes that Aristotle is confused, and mixes together something self-interested and something inconsistent with egoism, namely a 'for its own sake' that allows nothing of self-interest. (More on Aristotle in our Epilogue.) We admit that the *Rhetoric* statement remains. But there is hardly enough context to be sure just how seriously to take 'not for one's own sake' there. The point could have been as little as a point against *selfishness* – not caring for others at all – as opposed to a wiser self-interest. We discuss the issue of self-interest and desiring the good of one's friend in Chapters 11–12 below – but see also next note. And we discuss in Chapter 11, §8 the interpretation of 'for its own sake' in Socratic passages in Plato.

²⁶ Vlastos' initial understanding of 'for its own sake' in terms of 'intrinsic good' and even 'moral good' – along with the consequence for his account of love and friendship – emerges clearly enough at Vlastos 1969: 10, n. 24, where he says that

Aristotle's 'wishing another's good for his sake, not yours', though still far from the Kantian conception of treating persons as 'ends in themselves', is the closest any philosopher comes to it in antiquity.

Thus in Vlastos' account, loving someone is valuing that person *as a person*. We regard such a view of friendship as an unwarranted intrusion of morality into friendship and love. (Compare Kant's preposterous bit of *New Testament* exegesis at *Groundwork* sec. 1, para. 13, Ak. IV: 399, according to which the command to 'love your enemies' cannot be a command to *feel* love for your enemies, but can only be a command to act in certain ways – whatever you may feel – because it is morally right to do so. If Kant's view of what *feeling love* is leads him to this frankly incredible account of what Jesus had in mind, so much the worse, one may say, for his view of what it is to feel love for someone.) We grant, of course that it may well be easier for us to reject this implausibly morality laden Kantian and Vlastosian conception of love than it will be to reject a more common conception of love that also opposes any suggestion of the presence of self-interest in love – but without any necessary reference to morality. This is the idea that Rudebusch 2003: 131ff., esp. 131–2, has put as the idea that love requires of us a 'pure heart'.

assumptions – philosophical assumptions we ourselves reject, and believe Socrates and Plato would also reject. For Socrates and for Plato, and also for us (Penner and Rowe), love *is* in the end self-interested. Since we spend some time in the following two chapters justifying this claim, it will be convenient to proceed in the present chapter as if it can be made out. Then if Chapters 11 and 12 fail to convince, so will the present chapter. The present chapter restricts itself to other questions and difficulties.

But lest the position we here endorse seem a *total* non-starter against Vlastos' Kantian altruism – morally pure altruism – we offer some softening-up reflections on the following famous Kantian remarks on the topic of morality:

To be beneficent when we can is a duty; and besides this, there are many minds so sympathetically constituted that, without any other motive of vanity or self-interest, they find a pleasure in spreading joy around them and can take delight in the satisfaction of others so far as it is their own work. But I maintain that in such a case an action of this kind, however proper, however amiable it may be, has nevertheless no true moral worth, but is on a level with other inclinations, e.g. the inclination to honour, which, if it is happily directed to that which is in fact of public utility and accordant with duty and consequently honourable, deserves praise and encouragement, but not esteem. For the maxim lacks the moral import, namely, that such actions be done from duty, not from inclination. Put the case that the mind of that philanthropist were clouded by sorrow of his own, extinguishing all sympathy with the lot of others, and that, while he still has the power to benefit others in distress, he is not touched by their trouble because he is absorbed with his own; and now suppose that he tears himself out of this dead insensibility, and performs the action without any inclination to it, but simply from duty, then first has his action its genuine moral worth. Further still; if nature has put little sympathy in the heart of this or that man; if he, supposed to be an upright man, is by temperament cold and indifferent to the sufferings of others, perhaps because in respect of his own he is provided with the special gift of patience and fortitude and supposes, or even requires, that others should have the same – and such a man would certainly not be the meanest product of nature – but if nature had not specially framed him for a philanthropist, would he not still find in himself a source from whence to give himself a far higher worth than that of a good-natured temperament could be? Unquestionably. It is just in this that the moral worth of the character is brought out which is incomparably the highest of all, namely, that he is beneficent, not from inclination, but from duty.²⁷

We ourselves knew, and were lucky enough to have as a friend, a person who exemplified (in our view) exactly the characteristics Kant attributes to

²⁷ *Groundwork of the Metaphysics of Morals*: 398–9 (page numbers from the Akademie edition, (vol. iv)).

those ‘minds so sympathetically constituted that, without any other motive of vanity or self-interest, they find a pleasure in spreading joy around them and can take delight in the satisfaction of others’ – a person, alas, taken from us too soon, of whom his best friends will tell you that hardly a day passes that they do not feel the impulse to call him up to talk to him about some bit of philosophy or music or about some personal matter. We of course prefer to put Kant’s point about *having a sympathetic constitution* in the rather more Socratic terms of *having the wisdom, whatever one’s natural constitution, to understand the place of the happiness of those around one in one’s own happiness*. (On the other hand, we admit that we do not know, and probably never will know now, whether our friend would agree with us about what if any place morality plays in friendship.) It is at any rate clear than on Vlastos’ Kantian view, our friend would come in second to the person Kant describes as ‘not the meanest product of nature’. We, however, know whom we would rather have as a friend. And there is surely little doubt that any child knows which he or she would rather have as a parent.

It is true that Kant invites us to reflect on what would happen if people like our friend were to face a situation where their life was being clouded with sorrow – whether they wouldn’t turn egoistic in a nasty way, whereas the moralist will stay firm. But if one is to indulge in such attributions of possible motives, we too may wonder about the motives resulting from this disconnection – except in one’s moral commitments – from the happiness of others in Vlastos’ Kantian friend. The part played by morality in Western civilization is not without warnings of these sorts of aberrations. We recall the many horror stories of the upright Victorian father who, when his wishes are crossed by his children, seeks to impose his wishes on them from what he represents as purely moral motives. We think, in short, that one should not simply adopt Kantian, principled friendship in preference to the friendship that sees (what we take to be) the truth about the part played in our lives by the happiness of those around us. But we shall talk of this a bit more fully below.

Our next task, to be undertaken in §4 of the present chapter, will be to introduce and explain the psychology of action needed to ground the teleological, hierarchical conception of desire for one’s own good referred to in (IIIa) and (IIIb) above, as well as others of the claims in the same list. The psychology of action in question is that brand of psychological egoism²⁸ known among interpreters as ‘Socratic intellectualism’. As already

²⁸ Psychological egoism is the view that an agent’s every action *in fact* aims (ultimately) at the agent’s own good. Psychological egoism is usually contrasted with ethical egoism, the latter being taken

suggested above, this conception of the psychology of action is not explicit in the *Lysis* (though the teleological, hierarchical conception of *philia*, or love, for the first friend exhibits a parallel to it so remarkable as hardly to be coincidental). In any case, since our explanations will use this conception extensively in the account of the *Lysis* that follows in Chapter 11 – above all, in connection with the proper understanding of the teleological and hierarchical conception of *philia* – it will be desirable to have it before us. In the next section, therefore, we give a slightly extended account of that psychology of action. At first this section may seem disproportionate to our purpose of interpreting the *Lysis*. But if there is anything to our proposed explanations, i.e. (I)–(III) above, we believe it will have been well worth our while to have spent the time and effort required by the next section.

4 SOCRATIC INTELLECTUALISM INTRODUCED

Socratic intellectualism, as we construe it, is a psychology of action. It offers an explanation-scheme for every so-called ‘voluntary’ action whatever. It has at least two unusual features that need to be noticed right away. The first, which is quite familiar to readers of Plato, is the extremely intellectual-looking character of the explanations involved: every intended action conforms to the agent’s belief at the moment of action as to what is best for the agent in the agent’s circumstances. The second unusual feature is rather less familiar, though it lies in plain sight in the dialogues. This has to do with the question whether, if the agent does something he or she *thinks* is best for him or her in the circumstances, and the action turns out not to have been best for him or her, the agent wanted to do the action in question.

There are two different answers to be found in Plato’s dialogues. The first answer – that the agent does want to do the action that will in fact

as a doctrine not about how people *do* act, but about how people *ought* to act – a *normative* or *evaluative* or *moral* doctrine. That is, ethical egoism asserts that (by some normative, evaluative or moral system), one *ought* to seek one’s own good, or that some norms, values or morals make it the case that it is (morally) *good* that one seek one’s own good. As an example that will be familiar to some, consider the views of Ayn Rand: since it is her view that looking to your own interest (and indeed – a further matter – looking to your own interests *as against the interests of others*) is a moral imperative, she is a clear case of an ethical egoist.

Now Socrates believes that it *is* good that one seek one’s own good. But not, we believe, on the basis of any normative or evaluative or moral system. Rather, we suppose that as Socrates thought that (1) it is a matter of scientific fact that good knives are sharp and also a matter of scientific fact that cutting is the good end to which the sharpness of knives are means, so too he thought that (2) it is a matter of scientific fact that a good person has the knowledge that is virtue (excellence) and also a matter of scientific fact that happiness is the good end to which the virtuous (excellent) person’s knowledge that is virtue (excellence) is the best means. It is thus not a matter of norms or values or morals. Just a matter of fact. So, whatever one of us (Penner) has said in the past, the expression ‘ethical egoist’ should not be used for Socrates. We prefer simply to speak of a purely factual ‘Socratic ethics’.

turn out badly – is the only one most interpreters ever succeed in finding. This answer can be found wherever Socrates is working from beliefs of his interlocutors, and perhaps also, on some occasions, when he is simply, as Berkeley would put it, ‘speaking with the vulgar’. But it is arguable that such ‘speaking with the vulgar’ is just another case of working from beliefs of actual or potential interlocutors. In that case, this first answer to the question – that

VULG. the agent does want to do the action he or she does – even when it will turn out not to maximize the agent’s available happiness or good –

is never an answer Socrates will give in his own person. (According to this answer – which is *not* Socrates’ answer – when Helen ran away with Paris but that action turned out worse for her than other available alternatives, it remains true that, at the time, she *did* still want to do the action.) The second, and very different, answer is the one that Socrates, and also Plato, will give when speaking in terms of their *own* deepest convictions about human desire and human action. This is that

SOC. the agent does not want to do the action he or she is doing – the one that will turn out not to maximize the agent’s available happiness or good.

(According to this second answer – which *is* Socrates’ own answer – Helen only *thought* she wanted to run away with Paris. In fact, and unbeknownst to her at the time, she did not want to run away with Paris.) This second answer – rather puzzling, at any rate the first time one opts to take it seriously – will be deployed below, along with some textual evidence for it.

Socratic intellectualism²⁹ starts from the claim that every action whatever of the sort Aristotle used to call ‘voluntary’ results from a certain generalized desire, i.e.

²⁹ For the formulation of Socratic intellectualism that follows, see Penner’s O’Neil Memorial Lectures (= Unpub c), from which much of this material is taken, as well as a formulation now available in Gill 2005. The formulation derives in the first instance from *Gorgias* 466A–468E, a passage which tells us clearly enough why someone does an action if (Case 1) the action turns out to be [maximally] beneficial [in the circumstances]. For in that case, we are told, the agent is doing what he or she *wanted* to do. But, as will soon become apparent, if (Case 2) the action turns out to be harmful – less than maximally beneficial – Socrates in this passage (and a few others, to be discussed later) will say, paradoxically, that the agent *did not do what he or she wanted to do*, but rather simply did what *seemed best*. That raises the following problem (put to Penner most insistently, and most helpfully, by Antonio Chu): but if the agent didn’t want to do the action which unfortunately turned out badly, why on earth did he or she do it? Socrates never addresses this question. Yet it is a question that needs answering. How can a belief-desire account of the explanation of voluntary action account for the agent’s doing actions he or she *didn’t want to do*? Penner has provided such an account, first in Penner and Rowe 1994. That account will turn out to be embodied in the present treatment of Socratic intellectualism: see the discussion of ‘Case 2’ below.

DES. the desire for *whatever action may be* the best means currently available to me, in the circumstances I am in, to the end of maximizing the amount of happiness (or of ultimate good) that I will achieve over a complete life,³⁰

together with

BEL. the belief that *this action here and now* realizes the best means in question, and thus instantiates the general characterization given in (DES.).

(More about how this belief concerning a particular action results from other beliefs and the dialectical deliberative process later in this section.) Now, following up on our question just above, we need to divide our development of this basic position by cases – into the case where the belief in question, about which action is best, is *true*, and the case where the belief in question is *false*. This will be necessary if we are to capture Socrates' view of what agents want to do – that is, when he is not speaking with the vulgar, but in accordance with his deepest convictions.

Case I is where the belief (BEL.) is *true*. Then when the action identified in (BEL.) is substituted into (DES.), (DES.) is transformed, by a sort of identity through change, into what we propose to call an 'executive desire',

EXDES. the desire to do this action here and now which is the really best means to the agent's maximal happiness (maximal good),

³⁰ The analysis offered here begins from

(i) the desire to do *now* whatever action may be the really best means available to one's happiness or ultimate good.

This is a simplification from a fuller account which would generate that desire further back – from

(ii) the generalized desire for one's own greatest happiness or ultimate good over a complete life, together with the fact (which we take to be something like a law of nature) that

(iii) humans, unlike oysters, are beings whose reaction to their desires is to *act* in such a way as to fulfil them,

so that

(iv) with beings who desire their own ultimate good, they will, at any one instant, necessarily desire whatever particular action is the best action currently available to them in their circumstances as the means to that ultimate good.

But this is just (i) above, concerning desire for whatever action may be the best means. Thus, our simplification consists in beginning immediately with the latter desire, i.e. (i), for the present simply taking the part played by (ii), the generalized desire for happiness, for granted. The reason for calling this a 'generalized' desire is indicated by the phrasing '*whatever action may be* the best means . . .'. The idea is that that phrasing gives us no particular action. The agent wants *whatever* particular action is best at this particular instant prior to his or her wanting *any* particular action. For the agent hasn't as yet arrived at any belief as to the identity of the particular action that *is* the really best means available. This desire remains generalized until such time as the agent arrives at such a particular belief. At that point, as a result of the substitution for 'whatever action' of the particular action identified by the belief, the desire transforms itself into what we call below an 'executive desire'.

so that the action take place immediately.³¹ (An executive desire is a desire explanatory of a voluntary action which has actually taken place.) In this case, where the belief is true and so the action is the really best means, Socrates will say that in so acting the agent 'did what he or she wanted to do'.

Case 2, by contrast, is where the belief identified in (BEL.) is false. Then Socrates will say one of two things, depending upon whether (i) he is employing an interlocutor's premiss, and so 'speaking with the vulgar', as in (VULG.) above, or whether (ii) he decides to speak in accordance with his deepest beliefs about human motivation, as in (SOC.) above. When he speaks with the vulgar – when arguing from views expressed or held by others – Socrates will say something that to moderns looks very close to what he says in Case 1 – though in fact it is *not* the same thing, since it will involve having the agent go for the *apparently* best means rather than for the *really* best means. (For if, as in Case 2, the belief as to what the really best means is, is false, then the action identified by the belief as the best means *will not be* the really best means, but will only be *believed* to be: that is, it will only be the *apparently* best means.) Given this retreat from the real good to the apparent good – a retreat characteristic of virtually all post-Platonic philosophy, from Aristotle and Aquinas through to Anscombe and Davidson – substitution of belief (BEL.), even when (BEL.) is false, into generalized desire (DES.), is still taken – by these post-Platonic philosophers – to transform (DES.) into an executive desire which is

* EXDES./APP.³² the desire to do this action here and now which is the
apparently best means to the agent's maximal happiness
 (maximal good);³³

a desire which immediately brings about the action that the agent at that point apparently wanted to do – though, as already pointed out, because (BEL.) is false, the action is not in fact the really best action which the agent was said to want in (DES.).

³¹ We borrow the word 'immediately' here from the Aristotelian psychology of action (*Nicomachean Ethics* VII.3, 1147a28). Aristotle's 'immediately' (*euthus*), however, is qualified by a 'provided that nothing interferes' (1147a30–1, cf. 33–4). Aristotle is thinking here of a deliberate action hijacked by an irrational appetite that drags the agent into an akratic (irrational, weak-willed) action. In Socratic intellectualism, there is absolutely no provision for the kind of hijacking of the action by irrational desire of the sort that Aristotle is envisaging here. See also n. 2 above.

³² An asterisk before the name or number of an item or claim, here and in what follows, indicates an item, claim or belief that we (and, as we claim, Socrates and Plato) reject.

³³ It would be somewhat unsystematic on the part of Aristotelians, Thomists and various moderns not to take the end also as merely the agent's *apparent* maximal happiness. But the issue need not be raised here.

It is worth pausing here to recognize that with this retreat, for the case where the belief (BEL.) is false, from the really best means to the apparently best means, we are turning away from an actual means to a certain conception or description – a false appearance, in fact – of what the really best means is. The particular action that then results is the one that falls under this description or conception or appearance – *with no direct reference at all, included in that description or conception or appearance, to what is in fact the really best means*. According to the view that this is an appropriate retreat to make, it is enough, in order for us to be able to say that the agent *wanted* to do the action he or she did, that the agent *believed* the action best. Indeed, on the usual modern view – which, from the Socratic perspective, will be to ‘speak with the vulgar’ – the action the agent *thought* he or she wanted to do precisely *is* the action the agent wanted to do.³⁴

The reflection that if we are to speak with the vulgar, then the agent may no longer be said to desire the really best means, leads us to how Socrates speaks (in reaction to any such view) when he speaks in accordance with his deepest beliefs. What lies behind this Socratic mode of speaking is a very deep realism about the objects of psychological states. This realism, which lives on in Plato (for example, in such ideas as that of true and false pleasures which we find in both the *Republic* and the *Philebus*), imposes the requirement that what we and others are thinking about is *the real things that are there*: not things as they *appear* to us, *à la* Protagoras, but things as they *really are*, even if how they are is different from what we suppose them to be (or indeed different from how our language conceptualizes them). Thus Socrates denies that the agent ever stops desiring the really best means. In other words, he would repudiate any retreat from the really best means to the apparently best means. Instead he insists that the agent, in acting on the false belief, has *not* done the action he or she wanted to do, though he will say that the agent did what *seemed best* to him or her (the distinction so clearly laid out in *Gorgias* 466A–468E: see n. 29 above). It is only the action the agent wanted to do if the executive desire is the desire (EXDES.) to do the action which is the really best means to the agent’s maximal good. It follows that in Case 2, the executive desire which brought about the action is not (*EXDES./APP.), since that would involve Socrates’ retreating to the apparently best means. But neither is it (EXDES.) itself, since the action done

³⁴ This view – widely endorsed in one form or other ever since Aristotle, we have noted – is strongly reinforced by Protagoreanism in antiquity, and in modern philosophy by Cartesianism about our own inner states. (‘Under the description *D*’, since it requires the assent of the agent to the description, is – as a matter of fact – an invitation to Cartesianism: for example, to the doctrine that, at least in the simplest cases, I cannot fail to know what I believe, what I feel or what I desire.)

is not the really best means. What executive desire, then, could bring about the action that was in fact done?

This raises a crucial issue:³⁵ how the action ever took place at all, if the agent *did not want to do* the action he or she did. Neither Socrates nor Plato ever tells us how this question is to be answered – perhaps because the case of most immediate importance to them is the one involved in virtuous action, namely Case 1, the case where the agent's beliefs are true. But if we do not make the Aristotelian–Thomist–modern retreat from the real good to the apparent good, what are we to do?

In the absence of any answer in the Platonic text, Penner has constructed an account³⁶ which enables us to stick with desire for good as desire for the real good, while still allowing for the operation of an executive desire in producing the action – notwithstanding the fact that this new executive desire will not be a desire for the actual action done. There will be *another*, defective, sort of desire – which Plato might have called a 'false desire' – that will bring about the action which the agent did (though by Socratic convictions, he or she did not want to do it).

We get this other desire as follows: substitution into the desire (DES.) of the false belief (BEL.) that this action here and now is the really best means to fulfilling (DES.) should result in the following executive desire:

*EXDES./INC.³⁷ the desire to do this action here and now which is *both* the really best means to the agent's maximal happiness (maximal good) *and* the actual action done which the agent thinks to be the best means available (*though in fact it is not*).

It may be replied that there *is* no such action. But we grant that. It is just that the agent doesn't know there is no such action, so that he supposes there is. It is that false belief which creates, and is embedded in, the desire (EXDES./INC.). There is indeed an incoherence in this desire, as there is in the belief. The question is not whether there can be any such action, but only whether there can be a (defective sort of) desire to do such an action. We answer: yes.³⁸ The fact that there is no Santa Claus does not stop it being the case that some misguided child is waiting for Santa Claus. In just the same way the incoherent, defective executive desire (EXDES./INC.) brings about an action – this action here and now which the agent *thinks* to

³⁵ Antonio Chu's issue: see n. 29 above.

³⁶ First in print in Penner and Rowe 1994, 8–9, with n. 2 (pp. 1–2).

³⁷ 'INC.' for 'incoherent': see below.

³⁸ As is noted elsewhere, e.g. in Penner 2002, 208–9, n. 19, there is nothing particularly troubling for a philosopher of language in the idea of such an incoherence in an erring psychological state.

be the best means, even though it is not. This action, we submit, is not the action the agent wanted to do – whatever Aristotle, Aquinas, Anscombe or Davidson say. The agent wanted to do the action which was both this particular action *and also the action which constituted the really best means*. And there is no such action. *A fortiori*, the agent has not done the action he or she wanted to do.

Such is the view that we (Penner and Rowe) take of cases where the action the agent apparently wanted turned out not to be the best action available. On this view, to sum up (and repeat), Socrates will say, when speaking in accordance with his deepest convictions, that
 soc. the agent did not want to do the action he or she did – the one that turned out not to maximize the agent's available happiness or good. What is the evidence that this is how Socrates speaks when he is doing so in accordance with his deepest convictions? The evidence is not restricted to the passage in the *Gorgias* to which we have been appealing so far (466A–468E)³⁹ – and the message of which is surely irreproachable once one loses the idea that Plato is hopelessly confused (and/or the idea that the passage involves a special sense of 'want' – because what Plato says is inconsistent with what modern philosophers say). Penner and Rowe 1994 argues that the only way to understand *Meno* 77A–78B is to take it, too, as being in accordance with that view. That is, we propose, 'everyone desires the good' must be understood in the *Meno* too in terms of everyone desiring the *real* good, not the apparent good – though this has been doubted.⁴⁰ The same idea can be found in the *Republic*, at IX, 577D10–11: 'the tyrannical city least of all does what it wants to do', says Socrates there, a clear enough allusion to the *Gorgias* discussion (or the idea it advances). But most importantly, the idea in question is contained in the famous dictum 'No one errs willingly,' when this is filled out – as it must be – as 'No one errs willingly <sc. at getting what is best for oneself> .'. No one errs willingly at *this*: such is the claim, notwithstanding the usual view (e.g. Aristotle's), that people frequently, and quite willingly or voluntarily, do actions that turn out not to be best for them – either by virtue of the action's flowing from a (mistaken) rational desire (wish), or by its flowing from an appetitive desire – so that actions of both sorts are (on this Aristotelian, Thomist, modern view) willing (voluntary). Contrary to this more usual view, what we find not only in Plato's early dialogues, but also in late dialogues like the *Laws*,⁴¹ is

³⁹ See n. 29, and §2, above.

⁴⁰ E.g. by Mariana Anagnostopoulos in Reshotko 2003: 171–91; cf. also Anagnostopoulos' doctoral dissertation at the University of California, Irvine, 2001.

⁴¹ This appearance of 'No one errs willingly' in the *Laws* used to puzzle one of us (Penner) since he thought that 'No one errs willingly' should be quintessentially Socratic, and not Platonic. Here he

still that no one errs willingly. What this tells us is that where an agent does action *A* and *A* is not the best action (so that one *erred* in one's action) then one did *not* do *A* willingly. So the only action done willingly is an action that is in fact beneficial. This is Plato's steady doctrine – when he is not 'speaking with the vulgar'⁴² – whether in Socratic or Platonic parts of the corpus.

So there is ample evidence outside *Gorgias* 466A–468E for the view that according to Socrates, and Plato (where they are not speaking with the vulgar, but rather in their own person),

NACT. No one ever wanted to do an action he or she did which did not maximize the available real happiness or good.⁴³

was too influenced by the assumption that Plato thought of voluntary action in the way Aristotle did, so that Plato, like Aristotle, should allow that one does often err willingly (voluntarily) when an action flows from irrational desire after the manner described in n. 2 above. But if Plato continues to have even the Athenian Stranger affirm (at *Laws* IX, 860C–861D) that no one errs willingly, then he too must be supposing that in any case where any particular action *A* does not in fact maximize the agent's available real happiness or good, the agent does not willingly (voluntarily) do *A*. (If there is any other available basis for the claim, we – Penner and Rowe – do not see what it is.)

This is an example of how careful we need to be in putting Plato's thought in modern terms. It is a complete misrepresentation of Plato to say that Plato, in agreement with Aristotle (and contrary to the view of Socrates, where there are no irrational voluntary acts), holds that actions proceeding from irrational desires are also voluntary. What is true here is that Plato agrees with Aristotle that the *goodness or badness of individuals may be judged from actions that proceed from irrational desires* (where Socrates will disagree, since he does not think there are any such actions). What is *not* true is that Plato holds that any such actions are voluntary (willing). As we have just said, Plato does not even hold that actions issuing from the rational part of the soul, but based upon false beliefs, are voluntary. This is all a consequence, not of anything that is different between Socrates and Plato – so that we would need to wonder how the Socratic 'No one errs willingly' can still appear in the *Laws* – but rather of the view that if an act turns out not to maximize the agent's available real happiness or good, then the agent did not want to do it. This is the view we have been identifying in the main text as the view that reflects Socrates' deepest convictions about human motivation. Those particular convictions remain in the thought of the mature Plato, in spite of his acceptance of a parts-of-the-soul doctrine. (This, incidentally, resolves the crux at *Republic* VI, 505D10–506A4, where it is said that we all, always, pursue the good – as if Plato thinks that no one ever pursues what he or she is led to by irrational desires! 'Pursue' here – as we ourselves for the most part failed to see up till now – is simply to be understood as 'willingly pursue' – an understanding that is surely natural enough in itself. See Rowe forthcoming, a preliminary essay on the *Republic* passage in question.)

⁴² So e.g. in the *Lysis*, when Socrates simply allows that Lysis wants to do things that his parents see will be harmful to him (207D–210D), he is speaking with the vulgar – drawing out consequences of how *the boy Lysis* sees desire. (As elsewhere in the present book, we – Penner and Rowe – feel not the slightest embarrassment about supposing that Plato and the Socrates of the *Lysis* want to say the same things. That is, we hold, perhaps in current terms unfashionably, that the Socrates of the *Lysis* is Plato's *portavoce*; and even a splendid book like Blondell 2002 has not inclined us towards any shift of position on the issue.)

⁴³ Another clear illustration of this view of Plato's may be seen in the discussion of the willing and unwilling discarding of beliefs at *Republic* III, 412E10–413C3. The only case of willingly discarding a belief is the case of discarding a false belief. Any case of discarding a true belief for a false one is taken to be *unwilling* (412E11–413A1). The discarding of a true belief for a false one occurs in three ways – by theft (by being persuaded otherwise, or by forgetting), by bewitchment (by pleasure) and by compulsion (by fear). The key point in the passage is that no one willingly believes something false. But then neither can anyone willingly do an action flowing from a false belief.

This remarkable approach to what we *want* to do, as we have already made clear enough, has parallels in Socratic/Platonic thought concerning what people believe and what people refer to.⁴⁴

To sum up on our proposal as to how Socrates *would have* handled Case 2 when he is speaking in accordance with his deepest beliefs: we admit that there is nothing like (EXDES./INC.) in the text of Plato. But at the same time we are not aware of any other workable proposals as to how it can be the case that all three of the following are true together:

- A. no one errs willingly in (what Aristotle will call) voluntary action, and that
- B. no agent ever wanted to do an action he or she did which ended up being harmful to him or her, and yet
- C. the agent in some way *opted to do* the action.

So in advance of the production of some other account of how we are to deal with such claims, we shall assume that the present account embodied in (EXDES./INC.) is along the right lines.

A few remarks now about the belief (BEL.) that this action here and now realizes the best means, and so is the really best action. This belief does not just come out of the blue, but is the product of a deliberation that involves – either as explicit premisses or as background assumptions – at least the following sorts of beliefs, in the agent:

- BEL.1. general beliefs about the human good;
- BEL.2. general beliefs from sciences other than the science of the good, e.g. medicine, sailing, carpentry, farming and so forth;
- BEL.3. general beliefs about the agent and his or her circumstances; and
- BEL.4. particular beliefs about the agent and his or her circumstances.

Of these beliefs, we shall say that general beliefs about the good, of the sort (BEL.1), are beliefs about the means to happiness of a *higher level* than any of the other beliefs, since these general beliefs about the good are most likely to show up over a wide range of deliberations. Thus, within those general beliefs about the good and means to the good, the belief that the happiness of one's children is important to one's own happiness is a higher-level belief about means to one's own happiness than the belief that having a car or a stereo is important to that happiness – the cars and stereos just getting into considerations of the good because of special features of one's particular circumstances. (The basic idea here is that the happiness of one's children – and one's friends – are means that are present, at least implicitly,

⁴⁴ This is what we call the 'principle of real reference': see §2 above.

in practically all one's deliberations, being more like major premisses in the account of means, while cars, stereos and the like are present only by virtue of minor premisses that speak to other less general desiderata.) We might indeed speak of children, parents and friends as 'at very nearly the highest level of means to our happiness'.⁴⁵

Turning now more generally to (BEL.1)–(BEL.4) – beliefs of the sorts that we have suggested are involved in the production of the belief (BEL.) as to which particular action is the really best means to one's maximal good – we claim that there is hardly a belief in the agent's *entire web of belief* that is not either actually involved or potentially involved in deriving the identification, in (BEL.), of the action which constitutes the best means currently available to the agent's maximal available good. For any expression that occurs in (BEL.) – such as 'good', 'health', 'friend', 'love' and so forth – will implicitly evoke dissonance or consonance with higher level beliefs about the kinds *good, health, friends* and so forth.⁴⁶ And each obstacle or opportunity that presents itself in new perceptions of the situation at the

⁴⁵ The significance of this point will emerge later with respect to some apparent counter-examples to what we shall offer as the Socratic account of love. (See Chapter 11 below, e.g. n. 63.)

⁴⁶ Suppose that the beliefs in Penner's belief-structure to which he appeals in deciding to go and shovel his grandmother's sidewalk after a snowstorm include the bad effects on the health of older people of certain sorts of strenuous activity, the heaviness of the snow in this snowstorm, the good to Penner of his grandmother staying healthy (his love for her), the conflicting schedule of ice-hockey practice, his chances at an athletic scholarship, and so on. Then from these and other beliefs, he may conclude that all things considered it will be best for him to get over to his grandmother's house as soon as possible to beat her to the shovel, and thus to miss hockey practice. In that case, the action he is doing is identified by the totality of those beliefs, organized into the argument-structure he gives to them in deciding that this is the best action. The action is that action (he believes) that is best for him in the circumstances: the action of shovelling the recent heavy snow that will preserve his grandmother's health, her health being an important part of his own good, even though it means missing hockey practice, and so forth. It will then be easy to see how the sorts of considerations invoked above bring in such higher-level considerations as that of the good. For with the considerations of hockey practice and an athletic scholarship conflicting with other considerations, we must appeal to wider considerations of good, health, love and the like in order to decide on the action that is to be preferred.

That 'and so forth' – in the list, three sentences back, of the beliefs that entered into Penner's particular decision on the action to be done – is required because we do not have the action properly identified unless we have accounted for *every* consideration that, explicitly or implicitly, plays a role in our reasoning. Take another case. Jack murders Peter. The fact that the law-courts, in condemning Jack for this particular action, disregard many of these beliefs (the 'whole story', so to speak), provided only that certain conditions laid down in the law for first-degree murder are satisfied – their satisfaction being sufficient for condemnation – shows only what St Augustine said long ago: that the law is for sinful man while only God sees the truth because only He sees into men's hearts. All the law cares about is that the action – whatever the rest of the story about it – is *voluntary under the description* 'guilty in the first degree of murdering Peter'. (Both Anscombe and Davidson insist upon this idea of actions being voluntary under a description.) This is to say that the law *does not care* exactly what action was done – except to the extent that it falls under the conditions for legal responsibility for a particular law-defined crime. Similar remarks might be made about many (deontological) moral theories. What we are suggesting is that knowing what action was done requires knowing *all* the beliefs that played into arriving at belief (BEL.).

moment, and accordingly calls for fine-tuning of our selection of action, will require us to assess the relevance of new kinds to kinds already in our web of belief – to a few that are relatively more specific, and probably to most of the higher-level kinds. And the relevance of new kinds will in turn prompt the need for new examination of the particular circumstances of the situation.⁴⁷

Looked at from another point of view, what is explicitly involved in any particular deliberation may be fairly slight – e.g., there may be little more than the (surely nevertheless necessary) identification of this person here as someone who is loved and who is in need in a particular way. (We can hardly do without *this* much that is reasonably explicit.)⁴⁸ Even so, this deliberating will take place against a background that potentially includes all of my beliefs of all four kinds above, i.e. (BEL.1)–(BEL.4). For even if I act fairly straightforwardly and without much explicit thought, I am at least supposing that there is nothing in the rest of my beliefs of a sort that would make it better to delay action for purposes of further consideration; and *that* is, implicitly, making a judgement that involves my entire web of belief.⁴⁹

⁴⁷ In the first example in the preceding note, the presence of a university coach at this particular hockey practice, and the availability (or unavailability) of Penner's mother's car to get him more quickly from his grandmother's to practice, may both force reassessments – again leading to considerations of overall good. The essential point here is that practical reasoning, like Socratic theoretical reasoning that attempts to identify such things as justice, courage and the like, is essentially *dialectical*. We don't start with our most general beliefs and then plug in particular circumstances. Rather, what more general beliefs turn out to be relevant will change as we learn more about our particular circumstances: see on Pentheus in n. 49 below. Just so, new knowledge of our particular circumstances will show new kinds as relevant, and so prompt an interest in new generalities – and so on, till the costs of further deliberation are greater than the good of deciding now. (In this respect, we find much to agree with in Wiggins 1980.)

⁴⁸ A mother does not rush to save her child *purely on instinct*, whatever those with a sentimental regard for the power of maternal instinct may say. The particular action will hardly take place unless the mother comes to believe (a) that there is a child in the situation that is in a certain sort of distress, e.g. drowning, in pain, feeling threatened; and (b) that the child in question is *hers*. The love in question is a love intellectually directed at a particular child in a particular relation to this woman in particular circumstances. It cannot thus be *purely* instinctual – not unless perceptual beliefs about a situation one faces (in all its particularities) could themselves be purely instinctual (indefinitely many instincts?!). We (Penner and Rowe) take it that the loving motive is the desire to save this child in this relation to the woman in these circumstances. We do not countenance a division of labour between purely instinctual (undirected) love-of-child and directed desire to do such an act. For an indication of our reasons, see n. 50 below; and more generally on what we shall call 'brute' desires (to be introduced shortly below), Chapter 11, §6 below.

⁴⁹ Take the case of Pentheus in Euripides' *Bacchae*, deciding what to do with Dionysus once he has captured him. Pentheus' initial decision to throw Dionysus into prison does not explicitly consider the question of a scouting party to see what he, Pentheus, is up against with these orgies up on the mountain. But it is surely fair to say that he thinks he has considered what he needs to consider in order to decide to throw Dionysus into prison. So it is, at that point, an 'all things considered' judgement to throw Dionysus into prison. But when Dionysus brings up the question of just

So much, then, for purposes of introducing Socratic intellectualism as we understand it. Several consequences of this way of understanding action are worth singling out here. Take first the fact that Socratic intellectualism generates all action from the single desire, common to everyone, for his or her own maximal available happiness or good over the rest of his or her life – or, in the present version, for whatever particular action may constitute the best means currently available to his or her own maximum available happiness or good over the rest of his or her life. One very striking consequence of this is an utter rejection of the perspective involved in Plato's parts of the soul doctrine (as also in Aristotle's psychology of action, following Plato's). According to that doctrine, some actions of the sort Aristotle would call voluntary (cf. n. 41 above) proceed *not* from any generalized desire for the good at all, but from brute irrational desires such as those for drink, food, sex and so forth – desires that involve no reference whatever to the good.

This particular Platonic–Aristotelian belief–desire psychology offers explanations of the following sort (we here pick up and develop material already in n. 2 above):

- (i) pick any brute – irrational, good-independent – desire such as thirst; then
- (ii) find a belief as to where there is some available liquid that will satisfy that desire; then
- (iii) the action of availing oneself of that liquid will result.

Here, it is alleged, we have a belief–desire explanation of a particular act, where the desire-half of the explanation is a brute irrational desire. The reasons why we (Penner and Rowe) believe Socrates would have rejected this particular version of belief–desire psychology are to be found at Penner 1990.⁵⁰ Here we wish only to make clear that Socratic intellectualism has

such a scouting party, with an explicit appeal to military intelligence and an unstated appeal to Pentheus' prurient interest in these orgies, Pentheus adjusts his deliberations immediately to such considerations, and makes a new 'all things considered' judgement. In Gill (forthcoming), Penner speaks about this process as *tracing a pathway through one's web of belief*.

⁵⁰ The argument, in brief, is this: that an action cannot be generated purely as a result of the co-occurrence of *this* belief and *this* desire, since in general we will have *many* beliefs and *many* desires – on whatever principle of counting we employ for beliefs and desires. Suppose there are six other beliefs and seven other desires. Then from the co-occurrence of these six beliefs and seven desires there might be as many as forty-one *other* actions brought about simultaneously with the one that is supposed to be explained from *this* belief and *this* desire. Since that is evidently impossible, some further explanation besides simple co-occurrence will be what is needed of just why *this* belief and *this* desire get to be the pair that generates the actual action done. What is needed is not co-occurrence, but some *integration* of a relevant belief and a relevant desire. We have explained this integration above (through the substitution of a particular belief about what is best into a generalized desire for whatever is best). In his O'Neil Memorial Lectures, Penner has argued that there are similar difficulties

a perfectly good alternative account of acting out of a desire for drink, in which the desire-half of the belief-desire pair that generates the action is not the good-independent desire for drink in the schema above, but *the* generalized *desire for good*. The account runs as follows:

- (iv) a particular thirst occurs;
- (v) that thirst is *represented* in a *belief* as to the contribution that satisfying that thirst would make to the agent's overall good;
- (vi) that belief is incorporated into the totality of beliefs in terms of which the agent decides which particular action now available constitutes the really best means to the agent's maximal happiness over a complete life;
- (vii) the resulting belief as to which action is best – in this case, an action in which one drinks – coalesces with the generalized desire for good, and produces the executive desire to do this particular action of drinking.

There is thus no need for actions generated by belief-desire pairs in which a brute appetite like thirst constitutes the desire-half of the belief-desire pair. (In the schema just presented, the way thirst appears is not in the desire-half of the explanation, but *as represented in the belief-half*.) We are not saying that this appetite of thirst does not occur physiologically, nor are we saying that our perception of it fails to awaken our thought about what to do. Indeed we insist that it *does* do just this, as much as perception of any large rock put in our pathway would awaken our thought about which way to walk. The point is simply that the thirst does not function as the desire-half of a belief-desire explanation.⁵¹

Now for a second consequence of Socratic intellectualism. From the fact that all actions are generated by a single generalized desire, we get that good people do not differ from bad people because of any difference in fundamental desire, i.e. what we have labelled as (DES.): the desire for *whatever action may be* the best means currently available to me, in the circumstances I am in, and so on. That is, good people will not differ from bad people because of any difference in fundamental desire as that fundamental desire is *before* the substitution into (DES.) of the belief (BEL.) as to which particular action will be best for the agent.⁵² The result is that any

for the stock Butler–Hume–Sidgwick–Broad–Feinberg–Nagel arguments for the possibility of pure altruism. (We are not saying that this exhausts the possibilities for a defender of pure altruism. The point for the moment is just that the Platonic–Aristotelian options fail.)

⁵¹ See preceding note.

⁵² See n. 30 above, and text to n. 30. We need to make the distinction in question – between the desire as it is before the decision to do just this particular act has been taken, and the desire as it is once the belief that this particular act is the really best means available has been substituted in – because of our decision to declare that there is an identity through change involving the desire to do whatever means is best and the desire to do the action which is in fact the best. Cf. also *Meno* 78A6–8.

differences there are between people in their actions – aside from their being different individuals – must come from differences in beliefs of the sorts (BEL.1)–(BEL.4), from which the belief (BEL.) is derived, this belief (BEL.) then coalescing with the fundamental desire (DES.), and turning it into the relevant agent-particular executive desire (EXDES.). This is what allows the possibility of explaining how knowledge might come to be identified as the ultimate good that everyone desires and loves – the ‘first friend’. For knowledge alone assuredly secures the happiness that is the object of the fundamental desire (DES.). (Dumb luck is not going to be a real possibility in a complex world.)

There is presumably no need to add, third, that this account of the psychology of voluntary action also explains the extreme urgency of the remark that *the unexamined life is not worth living* (Socrates at *Apology* 38A5–6). For it is one’s beliefs alone that make a difference to whether one does better or worse in the actions one chooses to do. Hence, in terms of what we aim at in action, there is no difference between the course of action which aims at the happiness available in these circumstances and the course of action that aims at the knowledge that is appropriate to this situation. (This identity or near-identity between seeking happiness and seeking knowledge will reappear crucially in the penultimate section of Chapter II below.)

Fourth, the account we have given of Socratic intellectualism will allow us to explain why it is, given that *philia* and *erōs* are motives for acting in particular ways, and that all action is *via* desire for the ultimate good, that *philia* and *erōs* must in every case be understood in terms of an underlying desire for good that is present in all actions (i.e. the ultimate good; the ‘first friend’). Hence we have also made it possible to explain how it is that *philia* and *erōs* are each species of desire for good – and indeed species of desire for the agent’s *own* good. The fundamental desire underlying Socratic intellectualism is explicitly egoistic (*psychologically* egoistic, not ethically egoistic); and this in turn will explain why the accounts of *philia* and *erōs* which we attribute to the *Lysis* are (psychologically) egoistic.

From this fourth point flows a fifth, one which shows some of the commitments we – Penner and Rowe – make in this discussion. This is that on our account, desire for good, *philia*, and *erōs* are all to be understood and to be judged in terms of *action* – in terms of what people *do* rather than simply in terms of what they *feel*. (After all, we have been explaining Socratic intellectualism as a psychology of action.) This point should not be turned into any kind of endorsement of behaviourism – since it is central to identifying what an action *is*, on the Socratic view, *what inferences* from *what beliefs* go into that belief-structure that generates the crucial belief (BEL.) that in turn

identifies the action in question. This is no simple behaviourism, treating internal belief-structure as a kind of permanently sealed black box (and deriving what beliefs someone has from how they behave, rather than, as in Socrates, deriving what behaviour has occurred from what the agents involved believe). Nonetheless, feelings *not* acted on, or which *play no part in* generating actions are, on this account, highly marginalized. This will disturb those who want to insist, 'Love is about how I *feel*, don't you see?' Such protesters are missing the real issue. The real issue is not the existence of feelings, but the existence and importance of feelings *not at all influential on anything we do* – even on our *deliberations* (which are, after all, also actions) – and the question how the Socratic account relates what we call 'feelings' to beliefs and actions. For most such 'feelings' *are* intimately connected with beliefs and actions. Indeed, it is hard to imagine a feeling that does not somehow influence some belief the subject has – only recall our treatment of particular thirsts in (iv)–(vii) above. But it is surely clear that there is no belief that is not capable of influencing action – given the kind of dialectically holistic procedure deliberation actually is. So we feel perfectly free simply to ignore the supposed existence of feelings not at all influential on anything we do.

* * * * *

So much, then, by way of preparing the reader for a second trip through the *Lysis* (Chapter 11 below). We have warned the reader, first, that we shall be employing somewhat different assumptions about what is involved in Socratic dialectic from those employed by those who impose on the dialogues the rules of the so-called 'Socratic elenchus' (and that we expect thereby to gain a much greater unity for the *Lysis* as a whole); second, that we shall be identifying the 'first friend' – the sole ultimate good sought in all love or friendship, all *being in love*, and all desire – as a certain knowledge or wisdom; third, that both love (friendship) and being in love will be identified as species of that desire for good; and, fourth, that the desire for good – especially when we see it in its generalized form – will turn out (a) to be the origin of all voluntary action whatever, and (b) to be identical with the desire for the agent's *own* good.