Plato on the Attribution of Conative Attitudes

by Rachana Kamtekar (Tucson)

Abstract: ####.

1. Introduction

Plato’s Socrates famously claims that we want (βουλέσθαι) the good, rather than what we think good (Gorgias 468bd). My paper seeks to answer some basic questions about this well-known but little-understood claim: what does the claim mean, and what is its philosophical motivation and significance? How does the claim relate to Socrates’ claim that we desire (ἐπιθυμεῖν) things that we think are good, which

1 A brief note on terminology: This paper is about conative attitudes in general, and while ‘desire’ is the most felicitous generic term for conative attitudes in English, my generic term will have to be ‘conation’, for a number of reasons. Plato sometimes distinguishes between conative attitudes by their objects and sometimes reserves particular Greek verbs for attitudes so distinguished. So, for instance, at Charmides 167e, Socrates says that βουλέσεις is for the good and ἐπιθυμεῖα is for the pleasant; it has become fairly standard to translate these, respectively, ‘want’ (or ‘wish’ or ‘will’) and ‘desire’. Plato may not always have the same distinction in mind: in the Gorgias Socrates says that wicked tyrant may kill a good and fine man if he wants (ἀν βουλέσθαι) (511b) and in the Republic that a person may do everything he wants (βουλέσθαι) except what will make him virtuous (445b), but obviously what is wanted in these cases is not good; perhaps, however, it appears good to the agent (cf. Republic 360bc and 362b). It has been argued that Republic 437b-38a distinguishes among types of conation by saying that we βουλέσθαι that which appears good to us whereas we may ἐπιθυμεῖν something even if it appears bad to us (Irwin 1995, 206 f.). On the other hand, it has been argued that this passage only says that desires ought to be individuated and defined by their natural objects, saying nothing about the existence of good-independent desires (Carone 2001, 117 f.). Saying that thirst is for drink, not good drink, might be compatible with saying that thirst (insofar as it is a desire) is for drink, which appears good to the thirsty person insofar as she is thirsty. (I’m grateful to Stephen Menn for discussion on this point.) So even if the thirst may be only for the drink, it may still be that the person desires the drink as good,
are sometimes in fact bad (Meno 77de): can these claims be reconciled (how?), or do they conflict and indicate a change of mind (about what, exactly?)?

In this paper, I argue that Plato’s various accounts of the relationship between our conative attitudes and the good share a core commitment to understanding the norm-responsiveness of our conative capacities and activities. And I elaborate this core commitment by drawing on similarities in Socrates’ treatment of conative and cognitive attitudes.

The paper proceeds as follows. Beginning with the text from Plato’s Gorgias in which Socrates argues that we want the good rather than what seems good to us, the remainder of this introduction raises questions for this argument which any interpretation must answer. Section 2 discusses, and rejects on textual and philosophical grounds, the currently influential trends in scholarly interpretation of the Gorgias text. My own account, beginning in section 3, discusses parallels between Socrates’ attribution of cognitive and conative attitudes to suggest two reasons to hold that conation is for whatever is really good. First (minimally), as belief aims at the true, so desire aims at the good; second (more substantially), as the doctrine of recollection attributes to us beliefs other than those we avow, in order to explain our cognitive behaviour, so the attribution to us of a standing want for the good explains some of our conative – and indeed cognitive – behaviour: not only what we try to get, but also what we are satisfied by, and finally, what we want to know, is best explained by our wanting the real good. Section 4 contrasts the roles played by the real and apparent good in two styles of explanation: natural teleological explanation, and explanation that seeks to make intelligible the agent’s point of view.

The text from Plato’s Gorgias is part of Socrates’ attempt to convince Polus that although expert orators are able to do what they judge best, perhaps conceiving of the good as the pleasant. And conflicts between what appears good to e.g. thirst and what appears good to reason may be what accounts for akrasia. Finally, Diotima’s speech at Symposium 205d f. seems to treat these types of conation as equivalent. Nevertheless, because a distinction between ‘θέλεσθαι’ and ‘ἐπιθυμεῖν’ will turn out to be important to two discussions of the relationship between conation and the good, in the Gorgias and Meno, I will always render Plato’s ‘θέλεσθαι’ by ‘want’, ‘ἐπιθυμεῖν’ by ‘desire’ and conative attitudes generally by ‘conation’. Kahn 1974 points out that Plato does not have a generic term for the conative, as Aristotle has ᾧσίς; at least the awkwardness of ‘conation’ will remind us of that.
they are not able to do what they want, because they want to get or do what is good but do not know what that good is.

[Socrates:] Does it seem to you that people want (βολέσθαι) the thing that they are doing on each occasion, or that for the sake of which (θεῖκοι) they are doing what they are doing? For example, those who drink medicine on the advice of their doctors, do they seem to you to want what they are doing, drinking the medicine and suffering, or that other thing, health, for the sake of which they are drinking [the medicine]? [...] And those who sail, and who make money by other means, what they do on each occasion is not what they want (τῷ τοῦτῳ ἐστίν ὁ βολέσθαι, ὁ ποιῶσιν ἐκάστῳ) (for who wants to sail and take risks and take on troubles?) but rather [what they want is] that other thing for the sake of which they sail, namely, to be wealthy [...]. Is it not also so in all cases? Whenever someone does something for the sake of something, he does not want this thing which he is doing, but rather that other thing for the sake of which he is doing this (τῷ τοῦτῳ βολέσται ὁ πράττει, ἀλλ' ἐκέλευ ὁ ἐνέκα πράττει) [...]. Now among the things that exist, is there anything that is not either good, bad, or intermediate between these, neither good nor bad? [...] Do you call ‘good’ wisdom and health and wealth and other such things, and ‘bad’ their opposites? [...] And do you call ‘neither good nor bad’ such things as sometimes share in the good, sometimes in the bad, and sometimes in neither, for example, sitting and walking and running and sailing, and again such things as stones and sticks and other such things? [...] Then whenever people do things, do they do the intermediate things for the sake of the good things, or good things for the sake of the intermediate things? [...] Then it’s pursuing the good (τὸ ἀγαθὸν ἀρα διόκοντες) that we walk whenever we walk, thinking it to be better (ἀμένοι βέλτιστον εἶναι) [to walk]. (468b1-2) [...] Then we also kill, if we kill someone, and exile and confiscate his wealth, thinking it to be better for us to do these things than not? [...] Then it’s for the sake of the good that those who do all these things do them. (468b8) [...] But we agreed that we do not want the things which we do for the sake of something, but rather those other things for the sake of which we do these things? [...] Then we don’t just simply (ἐκλέων οὖσι) want to slaughter or to exile from cities or to confiscate wealth, but rather we want to do these things if they are beneficial; if they are harmful we do not want to. For we want good things, as you say, but we do not want things that are neither good nor bad, nor things that are bad. (468c7) [...] Then if you agree with these things, if someone, tyrant or orator, kills a person, or exiles him from the city or confiscates his wealth, thinking it better for himself, when it really turns out to be worse, I suppose he is doing what seems best to him (ὁ δεικνύει ἄνδρον)? [...] Then is he also doing what he wants (βολέσται), if these things turn out to be bad? Why don’t you answer?
[Polus]: But it doesn’t seem to me that he is doing what he wants (βούλεται).

(Gorgias 467c5-68d7)

This passage raises a number of puzzles. One is the question of just what the differences are between wanting something and thinking it good (literally, ‘better’ or ‘best.’) Socrates points out one difference – we can only want something if it is in fact good, whereas we can think something good when it is in fact not – but is this the only difference? Are the two meant to be contrasted as a conative and a cognitive attitude? There is also the question of just what relationship (and not just what difference) Socrates is suggesting between wanting something and thinking it good. He says that we do what we think good for the sake of what is really good, or in pursuit of some real good. The relationship between what we do and what we do it for the sake of (or in pursuit of) bears explaining, especially as Socrates seems to identify the things that we think good with the things we do for the sake of something else (I’ll call this ‘something else’ an ‘end’). Yet it would seem that our ends may also be thought good, and that the things we do for the sake of our ends may be really good and not only thought to be so. Further, Socrates at first uses ‘want’ as if short for ‘want as an end’; this is what allows him to ask, as if they are mutually exclusive options, ‘do people want the thing they are doing, or that for the sake of which they are doing it?’ and to deny that people want what they are doing for the sake of something else however, he goes on to extend the scope of wanting to actions which have good consequences, so to things done for the sake of something else when that something else is good (468c3-4). It is tempting to regard Socrates’ initial claim that we do not want the things

2 Unless otherwise indicated, all translations are my own.
3 Similarly, in the Symposium, Diotima says that love is not of the beautiful, but rather of giving birth in the beautiful (206e3-5) soon after saying that the loveable (τὸν ἀγαπητὸν) is what is really beautiful (204c4, in agreement with what Socrates says at 201a9-10). At 204d3, Diotima starts to explain how love can both be and not be of the beautiful: love is of beautiful things, but if someone asks, “why is love of beautiful things?” the answer is that it enables giving birth in the beautiful. So the point of love, or what love is for the sake of, is giving birth in beauty, not just possession of beautiful things.
4 In the Republic, Plato suggests that we may value things for their own sake or for their consequences or both; justice turns out to be valuable for its own sake and for the sake of its consequences (357bc).
we do for the sake of something else as a hyperbolic expression of the point that they are wanted only conditionally, or wanted only as means, i.e., that we want to kill, banish or confiscate goods (these actions being means) in the belief that these actions are beneficial. But as Terry Penner has pointed out, the Gorgias passage says that we want to kill, banish, etc. only if these actions are beneficial (468c2-5). That is, we want the actual means to our ends, not what we believe are means to our ends (likewise, our ends, to be wanted, must themselves be good). If this is right (and I think it is), then it seems to me that on Socrates’ account of wanting, whether or not we want something will depend not only on facts about our narrowly-speificed psychological states but also on facts about the objects in the world to which these psychological states are related. Aristotle’s discussion in the Nicomachean Ethics of whether the object of wanting is the good or the apparent good raises problems for such an ‘externalist’ account of wanting:

That wanting (ἠθική) is for the end has been stated; to some it seems that it is for the good (πάρασηδον), to others that it is for the apparent good (τὸ φαινόμενον ἀγαθόν). But it follows for those who say that the good is the object of wanting (τὸ θεωρεῖν) that that is not an object of wanting which the man who chooses incorrectly wants (μὴ ἔχοντας θεωρεῖν ὅθεν νεότερος οὐκ ἔχων) – for if it is to be the object of wanting, then it must also be good, but it was, if it turned

---

5 According to Terence Irwin, it is Socrates’ ‘unclarity […] about wanting under a description’ that leads him to say inconsistent things about whether or not we desire the means to our ends: (1) whenever we do x for the sake of y, we want y, but not x (467d6-c1; 468b8-c1), and (2) if we want x and x is not identical to the good, then we want x for the sake of the good (468c3). Irwin’s diagnosis is that Socrates says we do not want various things we do as means to some good on the basis of his descriptions of those things, but it is clear that we do want them under, for example, descriptions which include the end for the sake of which we do them (Irwin 1979), 141 f).

6 Penner 1991, 178 f.

7 I take it that by ‘the apparent good’ Aristotle means ‘that actually existing object which (rightly or wrongly) appears to us to be good’ – and not ‘that (perhaps nonexistent) object which has all the features, including goodness, we think it to have’. For Aristotle thinks it should turn out that the object of wanting is what the man who chooses incorrectly ‘wants’. That the object appears good doesn’t have to be built into the description of the object wanted, i.e., when someone asks, “what does he want?” the answer doesn’t have to include the way it appears to him, because that fact, that the object appears good to him, can be included in the answer to the question, “why does he want it?”
out like that (ἐὰν οὖν ἔστω), bad. But for those who say that the apparent good is the object of wanting, [it follows that] there is no natural object of wanting (ὄνομα βουλήτου), but for each person, it is what seems to be [good]. One thing appears [good] to one person, another to another, and if it so turns out, vice versa. (EN 1113a15-22)

Here, Aristotle objects that the result of the sort of position adopted by Socrates in the Gorgias passage, that we can only want something if it is in fact good, is that the person who chooses incorrectly wants something which is not an object of wanting. Yet, one might continue the objection, it is reasonable to think that one has some positive conative attitude towards the thing that appears good to one, and that this attitude is, psychologically (both phenomenologically and functionally), the same whether one’s judgment about its goodness turns out to be correct or not. So if the attitude is called ‘wanting’ in the case where one’s judgment is true, it ought to be called ‘wanting’ even in the case where one’s judgment is false; but if wanting must have as its object the real good, then it cannot be so called. Indeed, on this view, one could have the same experienced, felt, judged, and action-guiding attitude towards two objects and want one but not the other.

These objections to the Gorgias passage’s rule restricting ‘wanting’ to actually good things are compelling and obvious enough to raise the question: why would anyone observe such a rule? Aristotle’s worry that if the Socratic rule is violated, then there is no natural object of wanting, points in the direction of a motivation, but it is not entirely clear what that motivation is. Is the ‘natural’ object of wanting that which we ought to want, something correct wanting is for? In that case, restricting wanting to actually good things looks like overkill: to be able to evaluate our wants as correctly directed or misdirected, it is not necessary to deny

---

3 Stephen Menn has suggested to me that the βουλήτου, the object of wanting, could be the proper object of wanting (rather than, as I take it above, a possible object of wanting). In this case, Aristotle’s criticism would be that if wanting is for the good, then some people, namely those who choose incorrectly, want something other than the good that is the proper object of wanting. But this is not as powerful a criticism as the one developed above, and Aristotle is unlikely to see it as much of a criticism at all, since it is his own view that people who choose incorrectly want something other than the proper object of wanting — a view allowed by his own preference for saying that wanting is for the apparent good (see below).
that the misdirected ones are wants at all, and denying that they are
wants at all saddles one with the additional burden of explaining why
these attitudes should be judged by the standards of wanting in the first
place.

Consider Aristotle’s own answer to the question of whether the object
of wanting is the real or apparent good:

[...] should we say that the object of wanting without qualification (ἁπλῶς) and in
truth is the good, but for each person [it is] the apparent good; and that to the
good man [it] is the truly [good] while to the bad [it is] any chance thing (τὸ
τὸ[...] – just as in the case of bodies such things are truly healthful as are
healthful for the well-disposed (τοῖς ἀνδρῶν), but different things [are
healthful] for the sickly, and similarly too bitter and sweet and hot and heavy and
each of the others? For the good man judges each thing correctly, and in each
thing the truth appears to him. For in accordance with each disposition (εἴς
there are distinct fines and pleasanths, and perhaps the good man differs most in
seeing the truth in each thing, being, as it were, the standard and measure of
them. For most people the deception seems to come about as a result of pleasure:
for things that are not [truly good] appear to be something good. They choose,
then, the pleasant as good, and avoid the painful as bad. (EN 1113a22-1113b2.)

Aristotle endorses the view that each person wants that thing which
appears to him to be good, but rejects the implication that there is no
natural object of wanting by invoking the good person as the standard:
just as the healthy person’s judgment of sweet, bitter, hot and cold is the
standard of correctness in judgments about these things, the good
person’s wanting is for the ‘true object of wanting’, what wanting
should be for, since the good person’s faculty of wanting functions well
and what appears good to the good person is truly good.6

By contrast with Aristotle’s claim that people want what appears good
to them, should want what is truly good, and would want what is truly
good if their faculties were functioning well, Socrates’ claims in
Gorgias 467c-68d, that we might judge something best and yet not want
it, and that we want things only if they are in fact good, seem bizarre.7
Why should Socrates make such claims?

10 Contemporary philosophers sometimes defend the claim that we want the real as
opposed to the apparent good by saying that when we want something, it is
something in the world rather than something in our head. There is no dispute
2. Trends in Scholarly Interpretation

Perhaps the most common reaction to Gorgias 467c-68d is a sense that Socrates’ argument is invalid, and so I begin my discussion of current scholarship by considering the validity of the argument. According to Kevin McTighe, in Gorgias 467c-468d, Plato is deliberately putting a fallacious argument in the mouth of Socrates so that Socrates can refute Polus, and the fallacy consists in an elision between a de dicto and a de re reading of ‘wanting the good’, and between a subjective and an objective reading of ‘good’. When Socrates and Polus agree that we want the good at 468c5-7, Polus understands by this that we want those things we believe to be good (so at 468a5-8, when Polus agrees that men do intermediate things like walking, running, expropriating and slaughtering for the sake of good things like skill, health and wealth, his agreement is to the claim that we want skill, health, etc., believing them to be good), but by 468d5-7, Socrates is saying that if there is some x that we believe to be good, but this x is not in fact good, then we do not want it.

To establish that Socrates makes the first supposedly fallacious shift between wanting what we believe to be good and wanting what is actually good, McTighe claims that Socrates says, at 468a5-b8, “All

---

about this between Plato and Aristotle; when Aristotle says we want the apparent good, of course he doesn’t mean by this that we want something in our head. Nor does the view that the good is objective determine one’s position on the issue: both Aristotle and Plato believe that the good is objective, but on the matter of what is the object of our psychological attitude of wanting, Plato says it is the real good, Aristotle that it is the apparent good. Nor does the intuition of many of our contemporaries, that they would reject a life of pleasure induced by artificial means (e.g. a machine stimulating our so-called ‘pleasure receptors’), tell one way or the other. A contemporary Aristotelian could explain the intuition by pointing out that the life of artificially-stimulated-pleasure-sensations doesn’t appear good to us. The issue between Plato and Aristotle is, rather: if I think that the good is honour, but the good in fact is virtue, then do I want honour or virtue? The passage above suggests that Aristotle would say, honour (because I believe honour is good), and Plato would deny this. One could attribute to Aristotle the view that because I want honour as good, my attitude of wanting-as-good picks out what is genuinely good in the world; I do not pursue this interpretation of Aristotle here but, later in the paper, I do discuss whether such a view can be attributed to Plato.

11 McTighe 1984, 205 f.
men do intermediate things [...] for the sake of what they hold to be good for themselves [...] but what the text actually has is (1) Polus admitting that when people perform intermediate actions, it is for the sake of the good (468a5-6) (rather than for the sake of what they hold to be good, as McTighe claims on the grounds that this is all Polus would assent to), and (2) Socrates concluding that people perform these intermediate actions, thinking them to be better, for the sake of the good (468b1-4). In (2), McTighe requires it to be the results of the actions that people think are better; while this is possible, the more natural reading is that it is the actions which people think are better: the referent of “οἷμεν βελτίων εἶναι” at 468b2 must be “βασίζομεν ὀταν βασίζομεν”; similarly, the referent of “οἷμεν άμεινον εἶναι” at 468b6 must be “ἀποκτείνομεν, [...] ἐκβάλλομεν καὶ ἀφαίρομεθα χρήματα” at 468b4-6. One might object on behalf of McTighe that if people think their actions better it’s surely because they have good results. But even if that is right, what is relevant to McTighe’s claim that Socrates slips from one object of desire to another is what Polus says people think better, not why they think it better.

A better interpretation than one convicting Socrates of equivocation would be one that could make sense of Socrates inferring, from the fact that we choose or do things that we believe to be good, that we have the goal of coming to possess what is actually good. Polus can certainly agree, without being tricked into it, that we want x, y, and z which appear good to us, and also that we want the real good. He initially

---

12 Ibid., 203.
13 This objection is due to Roslyn Weiss.
14 McTighe argues that to win Polus’ assent, Socrates must initially claim only that we desire the apparent good, and that to this end, Socrates uses examples of goods for the sake of which we act (health, wealth, power), which Socrates believes are not always good. But it doesn’t follow from the fact that they are not always good that they are only apparently good; for example, at Euthydemus 278e-82c Socrates seems to hold that health, wealth, and so on become genuinely good when accompanied by wisdom. And must Socrates only claim that we desire apparently good things to get Polus’ assent? Tad Brennan has suggested to me that if the good for the sake of which we act (in this argument) were objective, then to Socrates’ question, “whenever people do things, do they do the intermediate things for the sake of good things, or good things for the sake of intermediate things?” (468a5-6), Polus would reply ‘sometimes they do intermediate things for the sake of good things, but sometimes they do good things for the sake of intermediate things, and indeed, sometimes they do good things for the sake of intermediate things, and indeed...
takes the two to be equivalent; once they have been distinguished, it is
open to him to say (in the spirit of Cleitophon at Republic 340b) that
what tyrants and orators want is what they consider to be good for them.
He (wisely) doesn’t make this move, for while the move might enable
him to maintain that tyrants and orators do what they want, it would do
so at the cost of his dialectical goal, which is to recommend rhetoric as
really good for its practitioners and not just as believed by them to be
good (as Thrasymachus in the Republic wants to recommend injustice as
really advantageous for those who practice it and not merely to appear
so to them or others). Polus has said, after all, that rhetoric is ‘the finest
of the arts’ (448c), and that it empowers those who use it, power being
something good (466b). Unfortunately for his dialectical goal, Polus has
also allowed that an expert in rhetoric can be ignorant of what is good
(461b). Presumably this is why when Plato has Callicles take over the
argument, he has him take up the position that the good is pleasure or
desire-satisfaction (491e f., explicitly at 495d), for this would seem to
close the gap between what appears to be good and what is good. So a
close examination of Gorgias 467c-68d and the progress of the Gorgias
both indicate that rather than cheating his interlocutor, Socrates is
arguing from a shared but substantial premise.

I turn now to two trends in scholarly interpretation that find in
Gorgias 467c-68d a philosophical doctrine of some seriousness. Accordin
to the first, the passage identifies a special conative attitude, wanting (βολητης), which can have as its object only what is really
good – by contrast with other conative attitudes, for example desire
(ἐπιθυμία), the objects of which might be bad even if they appear to be
good. This special conative attitude explains how the Socratic paradox,
‘no one errs willingly’, could be true: when one errs, one contravenes
one’s want for the real good. At one point in the Gorgias, Socrates
makes this connexion explicit:

things for the sake of bad things [...].’ But Polus is not answering the question,
‘those things which they think best, i.e. to bring about by their actions, are they
[objectively] good, bad or indifferent?’. Rather, he is answering a question about
the ultimate object(s) of our pursuits; in the main text above, I argue that Polus
has good reasons not to maintain that we only want what appears good to us.

15 There is a comprehensive survey of interpretations which distinguish desiring
and wanting from antiquity to the 1980’s in McTighe 1984, 195 f.

16 Note that there is no linguistic connexion between the Greek βουλητης
[wanting] and ἤκων [willing] or ἁκων [unwilling].
Why don’t you tell me this, Callicles, whether or not you think Polus and I were rightly compelled to agree in the foregoing arguments, when we agreed that no-one commits injustice wanting (διασκορπιζόμενον), but all who commit injustice do so unwillingly (δέκομαι)? (*Gorgias* 509e2-7)

Socrates has just suggested that to protect oneself from the harm of committing injustice, one needs, in addition to wanting to avoid committing injustice, a power and craft. At *Gorgias* 467c-68d, Polus and Socrates had agreed that we do not want to do things that are not in fact good; in subsequent discussion they agreed that doing injustice, especially when it goes unpunished, is the worst thing there is (474c-79e). It would seem to follow that those who commit injustice do not do so wanting to; at *Gorgias* 509e Socrates glosses this as ‘committing injustice unwillingly’. So any wrongdoing would be unwilling because contrary to our want for the good. Similarly, at *Laws* 733b-34b, the Athenian says that we all want (διασκορπιζόμενον) pleasure, and that anyone wanting to live pleasantly (τὸν γε βουλέμενον ἥδεως ζῇ) must be licentious unwillingly (δέκομαι) (because licentiousness detracts from a pleasant life).

---

*EN* 1111a25-b3, Aristotle reports and criticizes a view (perhaps the proposal at *Laws* 860d-64c) that acts motivated by anger (θυμοί) or appetite (ἐπιθυμία) are involuntary. First, animals and children, who are motivated by only these conative attitudes (they lack rational wants) nevertheless act voluntarily. Second, there are cases in which we ought to be angry or to have an appetitive desire for certain things, which implies that at least in these cases the angry or appetitive motivations are voluntary, but how can these be voluntary unless the angry and appetitive motivations which we ought not to have are voluntary as well? (Aristotle is not only saying that the actions motivated by appetite and anger are voluntary but also that if one ought to desire, feel angry at, and have an appetite for certain things, then these motivations must be voluntary as well.) Third, if acts caused by errors in calculation are nevertheless voluntary (this is a departure from the view that acts contrary to wish are involuntary), then the same should hold for acts due to erroneous appetites and anger. These arguments seem to be targeting a view which recognises several sources of motivation and regards action motivated by, or in accordance with, only one of them, wish (βουλήσι), as voluntary. However neither this discussion nor the parallel discussion at *EE* II.7 restricts wish to what is in fact good; the question of whether the voluntariness of non-rational motivations depends at all on their contravening wish seems to be independent of the question of whether wish is for the actual or apparent good.
Interpretations in this vein are elaborated in different ways. So for instance, in his great commentary on Plato’s _Gorgias_, Eric Dodds writes, “The distinction between what people think they want and what they ‘really’ want [...] evidently originated in the attempt to understand Socrates’ saying that no one does wrong willingly. But it is perhaps only fully intelligible in light of Plato’s later distinction between the ‘inner man’ who is an immortal rational being and the _empirical_ self which is distorted by earthly experience (cf. _Rep_. 611b f.) This interpretation supplies the two kinds of conation with two separate subjects. However, positing a true or rational self as the subject of a special kind of conation fails to explain why this conation has to be for the real, rather than the apparent, good. Why not maintain instead (as Aristotle does about the good person) that the true or rational self wants the apparent good and its appearances are correct?

Norman Gulley avoids positing separate subjects, but instead argues for a hypothetical interpretation of wanting. To want something it is necessary to think it good, so those who do not yet know the good could not strictly speaking want it; however, people in this condition can still be said to want the good and so to do wrong ‘unwillingly’ in the sense that their actions are contrary to the want they would have had had they known the good (although not in the sense that their actions are compelled or contrary to the actual desires they have). Yet Socrates speaks of wanting as actual rather than hypothetical, and as unrestricted: in _Gorgias_ 467c-468d, people in general (not just those who know the good) want (actually rather than hypothetically) that for the sake of which they are acting: health, wealth, the good. The conative attitude contravened seems to be actual.

Taking the road not taken by Gulley, Heda Segvic proposes that for an agent to want something, he must know it to be good: an agent desires (ἐπιθυμεῖν) something

---

18 Dodds 1959, 236. Dodds goes on to complain, “Whatever its theoretical justification, in practice it too easily becomes an excuse for dictation [...] [quoting Barbara Wooton] ‘Sooner or later what I “really” want to do turns out to be a polite paraphrase for what you think I ought to want to do.’”
19 See Gulley 1965; on the hypothetical sense of “wish”, see esp. _ibid_. 89 f.
20 George Rudebusch charges that the distinction between actual and hypothetical (or ‘counterfactual’) desires cannot make sense of the experience of coming to appreciate one’s true desires; the distinction that helps us make sense of both this experience and of Socrates in the _Gorgias_, Rudebusch proposes, is between felt desire (desires I believe or feel I have) and true desire (desires I really have – and can discover in the process of cognitive enlightenment). (Rudebusch 1999, 44 f.) I am not sure whether this coming to appreciate our true desires is an experience or a belief about an experience.
just in case he believes that thing to be good, and he wants (βουλεύειν) the thing he desires just in case that thing is good in the way he believes it to be good. But this cannot be right: since Socrates attributes to orators and tyrants wants despite their ignorance (they do just about nothing they want to [οὐδὲν γὰρ ποιεῖν ἄν χωρεύει], 466d8-e1), it cannot be necessary for someone to know or even believe that a thing is good in order for him to want it. On the basis of her account of wanting as requiring knowledge, Segvic infers, from Socrates’ 509e5-7 claim that no-one ever does what is unjust wanting to (βουλεύειν), that when Socrates asks Callicles whether wanting to avoid wrongdoing suffices for avoiding it or whether one also needs some power (δυναμιν τις καὶ τέχνη) to avoid it (509d7-e2), the correct answer is ‘wanting to avoid wrongdoing suffices for avoiding wrongdoing.’ But Socrates’ own answer is that to avoid wrongdoing, one also needs to acquire some power or skill (510a3-4). And it seems right: surely one could want to avoid wrongdoing, even on the basis of knowledge of the badness of wrongdoing, without knowing how to avoid it.

Indeed, the Socratic paradox ‘no-one does wrong willingly’ does not need to be explained by appeal to a contravention of the special conative attitude of wanting, which gives us reason to think that that doctrine is not the motivation for claiming that we only want what is in fact good. Perhaps an action done in ignorance is an unwilling action because the agent might not have done it, or might not have wanted to do it, had he had knowledge. In the Protagoras, Socrates says nothing to distinguish kinds of conation, but instead supports his claim that no-one willingly (ἐξονύ) pursues what is bad or what he believes is bad (358d) by arguing that pursuit of what is bad is always due to ignorance (rather than to any non-rational motivation). As far as the question whether the object of

21 Segvic 2000, 10. ‘Wanting’ (like ‘knowing’) names a state that is at once conative and cognitive: the soul’s receptivity, mediated by a correct conception, toward the property of goodness in the object of desire. Segvic yokes this account of wanting to an attractive interpretation of the Socratic analysis of akasia as a condition of ignorance (ignoresence of what is best, which allows one to believe false appearances, as well as ignorance of one’s own ignorance).

22 Cf. Meno 78b, where Socrates concludes that the difference between virtuous and vicious is in their power to get good things, for the wish for good things is common to all.

23 Cf. Protagoras 345de: “For I pretty much believe that no-one among the wise thinks that any person willingly (ἐξονύτατα) makes a mistake or willingly (ἐξονύτατα) does shameful and bad things, but they know that all who do shameful and bad things do them unwillingly (ἀσυνετές).” Cf. also Republic 565bc, where the naturally best organized temperaments become oligarchs ‘unwillingly’ under the influence of the drones.
desire is the real or the apparent good, the text of the Protagoras seems neutral. For Socrates’ analysis of what happens when a person seems to act contrary to his judgment of what is best (which most people think is due to the person’s being overcome by pleasure) is that the person’s judgment of the goodness or badness (or pleasantness or unpleasantness) of a particular outcome is distorted by appearances, since this judgment is not formed on the basis of a knowledge by which one could measure good and bad, or pleasant and painful. The result is that, for example, the proximate pleasure appears greater than the distant one when it is in fact not (354e-357a). Socrates is here defending a theory of motivation according to which we want, and in our actions aim at, our own good, but this could be understood as the view that we want, and in our actions aim at, what is really good, or as the view that we want, and in our actions aim at, what appears best to us. Both views are equally intellectualist; on both views, so-called akratic actions are the result of our acting on a false appearance of goodness. (Socrates concentrates on false appearances that are mis-assessments of amounts of pleasure, but someone could have and act akratically on the false appearance that some pleasure is bad, thinking, like Protagoras, that pleasures can be good or bad [351d].) Further, Aristotle’s discussion of the distinction between the non-voluntary and the involuntary in the Nicomachean Ethics (1110b16-24) targets the claim we seem to have found in the Protagoras, that ignorance is sufficient for involuntariness. According to Aristotle, an action performed in ignorance is non-voluntary but need not be involuntary; for a non-voluntary action to be involuntary, it would have to be accompanied by pain and regret on the part of the agent – these, presumably, being evidence of a contrary conation. Of course, one might argue on the other side that Gorgias 509e is evidence that Plato too recognized the insufficiency of ignorance for involuntariness and required contravention of an actual conation. And Aristotle’s target may be not the Protagoras, but Gorgias’ Encomium of Helen, which argues that if Helen was persuaded to elope with Paris (an

---

24 For contemporary interpretations of involuntariness in terms of ignorance rather than contravention of a want, see McTighe 1984, 228-236, and Santas 1964. Santas observes the distinction between desiring and wanting but does not understand involuntariness in terms of the contravention of wanting; rather, he takes ignorance (of what virtue is, or that virtue is good for one, or both) to establish the involuntariness of vice.
opinion she would have formed because ignorant), then her eloping was blameless and akin to being abducted by force. This would give an immediately obvious point to Aristotle’s requirement of pain and regret on the part of the involuntary agent.

Perhaps we cannot settle on textual grounds the question of whether or not Plato introduces our wanting the real rather than the apparent good to explain the Socratic paradox. However, it does seem more likely that the doctrine that wrongdoing is unwilling should be a consequence of some deeper philosophical commitment about our orientation towards the good, rather than the other way round. But just what is the nature of this orientation such that Plato characterizes it in terms of wanting, and why does he think we have it?

According to Terry Penner, the Gorgias passage presents a Socratic theory of action according to which the desire or want (Penner does not distinguish these) that motivates any action is for the real good, so that what we do is always the result of our desire for the real good combined with our true or false beliefs about the means to or constituents of that good. Penner glosses Gorgias 467c-468d’s puzzling restriction of desire to the real good as follows:

(i) We wish an end only if the end desired is the best (most beneficial) end available in the circumstances – the greatest amount of happiness attainable starting from where one is at right now – and

(ii) We wish a given sequence of means to that end only if it is the best sequence of means available in the circumstances.

If either the end desired or the sequence of means desired is not the best available in the circumstances, then the sequence of means is not wished, but merely seems best.\footnote{Penner 1991, 182. Penner presents the above as “correct[ing]” (181) an earlier account, (W), according to which “A wishes action X if and only if: (i) it seems best to A to do X, and, in addition, (ii) doing X will in fact be the best thing for A to do in the circumstances A is in” (173 f.). Since Penner then says “I shall argue that the account of wishing in (W) does not constitute a redefinition of ‘wish’, but is a claim about what wishing really is” (177), I had originally supposed that his correction amounted to introducing, into clause (ii), the distinction between wishing for an end (only if it is the most beneficial in the circumstances) and wishing for a means (only if it leads to that good end). However, Penner (personal communication) tells me that...}
To explain Socrates’ paradoxical claim that if you do not get the good that you wanted by doing the action you thought best, then you did not do the action you wanted to do, Penner attributes to him the following view about the identity of actions: the identity of an action is given by its entire means-ends structure, all its actual or intended consequences. So (to abbreviate one of Penner’s examples) if the tyrant thinks it best to kill his minister in order to consolidate his own power, but in fact killing his minister brings about a coup as a result of which he is dethroned, then the action the tyrant wanted to do (killing his minister with the consequence of consolidating his power) is not the action he did (killing his minister with the consequence of losing his throne).

However, the thesis that the identity of an action is determined by all its actual and intended consequences is incompatible with the text. Plato’s text requires an interpretation according to which the action judged best is the same as the action done (but not wanted). According to Gorgias 468d, if a tyrant puts someone to death because he thinks that doing so is better for himself (although it is actually worse), then the action he is doing is the action that seems best to him. But on Penner’s way of individuating actions, the action judged best and the action done (but not wanted) would also have to be different actions: the action judged best would be killing the minister with the consequence, consolidation of the tyrant’s power; the action done would be killing the minister with the consequence, the tyrant dethroned. Further, since the intended consequences which identify the action wanted (on Penner’s account) don’t obtain, the action wanted does not exist, and so, a fortiori, is not good. This is an unfortunate consequence for a view according to which all desire is for what is really good.

Fortunately, the claim that the object of desire must be good in fact does not require the thesis that actions are to be identified by all their

---

26 Penner 1991, 188 f.
27 For example, Penner 1991, 188 f., writes, “The action of killing his chief minister that the tyrant wants to do is not the action of killing that the tyrant has the ability to do”; and, in reference to another example, “The army she [Private Benjamin] wanted to join was the army with the yachts and the condos (which the recruiter lured her with); but the only army she was able to join was the army in the real world [...]”.

consequences. Penner himself motivates the claim by appeal to Donnellan’s account of the referential use of definite descriptions. According to Donnellan, when one uses a definite description such as ‘the man drinking a martini’ referentially (as opposed to attributively), for example, when one asks the question, ‘Who is the man drinking a martini?’ of some particular individual, then, even if that individual is in fact drinking a glass of water, one’s words, ‘the man drinking a martini’, pick out that individual, because one intends one’s words to refer to him. Similarly, Penner says, one’s desire picks out what is in fact beneficial – even if this conflicts with one’s conception of what is beneficial.

We should ask whether it is appropriate to model desire on reference. One consideration against doing so is that ‘refer’ (like ‘see’ or ‘know’) is a success-verb, whereas ‘desire’ seems more like ‘believe’: one might think that desire aims at the good as belief aims at the truth and that this means that there are restrictions on what one can desire or believe, but nevertheless, just as belief’s aiming at the truth does not guarantee that the beliefs we actually have are true, desire’s aiming at the good does not guarantee that the desires we actually have are for the good.

---

28 Donnellan 1966. By contrast, one can use the same definite description attributively: the president of Teetoaller’s club, misinformed that someone at the party is drinking a martini, may ask, ‘who is the man drinking a martini?’ In this case, because his intent is to pick out whoever is drinking a martini, and no-one fits the description ‘the man drinking a martini’, his expression fails to refer.

29 What is one’s conative attitude if the thing judged best is not in fact good? Penner 1991, 186, calls it an apparent desire, but in a later paper, he explains that when one performs some particular action, one’s desire picks out that action (in addition to whatever is actually good), with all its actual consequences – even if these conflict with one’s conception of its consequences – and when these consequences are bad, there is an ‘incoherence’ in one’s desire (Penner / Rowe 1994, 6 f). But conceding that the actually-bad action is desired, even if incoherently, seems to undo the original contention that the objects of desire must be actually good. On the other hand, withholding the appellation ‘desire’ from the attitude the agent has towards the action he wrongly judges best seems arbitrary, since this attitude plays the functional role of a desire in the production of action.

30 I do not understand Penner’s justification for extending Donnellan’s account of reference to psychological states (especially since some psychological states, such as imagining, seem to be indifferent to the real thing). Penner may think that we are somehow constrained to desire that which is good just as (and here I quote the Parmenidean metaphysical and epistemological thesis he attributes to Plato) “what we [find ourselves constrained to] think and speak of [we should admit]
Further, while *Gorgias* 467c-468d seems to prohibit saying that we can ‘want’ things that are not in fact good, later in the dialogue, Socrates brings up cases of desires for things which do not seem to have to be good. He tells of a part of the soul in which desires (ἐπιθυμία) reside, saying that in people who lack wisdom, this part is undisciplined, like a leaky jar which constantly needs to be refilled (493a-494a). Examples of this sort of desire include hunger and thirst, painful conditions the relieving of which is pleasurable (494ab). But pleasures may be good or bad (495a), and the fact that these desires are in need of restraining (491de, 493d-494a) suggests that they are not simply for what is in fact good. Similarly, in the *Lysis*, hunger and thirst are desires that may be beneficial or harmful, in themselves neither good nor bad (221ab), which suggests that their objects may be good or bad for the desiring subject. And a passage in the *Meno* (discussed below) allows that we may desire actually bad things, albeit thinking that they are good.

If Socrates does not think that all desires (or positive conative attitudes) are for the good, and if the claim that no-one errs willingly is itself a consequence of some doctrine that we have a special orientation towards the good, what is that doctrine? What is the nature of our orientation, and why does Plato suppose we have such an orientation? Answering these questions is the task of the second half of this paper.

---

must be” (Penner 1987, x; Penner’s interpolations). This move, however, makes false belief as mysterious as misdirected desire.

31 Penner has two ways of dealing with such texts. First, he acknowledges that the *Gorgias* has non-Socratic ‘elements’, of which this passage is an example (Penner 1991, 150, n. 3). Any interpretation of the dialogue that can either reconcile or explain the tension between the two claims about desire will be more convincing than an attribution of an incorrect attribution of the two claims to different authors. Second, Penner allows that there may be desires for things that are not actually good, but suggests that these are restricted to hankeries, itches, and drives for pleasure – desires which cannot motivate action – with the consequence that his account, which is an account of the desire to do some particular action, stands (ibid., 201 f., n. 45). Presumably the reason to consider some of non-motivating desires harmful, and to seek to restrain them, would then be that their presence creates psychic disturbance, rather than that they can lead to bad actions. But why would Plato insist that desires which explain actions must be for the real good while allowing that desires which explain (say) how we feel need not be?
3. Grounds for Attributing Psychological Attitudes: Parallels between Cognition and Conation

We may begin by considering some passages in which Socrates seems to make claims about belief and assertion very like those we have seen him make about wanting in Gorgias 467c-468d. For although Socrates often accepts that what his interlocutors volunteer as their beliefs are genuinely their beliefs, he also routinely attributes beliefs to them that contradict the beliefs they volunteer, and denies that they have some of the beliefs they volunteer.

The following is an exchange in the Gorgias just preceding 467c-468d, in which Socrates, having introduced the distinction between orators and tyrants doing what they want and their doing what they judge best, affirms that they do what they judge best, and then attributes to Polus the view that they have no power, whereas Polus’ stated view is that they have great power.

[Socrates:] [...] I say, Polus, that both orators and tyrants have the least power in their cities, as I was just saying. For they do nothing they want, so to speak, yet they do what seems best to them.
[Polus:] Then isn’t this having great power?
[S:] That’s not what Polus says.
[P:] I don’t say it? I do too say it!
[S:] [...] you do not, since you say that having great power is good for the one who has it.
[P:] I do say that.
[S:] Do you think it’s good, then, if someone does things which seem to him to be best, when he has no intelligence? Do you call this ‘having great power’ too?
[P:] I don’t. (Gorgias 466de)

Similarly, in a later passage Socrates tells Polus, who is maintaining that suffering injustice is worse than doing it, that not only he, Socrates, but also Polus and everybody else believes that doing what’s unjust is worse than suffering it and that not paying one’s due is worse than paying it.

[Socrates:] [...] I think that you and I and all other human beings consider doing injustice to be worse than suffering it, and not paying the penalty to be worse than paying it. [...] Which seems to you worse, Polus, doing injustice or suffering it?
[Polus:] To me, suffering it. (Gorgias 474bc)
What licenses Socrates to attribute the beliefs he does to Polus? And what licenses Socrates to deny attribution of beliefs to Polus when Polus himself avows them? One possibility is that ‘X believes p’ is a way of saying, ‘X ought to believe p, given what else he believes’ or ‘X will believe p, when he sees that it is entailed by q, r, and s, which he believes’ or ‘X would believe p if he just paid attention to the evidence and thought it through’. Beliefs are dispositions to make particular judgements or avowals in particular circumstances and we may not know what we would judge or avow in these circumstances. For example, our beliefs (or their expression in particular circumstances) are typically governed by rules of consistency and logical entailment, and we may be unaware of the bearing of these rules on our beliefs – which ones conflict with one another, and which ones we will abandon when we find they conflict, and so on. Socrates’ attributions (or denials of attribution) may be based on prior dialectical experiences with interlocutors who have similar belief-sets.

Contemporary philosophers sometimes say that accepting such norms is constitutive of believing – so that I cannot get into the business of forming beliefs without also signing on to rules of consistency and entailment. In addition to these rules, our beliefs seem to be subject to other norms. It is not only difficult, but usually a fault, to adopt beliefs at will, that is, to adopt a belief in the absence of sufficient evidence. This indicates that we subscribe to some norm such as ‘believe p only on the basis of sufficient evidence’. William James’s classic examples of believing at will – believing in God, or believing, and thereby making it true, that you can perform some athletic feat – are, as James himself points out, non-standard: limited to circumstances in which our options are live, forced, and momentous, and in which what we believe cannot be settled on intellectual grounds[3]. In many cases, it may not even be

---

32 Cf. Gorgias 516d, where Socrates says Callicles is committed to the view that Pericles was a bad statesman given what he has asserted, and 495e, where Socrates says Callicles doesn’t agree with his own claim that pleasure and the good are the same and that knowledge and bravery are different, when he comes to see himself rightly. A number of examples of cognitive and conative attribution are discussed in Brickhouse / Smith 1992; they explain these as attributions to the ‘true self’ which elenchus helps us to come to know.

33 See James 1911, esp. 3, 11.
possible for us to adopt a belief at will – e.g. believing that the number of stars is even (or odd) – in the absence of evidence and expert testimony. At any rate, we judge beliefs formed in the absence of evidence (broadly construed) to be faulty; we want our beliefs, as far as possible, to be true; and we justify our beliefs by adducing evidence for their truth. Some contemporary philosophers explain the norms and expectations that regulate or at any rate characterize believing by saying that belief has a constitutive aim: it is because, in believing, we seek to know or to believe the truth, that we don’t (in most cases) believe at will, or believe contrary to our evidence.

By analogy, might we suppose that it is a constitutive feature of (some kind of) conation that it aims at the good? After all, we fault conative attitudes for being whimsical; we think we ought to want or desire particular things only if they are good (or appropriate) objects of conation; we try to regulate our conative attitudes so that they are consistent, and also for appropriate or good objects. Our conations are at least somewhat responsive to the apparent goodness or badness of their objects; we may even justify them by adducing evidence for the goodness or appropriateness of their objects; finally, as we cannot acquire beliefs at will, we also cannot acquire desires or wants at will.

---

34 The example is from Epictetus, Diss. I.28.3, quoted in Burnyeat 1980, 40.
35 Prominent among those who hold that the constitutive aim of belief is truth are Williams 1973 and Velleman 2000b. (In order to distinguish belief from other propositional attitudes, such as imagining, supposing, or assuming, Velleman says that belief involves regarding a proposition as true with the aim of getting its truth-value right.) Williamson 2000, 238 f., argues that the norm of belief (or its expressed counterpart, assertion), is knowledge.
36 Epictetus makes the analogy explicit: just as every soul by nature (πένθους δὲ πάσα νοημ) assents to the true, so too every soul is impelled (ἀρχηγοίς κωπέοις) towards the good (Discourses 3.3.2-3); just as it is impossible to reject what appears to be true, so too it is impossible to turn away from what appears to be good (3.7.15). Thanks to Brad Inwood for bringing these passages to my attention.
37 There is a widespread contemporary opinion that cognition and conation differ profoundly in having different ‘directions of fit’, belief aiming to fit the world and desire to have the world fit it; the observations above are of ways in which desire has the same direction of fit as belief.
38 Millgram 1997, 11 f., argues that we cannot desire just anything, or desire at will, and that this is no contingent fact about desire. Desires involve inferential commitments: we appeal to them to figure out what to do, and so realizing that we acquired a so-called desire in a way that gives us no reason to think its object
In one of his *Cratylus* etymologies, Plato himself suggests that believing and wanting are alike in being goal-directed activities.\[39\]

[Socrates:] ‘Belief’ (Δοξα) is named after either the pursuit (δουξι) of the soul pursuing the knowledge of how things are, or the shooting (πολη) of a bow (τσοου) – but the latter seems more likely. At any rate, ‘opinion’ (διψησ) harmonizes with the latter. For this seems to be demonstrated by the movement (δον) of the soul towards all things, how each of the things that are is, just as indeed deliberation (ποληδ) is somehow towards shooting (πολη) [at something], and ‘wanting’ (ποληδθαι) and ‘deliberating’ (ποληδθαι) signify aiming (ποληθαι) [at something]. All these seem to follow belief (δεινη) insofar as they resemble shooting (ποληγ) [at something]. *Cratylus* 420bc\[39\]

Perhaps wanting aims at the good and so any occurrent want is for the apparent good just as belief aims at the true and so any particular belief is of that for which there is the most evidence. If the aim of belief is knowledge rather than simply the truth, then, by analogy, wanting would aim not just at the good, but at being in a certain relationship to it – a relationship that is, by analogy with knowledge, infallible and stable.

desirable, we will not base our inferences (or actions) on that so-called desire; it will fail to be a desire (it might still produce an urge which we might indulge, but in a cast of mind very different from that in which we act on a desire). However, whereas the difficulty of believing at will would seem to have to do with the constitutive aim of belief being truth, the difficulty of desiring at will identified by Millgram does not have to do with the constitutive aim of desire being for the good; instead, it comes down to the norm-governedness of belief: it is because believing that an object is desirable requires evidence of its desirability, and in this case one lacks such evidence (one’s desire having been acquired in such a way that it provides no evidence of desirability), that one is not willing to act or make inferences based upon one’s so-called desire. I’m grateful to Katy Abramson for discussion on this topic.

39 I take it that the etymologies of the *Cratylus* are supposed to reconstruct the beliefs of the name-givers, who supposedly based names on what they took to be the natures of the things they were naming; so, to the extent that the name-givers were right about the natures of the things they named, etymology is a route to the natures of things. Cf. Sedley 2003, ch. 2.

40 In the *Theaetetus*, Socrates proposes that in believing something, the soul is trying to line up what it perceives with some impression (knowledge) in the soul, or to catch one of the pieces of knowledge one possesses (193c). In these passages, the standard for belief is set by knowledge, rather than just by the truth. Perhaps the thought is that only an infallible power may be authoritative over another.
In the *Gorgias*, wanting seemed to be the infallible attitude; in the *Cratylus*, the suggestion is that it *aims at being* an infallible attitude. Without minimizing the difference between these positions, I want to note a similarity: in both cases, the thing grasped in cognition or conation is plausibly grasped *for the sake of* the thing aimed at. In the *Topics*, Aristotle notes that in determining whether a desire (or science or other term that has more than one object) is ‘of’ or ‘for’ something, we should bear in mind that the desire for something can be the desire for it as an end, as a means to that end (πρὸς τὸ τέλος) or accidentally (I.3, 111a1 f.). Later, having said that the object of desire is the pleasant (VI.3, 140b27), Aristotle follows his own advice: “[See] if what has been said about a relative (πρὸς τι) has not given its end (τέλος). The end in each thing is the best or that for the sake of which the other things [are]. Indeed the best or the last should be stated, for example, *desire is not for the pleasant but for pleasure: for it is for the sake of this too that we choose the pleasant (ταύτης γὰρ χάραι καὶ τὸ ἥδυ αἰρειμένας).’* (VI.8, 146b9-12, my emphasis).

On the other hand, the theory of recollection suggests a way to take literally the claims of *Gorgias* 466de and 474bc that people have the beliefs Socrates ascribes to them. According to the theory of recollection, people latently believe all truths (although they also believe various falsehoods). When we learn, we are recollecting, or making explicit, the truths we latently believe; this is how it is possible for us in

---

41 An obvious example of Socrates attributing beliefs to an interlocutor contrary to those professed by the interlocutor is found in Socrates’ questioning of the slave about the length on which an eight (square) foot square is based, at *Meno* 82d-85c.

42 In the *Gorgias*, Socrates suggests that the elenchus can show not only the inconsistency of an interlocutor’s beliefs but can also help him find the truth (473b, 482b). According to Vlastos 1994, Socrates’ practice of the elenchus assumes that whoever has a false moral belief also has true beliefs which can be shown to entail the negation of the false belief, so that when one’s beliefs are consistent, they are also true. However, in the *Symposium*, Diotima tells Socrates that he is one of those who say that love is not a god (my emphasis), and from this concludes that Socrates “says” love is not a god: Socrates says that gods are happy and beautiful; Socrates says that those who possess good and beautiful things are the happy; Socrates says love lacks good and beautiful things – so Socrates says that love is not a god (202cd) – this although Socrates has just been saying that love is a god (201e4-5). In this case, the attribution of beliefs (or really assertions) to Socrates seems to be based on entailment rather than recollection – unless people can vary in what latent beliefs they have (which might be consistent with the way recollection seems to work in the *Phaedrus*, according to which different disembodied souls see more and less of the Forms, cf. *Phaedrus* 248ae).
inquiry to aim at the truths we do not yet know and to recognize truths as such when we come upon them (Meno 80d-86b). So Socrates may attribute beliefs to his interlocutor on the ground that they are truths rather than on the ground that the interlocutor’s expressed beliefs logically commit him to the attributed beliefs. Note that in Gorgias 466de and 474bc, Polus has not yet said the things that will commit him to the views Socrates attributes to him when he attributes these views to him. If the interlocutor latently believes all truths then Socrates can seek to bring his true beliefs to the surface by showing that their negation drives the interlocutor into contradiction. If this is right, it might not always matter whose asserted beliefs figure in a refutation. In any case, the doctrine that we latently believe the truth would explain some of our cognitive behaviour: most notably, how, given that we engage in inquiry from a position of ignorance, it is possible for us to recognize the truth.

If we elaborate the cognition-conation parallel in the terms of the recollection account of cognition, then we may say that as people latently believe truths, so they latently want good things. (Since Socrates’ account of recollection doesn’t name a special faculty or attitude, in addition to belief, as the holder of our latently held true beliefs, I will not do so in the conative case, either.) This elaboration suggests a new way of understanding the passage from Plato’s Meno in which Socrates affirms that we desire good things and things we believe to be good, and both affirms and denies that when we are mistaken, we desire actually bad things. The passage begins with Meno’s attempt to say what virtue is.

[Meno:] Virtue seems to me, Socrates, to be as the poet says, ‘enjoying beautiful things and being powerful’. And I say that this is virtue: for the one who desires (ἐνισθέοντα) beautiful things to be capable of procuring them for himself.

[Socrates:] Do you mean that the one who desires beautiful things desires good things?

[M:] Exactly.

[S:] [Does it seem to you] then, that there are some who desire bad things, and others who desire good things? Doesn’t everybody, good man, seem to you to desire good things?

[M:] Not to me.

[S:] Rather, some [desire] bad things?

[M:] Yes.
[S:] Do you mean thinking that bad things are good, or even knowing that they are bad, they nevertheless desire them?
[M:] It seems to me that there are both kinds.
[S:] Does it seem to you, Meno, that anyone, when he knows that bad things are bad, nevertheless desires them?
[M:] Certainly.
[S:] What do you say he desires? Is it that they come to be his?
[M:] Yes, that they come to be his – what else?
[S:] Is it thinking that the bad things benefit him who possesses them, or knowing that the bad things harm their possessor?
[M:] There are some who think that the bad things benefit, and others who know that they harm.
[S:] And does it seem to you that those who think that the bad things benefit know that they are bad?
[M:] No, I don’t think that at all.
[S:] Then it is clear (1) that, on the one hand, they (δῆλον ἃν οἶνος) do not desire bad things (ὦ τῶν κακῶν ἐπιθυμόντων), who are ignorant about them (that they are bad), (2) but [rather, they desire] those things which they think are good (ἰκαίων ἃ ἱώντα ἄγαθα ἔχοντα); (3) and that, on the other hand, these things are in fact bad (ἴσταν ὧν ταύτα γινόμενα). (4) The result (ἂστε) is clearly that those who are ignorant about these things and think they are good desire good things. No? (77ε3)
[M:] They, at any rate, probably do.
[S:] Well then, do those who, as you say, desire bad things, thinking that bad things harm their possessor, know that they will be harmed by them?
[M:] They must.
[S:] But do they not think that those who are harmed are wretched insofar as they are harmed?
[M:] This too, is necessary.
[S:] And [do they not think] that the wretched are unhappy?
[M:] I think so, anyway.
[S:] Is there anyone who wants (βοιλέται) to be wretched and unhappy? (78α4)
[M:] I don’t think so, Socrates.
[S:] Then, Meno, no one wants (βοιλέται) bad things, unless he wants (βοιλέται) to be such. For what else is it to be wretched than to desire (ἐπιθυμέων) bad things and to get them?
[M:] You are probably saying what’s true, Socrates, and no-one wants (βοιλέται) bad things.
[S:] Well did you say just now that virtue is wanting (βοιλεθαι) good things and having the power [to get them]?
[M:] I did say that.
[S:] Well this having been said, wanting (τὸ μὲν βοιλεθαι) belongs to all, and in this way no-one is better than anyone else?
[M:] It appears so.
[S:] But it is clear that if one person is better than another, it would be in respect of having the power [to get good things]. (77b2-78b8)

We can describe the relationship between this passage and the passage from the Gorgias with which we began, as follows: while the Gorgias passage insisted that we could only want (πολιτεία) actually good things and remained silent about our conative attitude towards things that we only believe are good (save to say that it is not wanting), the Meno passage both names and explains the logic of this second conative attitude, desire (πιθανόν).

To understand this logic, we should first note that when Meno initially defines virtue as desiring good things and having the power to get them, Socrates asks him, ‘doesn’t everybody desire good things?’ – and not, ‘isn’t all desire for good things?’ Meno’s mistake is not, in the first instance, to suppose that some desires are for things that are actually bad: Socrates denies not this, but that the people who desire actually bad things thinking they are good desire the actually bad things, tout court. So Meno’s first mistake is to suppose that there are people who don’t desire actually good things. The cognitive parallel would be to

43 The difficulty with drawing a distinction between wanting and desiring in this passage is that while Meno initially speaks of virtue in terms of desiring good things (77b4), he is described by Socrates as saying that virtue involves wanting good things and seems to see no difference between these two formulations (78b3-4).

44 George Rudebusch has objected that on my interpretation, at 78a (“what else is being wretched than desiring bad things [...]?”) Socrates is affirming the truth of the part of Meno’s definition of virtue (“desiring good things”) he criticised at 77bc. My response is that Socrates’ criticism need not be of Meno’s words, but instead of the assumption underlying those words, that there are people (namely, vicious ones) who do not desire good things. It may also be significant that Socrates restates Meno’s definition of virtue as “wanting (πολιτεία) good things and having the power [to get them]”, and then says by way of correction, “wanting belongs to all, and in this way no one is better than anyone else [...]” [i]f one person is better than another, it would be in respect of having the power [to get good things]” (78b). Perhaps Socrates distinguishes between ‘desire’ and ‘wanting’ as follows: wanting belongs to all and so does not distinguish the virtuous from the vicious, but desire does distinguish them because it is dependent on beliefs about goodness which may be true or false. However, at 77de, Socrates seems to say that even people who desire actually bad things desire good (and not only apparently good) things – for the argument, see above.
suppose that there are people who don’t (even latently) believe true things, who are totally cut off from the truth. But according to Socrates, everyone, virtuous or vicious, desires good things.

Interpretations of this text have focussed on the passage at 77de, where Socrates reasons that if people desire things that are actually bad thinking they are good, then they (‘clearly!’) desire good things. Not only is the basis for this inference unclear; it looks as if in the process of making it, Socrates both affirms and denies that ignorant people desire bad things.

In a classic paper, Gerasimos Santas has argued that since the object of desire is always desired under a description, Socrates is simply insisting at 77de that this description must always be ‘good’, even if the objects one desires as good, i.e., thinking they are good, may be actually bad. So, Santas proposes, Socrates is substituting equivalent expressions, ‘[apparently] good things’ for ‘things believed to be good’, in order to make the following inference.

Some people desire things they believe to be good [= (2) at 77de].
They desire good things [= (4) at 77de].
(stated more carefully:)
They desire apparently good things [= (4) at 77de].

This is a possible reading of 77de, but it fits uneasily with some aspects of the passage. First, given that Socrates has just distinguished between the really good things and the apparently good things we may desire,

---

45 Santas distinguishes desiring (ἐπιθυμεῖν) from wanting (βούλεσθαι) in this passage (and elsewhere), saying that while the intended objects of ἐπιθυμεῖν must always be good, the actual objects may be good or bad, but the actual and intended objects of βούλεσθαι must be good. Now on this view, the bad things that no-one desires have to be apparently bad things; but Santas says, “[…] in the present argument […] [Plato] uses βούλεσθαι to deny Meno’s claim that there are people who desire bad things knowing that they are bad, and ἐπιθυμεῖν to deny the claim that there are people who desire what are in fact bad things not knowing that they are bad; his using ἐπιθυμεῖν in the latter case allows us to interpret the denial as not including a denial that the actual objects of these people’s desires are bad things.” (1964, 152 n. 15, my emphases). Santas does not explain why Socrates is not contradicting himself if he is (1) denying that there are people who desire things that are actually bad, and, at the same time, (2) restricting this denial so as to allow that the actual objects of these people’s desires are bad.

46 Santas 1964, 149 f.
one might have expected a shorthand that does not confuse this very
distinction (as does ‘good things’ for ‘things believed to be good’).
Second, if the inference consists in a substitution of semantically
equivalent terms, then Socrates’ emphatic ‘therefore’ and ‘clearly’ are
overly dramatic, even misleading. It may be that the excitement is not
about the substitution but about the further conclusion that desires for
bad things, and ultimately vice, are due to ignorance. Still, a more
substantive inference, if it was also plausible, would provide a better fit
with these signposts.

Terry Penner and Christopher Rowe argue that contrary to appearances, 77de
does not allow that we desire things that are actually bad; rather, when we seem to
desire bad things, what we desire are the benefits which we wrongly believe result
from possessing the bad things in question. Penner and Rowe describe the Meno
passage as a whole as posing the question, ‘what could the object of desire be?’ and
then arguing, by elimination, that it is always the good. On their interpretation, at
77d7-e3 Socrates says,

(1*) these people [whatever we may go on to say about the others: μέν] don’t
desire τὰ κακὰ, the people who don’t know them [i.e. that they are κακὰ].
Instead [ἀλλὰ: strongly adversative],
(2*) they desire those things which [we agree] they think good (ἐπιθυμοῦσιν …
ἐκείνων ἀ κακωτά ἀγαθά).
But
(3*) these very things [ταῦτα γε] in fact are [position of ἐστιν] κακὰ.
So, then,
(4*) [this first group,] those who don’t know them [τὰ κακὰ], and think that they
are ἀγαθά clearly desire τὰ ἀγαθά.

According to Penner and Rowe, in (1*) Socrates asserts that the object of desire is
not (ever) the bad; in (2*) and (3*), that it is not the apparent good; and so, what is
left is (4*), the real good. So in (2*) and (3*) (which correspond to 77e1-2), Socrates
is not saying, ‘they desire those things that they believe to be good but that are in
fact bad’; rather, he is saying, in Penner and Rowe’s paraphrase, ‘You might think
that the object of desire is, not ‘bad things’, but ‘things thought good’. But those
very things in fact are bad things. [So the new suggestion, that the object of desire is

47 Penner / Rowe 1994. Mariana Anagnostopoulou has an excellent discussion of
this paper in Anagnostopoulou 2003.
48 Penner / Rowe 1994, 18 f. (remarks in square brackets by Penner and Rowe).
things thought good doesn’t in fact get us anywhere, given that in ... [(1)] we have ruled out the claim that what we desire are (really) bad things.\[4\]

Penner and Rowe’s reading of 77d7-e3 makes nonsense of the *Meno* passage as a whole. How can 77de deny the very possibility of desiring things that are in fact bad, when, at 78a7-8, Socrates says that wretchedness is desiring (ἐπιθυμεῖν) bad things and getting them – unless no-one can be wretched? Surely, the overall point of the *Meno* passage is that virtue is not a matter of desiring or wanting good things – everyone does that – but rather, it is a matter of being able to get good things. Vice is then a matter of desiring actually bad things in addition to the actually good things everyone desires, and because when people desire actually bad things, they do so thinking these bad things good, vice is ignorance. Finally, wretchedness is getting the mistakenly desired actually bad things.

I propose that at 77de, Socrates is correcting Meno as follows: we should say, of people who are ignorant of the badness of the things they desire, not (simply) that they desire bad things, for this is only half of the truth and is misleading if we are trying to understand their desires. Instead, we should say that they desire bad things believing them to be good things. Then, from their desiring things they believe to be good, we may conclude that they (too) must want (actually) good things. Socrates is not simply recording the point that if we desire something it must appear good to us but adding that our attraction to apparent goods shows that we have some kind of orientation towards what is actually good for us. But is this inference plausible?

We would have one reason to suppose that (some of) our conative attitudes aim at the good (or at being in a certain relationship to the good) if this could explain our conative behaviour in the way that belief’s aiming at the true (or at knowledge) explains our cognitive behaviour.

The explananda Plato provides are both cognitive and conative. These explananda are made most explicit in a passage of the *Republic*.

---

49 Penner / Rowe 1994, 20. Penner and Rowe argue for a full stop between (2*) and (3*) on the grounds that if they were to be taken as a unit, (3*) should have ὅταν rather than ὅτως and that (3*) is no mere parenthetical addition to (2*) but is quite emphatic. But if ἀλλὰ is “strongly adversative”, as they claim, then it ought to be governed by δῆλον ὅταν, which they seem to deny. In any case, whatever one thinks of it, Penner and Rowe’s punctuation does not make 77e1 (Penner and Rowe’s (2*)) anything other than an assertion – there is no textual justification for their treating it as a “suggestion” which “doesn’t get us anywhere”. I am grateful to Clerk Shaw for discussion on this passage.
Socrates has just said that the education of philosopher-rulers culminates in learning about the Form of the Good, which is the cause of the beneficialness of other things, and knowledge of which is necessary for us to benefit from any other knowledge we have (505a), and then gone on to report some controversies about what the good is (is it pleasure? but there are bad pleasures, too. Is it knowledge? but knowledge of what? [505bc]). He now remarks:


[... ] in the case of just and fine things, many people would choose (ελεύθερο) things that seem to be [just and fine] (τὰ δὲκοιμότα), even if they are not; they [would choose] also to seem (δόκοντο) to do and to acquire these things. But in the case of good things, to acquire appearances (τὰ δικαίωσε) satisfies (δοκεῖ) no-one, but everyone seeks the real thing (τὰ δικαίωσι) and disdains mere belief (τὴν δοξὴ) [...]. The good is what every soul pursues (δοκεῖ), and for the sake of which it does everything, divining (αποδικεῖομαι) that it is something but

---

50 The relationship between the good that is the object of our conation and the Form of the Good is so big a topic that I can here only give preliminary sketch of the issues as I understand them. I take it that the good that we all desire is our own happiness, and that the Republic tells us that Form of the Good is (1) that which makes things like justice, which are valuable in some other way (perhaps to society, or to the cosmos), good for us, and (2) something we must know, if we are to be happy in the fullest sense. The Republic seems to suggest that we would want to know the Form of the Good as a means to being happy, or we see the relationship between knowing the Form of the Good and being happy; rather than that we directly want (to know) the Form of the Good. But my proposed elaboration of the parallel between conation and cognition in terms of recollection might suggest that as we have some direct cognitive contact with the Forms, so we have some direct conative contact with the Form of the Good. To say that the Form of the Good is the true object of our conation, I suppose we could draw on the Symposium’s treatment of love of beauty, but I don’t yet see how to deal with the disanalogies between the good and the beautiful.

51 Alternatively, ‘to do and acquire and believe these things’ (e.g. Grube / Reeve 1992). For discussion of these alternative translations, see Segvic 2000.

52 Πάντα πράτευν can have the sense ‘exert great effort’, which may be favoured by those who believe that among the nonrational motivations in Republic IV are good-independent conations. The present passage seems to be making a weightier claim than that getting good things is a very important goal for us, and since I understand the nonrational motivations as operating with conceptions of the good (albeit weak and faulty ones, see above, n. 2), I can accept that every part of the soul pursues the good and every action is performed for the sake of the good. Plato’s own uses of the expression at Charmides 159be and Gorgias 469c7 seem to require the sense ‘perform every action’; this is also so for Gorgias’ and Aristotle’s uses of the expression.
confused and unable to grasp adequately what it is or to attain any stable belief (μωμύμο πίστει) such as it has about other things [...]. (505d5-e4)

Republic 505de contrasts a common attitude towards justice and fineness with what it claims is everyone’s attitude towards good things or the good: everyone ‘divines’ that the good is ‘something’, that is, something objective; everyone seeks the good; no-one is satisfied with things that appear good. By contrast, many people are content to pursue what is conventionally regarded as just and fine, either in the

---

53 Republic 505de claims that the condition of the soul seeking the good is one of ‘divining’ (ἀπομακρυνθείνη) that the good is something, by which I understand, something objective and not merely conventional. Plotinus uses this same word to express the attitude we all have to the good even when we are unaware of it, and contrasts this attitude with our attitude towards the beautiful, desire for which requires that we know the beautiful: in the case of the good, what we have is a connatural desire (σύγκλονος φιλίας, Enneads V.5.12). Elsewhere in Plato, the word, like the associated μακρύνωσθαι, seems to refer to trustworthy but ungrounded hunches. For parallels: at Republic 506a6-7, Socrates says “I μακρύνωσμαι that no-one can sufficiently know just and fine things before knowing the good”; at Republic 516d, he says that the person who has seen the real natural world will not envy but pity the person who is quickest to make out the shadows and most successful in guessing (διονυσίωσα ἀπομακρυνθεῖνη) what is about to come; at Sophist 250c Theaetetus says we ἀπομακρύνωσθαι that Being is a third thing aside from Motion and Rest. In conjunction with the good, at Philebus 64a Socrates says that if one aims at the best mixture of reason and pleasure, and wants to discover the good in humans and in the universe, and to divine – μακρύνωσθαι – the nature of the good itself, one had better not associate reason with foolish pleasures; and Aristotle, at EN 1095b26 says that the good is something we intuitively believe – μακρύνωσθαι – to be our own and hard to take away.) Cf. also EN X.2 1173a4. At Phaedrus 244cd, Socrates characterises μακρύνη (prophecy) as a god-given power of reason. The earliest use is at Apology 21a4-5, where Chaerephon is reported asking the oracle to μακρύνωσθαι whether there is anyone wiser than Socrates. Throughout, the suggestion is of unearned epistemic success.

54 Stephen Menn has objected that the contrast between our attitudes towards the just and fine on the one hand and the good on the other may be that in case we are amoralists, we are content merely to appear just and fine to others, but in no case are we content merely to appear good (happy) to others; I am taking it that the contrast is between our being content merely to appear just and fine, whether to others or ourselves, and our discontent with merely appearing good, even to ourselves. One advantage of my reading is that it connects the thoughts of this passage, so that there is not a radical change of subject between everyone’s pursuit of (what appear to them to be) good things and the good which is the dimly-grasped object of every soul’s pursuit.
spirit of the amoralist who seeks to appear just to others (365bc), or because they believe that things are just and fine only by convention.

The special attitude we have toward the good appears in a number of other contexts in Plato: for example, Thrasymanus, who says justice is defined by the rulers’ laws, which the rulers frame to their own advantage, rejects Cleitophon’s suggestion that the rulers’ advantage, or good, is determined by the rulers’ say-so (Republic 340bc); Protagoras and his followers exempt from their thoroughgoing relativism (about justice, the properties of bodies, the authority of waking over dreaming, etc.) questions of what is better or worse, or advantageous, or useful (Theaetetus 167ad, 171e-72b, 177d-178a).

Republic 505de may overstate the difference between our attitudes toward the good and everything else: after all, the material possessions and cultural accomplishments people pursue have a lot to do with their societies’ assessments of what things are good, and people (including the people of Plato) do not as a matter of course question these assessments or subject them to wholesale scrutiny – as one might expect them to if they were determined on getting the real thing. But a slightly weaker claim is surely true and worth remarking: once we have come to question or have doubts about whether what we have considered good and pursued as good really is good, that question takes hold of us in a way that the question of what is really just, or really fine, need not. Gripped by the question, we may actively seek to find out what is truly good, or be prepared to accept that someone else might have a better grasp of the good than do we, or, at the very least, be disquieted by our condition of not knowing what is good. And that is what we should expect, if we desire things that appear good to us – the explanation for our desiring them is that they appear good to us, and the explanation for our desiring things which appear good to us is that we want what is truly good, and our appearances of goodness (especially when we seek to align these with reality by inquiry) are our route to what is really good for us.

Republic 505de also says that the soul will not reach a stable condition until it finds the real thing: no-one is satisfied with mere appearances of goodness, only really good things satisfy. So the passage is telling us that even if we succeed in acquiring things that appear to be good, if they are not really good, we will not be stably fulfilled by them. The object of wanting, then, is identified not only by what we strive
after, but also by what stably stills our striving, or by that in which the striving comes to rest.\textsuperscript{3}

Stable fulfillment is not the same as purely subjective satisfaction, for apparent goods might provide the latter as well. Plato acknowledges that people may be so ignorant that they are content with their situation (cf. \textit{Symposium} 204a). Now Plato could offer an objective account of satisfaction and say that thinking one is satisfied is not the same as being satisfied, just as thinking one wants something is not the same as wanting it, and thinking one knows something is not the same as knowing it. But the \textit{Republic} passage points us towards felt dissatisfactions: the instability of the soul in possession of only apparent goods; our disdain for mere belief about the good (so that if a real question arises as to whether what we have or are pursuing is good, it can take hold of us).\textsuperscript{4}

There is a now-familiar slogan coined by G.E.M. Anscombe, that “the primitive sign of wanting is trying to get”.\textsuperscript{5} To ‘trying to get’, then, \textit{Republic} 505\textsuperscript{d}e adds another (and often neglected) sign of wanting (or

\textsuperscript{3} Plotinus speaks of the coming to rest of our desire for the good at \textit{Enneads} I.4.4, I.4.6. That in which our striving comes to rest need not only be truth, but (as Chris Bobonich has pointed out to me) also pure and true pleasures, such as the pleasures of contemplation (\textit{Republic} 583b\textsubscript{f}, cf. \textit{Gorgias} 493e-94a).

\textsuperscript{4} Dan Devereux, and on another occasion Houston Smit, have objected that what explains the good’s being the object of our conation need not be anything psychological, but could be a fact about the relationship between our words or concepts and the world: if I want the best bagel in New York City, it is the description ‘the best bagel in New York City’ that needs to be satisfied by an object in the world, not anything in my soul. However, \textit{Republic} 505\textsuperscript{d}e’s discussion of the just and the fine shows that on Plato’s view, whether my wanting to do the just thing is determined by the truth about justice or by what relevant others think about justice depends on whether I am a conventionalist (or amoralist) or a realist about justice. So according to Plato at least, our conation’s being for the real good is due not to a general fact about language but to some special fact about conation.

\textsuperscript{5} Anscombe 1957, 67 (§36). Less often quoted, but an important supplement to ‘trying to get’: “Now saying ‘I want’ is often a way to be given something; so when out of the blue someone says ‘I want a pin’ and denies wanting it for anything, let us suppose we give it him and see what he does with it. He takes it, let us say, he smiles and says ‘Thank you. My want is gratified’ – but what does he do with the pin? If he puts it down and forgets about it, in what sense was it true to say that he wanted a pin? He used these words, the effect of which was that he was given one; but what reason have we to say he wanted a pin rather than: to see if we would take the trouble to give him one?” (ibid., 70 §37).
having wanted) something, namely, being fulfilled by that thing. We attribute thirst, or a conation for drinking, to a crying infant not because her crying is trying to get drink – the thirsty infant cries long before she can understand that crying causes her caretaker to give her drink; presumably she cries when thirsty by instinct. We attribute a conation for drinking to the infant because drink (and only drink, not being held or sung to) is what stops her crying for more than a moment. It seems undeniable that we have other bases for attributing conative attitudes to the infant than what she tries to get: for example, on the basis of the (relatively dehydrated) condition of her body, or on the basis of expert advice about how often and how much infants of that size and age should drink in the normal course of their development. Perhaps Anscombe would attribute ‘trying to get’ to a creature on the basis of such considerations; my point is simply to note that it is not only Plato, but we too, who attribute conative attitudes on the basis of other considerations than what the subject reports are his attitudes.

4. The Good and the Apparent Good in the Explanation of Action

I have been arguing that Socrates sometimes attributes conative attitudes to a subject on the basis of what is actually good for that subject, and that the theoretical justification for this attribution is that it best explains certain features of our conative behaviour – not only the way we represent to ourselves the aim of an individual action, but also what leads to the cessation of aim-oriented activity, the bases on which we revise our aims, and the fact that we engage in inquiry about what to aim at.

In the paper discussed in section 2 above, McTighe suggests that in Socrates’ statements that we do what we do for the sake of the good, the for-the-sake-of relation (ἐν ἰδίᾳ) names the ‘teleological lure’ that the good exerts on us. Like everything in the cosmos, our actions are for the sake of the good. However, McTighe contrasts this with the claim that the good is the object of our desire (or want; McTighe does not distinguish between the two), for McTighe thinks that

58 Cf. Reshotko 1996 for a defence of desire-attribution on the basis of termination of conative behaviour.
according to Plato, it is the apparent good that we all desire. In reply to McTighe, I want to suggest some reasons why Plato would want to draw conative attitudes into the sphere over which the good exerts its teleological lure.

Doing what we judge best is probably our best shot at getting what is actually good for us (cf. ‘it’s pursuing the good that we walk whenever we walk, thinking it to be better’ [Gorgias 468b1-2]), and so it is plausible to think that we do the former for the sake of the latter. Further, we exercise judgment, i.e., try to determine what will be best for us because we believe that this (rather than, e.g., tossing a coin) is the way to get what is actually good for us. It seems that if we did not want what was actually good (but were content with the apparently good), we would not exert the intellectual effort we do exert to determine what is actually good, reported in Republic 505de.

Still, the view that belief and wanting aim at (being in a certain relationship to) the truth and the good, or that we have latent beliefs and wants, and that one of our latent wants is for the good, have an occult air to them. We might prefer to dismiss the putative phenomena of Republic 505de and say with Aristotle that we want what appears good to us, and what is truly good appears good to the good person.

Aristotle’s solution raises a worry about how or why what the good person wants, or considers good, should motivate those who are not in his condition. Aristotle might be thought to be reasoning that since everyone wants to be happy, and happiness lies in good functioning and the good person functions well, everyone ought to follow the good person. There is a complication, however: Aristotle says that while everyone agrees that happiness is the good, this is an agreement about a word (EN 1095b17, ἴδεω), for people have different conceptions of happiness, some thinking that it is pleasure, others honour, others contemplation, and so on (EN I.4-5). In this case, won’t these people

---

60 Contrary to McTighe’s assertion that apart from Gorgias 467c-68d Plato never says that we want the real and not the apparent good, Laws 734b1-3 (where the good is identified as pleasure) says that it is not possible for the person who wants to live a pleasant life (τὸν γε βουλεύον ζῆν) to voluntarily (ἐκόψον) live a licentious life; this seems to require that the person want to live the life that is really pleasant, not merely the life that appears to him pleasant.

61 In Lear 2004 Gabriel Richardson Lear argues that in both Plato and Aristotle, the ‘for the sake of’ relation to the end includes both the relation of means to end and of approximation to unattainable ideal.
want not functioning well, but rather living pleasantly, or honourably, or whatever, with little regard to what the good person wants? Or would Aristotle follow Plato and say that in the case of happiness, people want the real thing? It seems that unless he follows Plato, Aristotle must conclude that while people ought to want functioning well, this ought is external to them.

Following Aristotle, Anscombe says that the connexion that exists between ‘wanting’ ... and ‘good’ is a conceptual connexion between wanting something and seeing that thing as desirable. According to Anscombe, one can fairly ask, about the (purported) object of anyone’s desire, ‘what do you want that for?’ until the subject has given a ‘desirability characterisation’ of the object in question. Once he has done this, it no longer makes any sense to ask, ‘what do you want that for?’ As Anscombe puts it,

the notion of ‘good’ that has to be introduced in an account of wanting is not that of what is really good but of what the agent conceives to be good; what the agent wants would have to be characterisable as good by him […]. Whereas when we are explaining truth as a predicate of judgments, propositions, or thoughts, we have to speak of a relation to what is really so, not just of what seems so to the judging mind.

Part of the appeal of saying that desire is for the apparent good (and saying that this is a conceptual connexion) is that this holds out the promise of making people’s actions and desires intelligible to one another. So, Anscombe says, even Satan’s ‘Evil be thou my good’ can be rendered intelligible by the explanation that good is impotent, slavish and inglorious. Appeal to an apparent good explains action from the

---

62 Anscombe 1957, 75 f. (§40).
63 Ibid., 69 f. (§37).
64 Ibid., 76 (§40).
65 Ibid., 74 f. (§39). According to Velleman, Anscombe is mischaracterizing Satan’s desire, which is meant to be perverse, a desire for bad things under the description ‘bad’. Velleman criticizes philosophers like Anscombe for confusing two senses of ‘good’ when they say ‘desire is for the good’: sometimes ‘good’ is nothing more than direction of fit, so that to say ‘desire is for the good’ is to say no more than that desire is for that which is to be brought about, but at other times, ‘good’ is treated as a substantial feature of the object of desire (e.g. its attractiveness, where the attractiveness might be a justification for desire and not
point of view of the agent and engages the explainer’s sympathy with that point of view (the explainer comes to see how it would be reasonable, in those circumstances, to φ, or how he too would φ, in those circumstances).

But Plato’s Socrates holds that explanation ought to show how the explanandum is as it is for the best (Phaedo 97c f.). If one takes the perspective of a student of nature on a creature of a certain kind, one will try to explain its capacities and characteristic activities in terms of the characteristic good of creatures of that kind. From this perspective, an animal which is able to represent things as good and go after what appears good to it is thereby equipped to get what actually is good for it — better equipped, for instance, than is a plant, which, lacking the capacity to represent things to itself and modify its attitude accordingly, cannot respond as flexibly to its environment. But from the fact that the plant does not have a representational capacity it does not follow that it has no orientation towards its own good condition: The student of nature may say that the plant wants water and sunlight just on the grounds that these are conductive to its being in a good condition. And in this case, it is quite reasonable for him to attribute wants on similar grounds to creatures that do have a representational capacity, and to say that such animals go after what appears good to them for the sake of what is good for them, or because they want what is good for them: It is for the sake of their good that the designer gave these animals the capacity to represent things as good and go after them. From this perspective, furthermore, it makes sense to think of a creature’s actions on the basis of what appears good to it as means: these are a monk the designer’s or nature’s means for getting it what’s really good for it. But in cases in which what appears good to it is really bad, the mistaken appearances are not means for getting what’s really bad for it — its desires and desires directed pursuit are not for the sake of its getting what’s bad for it.

merely another way of saying that the object is desired). See Velleman 2000a, 105 f., 118 f.

66 McTighe 1984, 214, rightly observes, “The fact that all men desire and seek the apparent good […] is […] explained by the [teleological] relationship which the good has to the soul. We might say that the good brings it about that men, qua having a soul, necessarily seek what they judge to be good.”

67 A parallel from Fodor 1987 may be useful here (Fodor acknowledges that the original insight is Plato’s). Fodor accommodates misrepresentation in the framework of a causal theory of reference by saying that misrepresentation is
From this perspective, then, it makes sense to deny that something really bad is what the creature wants, however the creature might experience it.\footnote{This paper has been long in the works and I have incurred many debts along the way. I am grateful to the University of Michigan for the assistant professor nurturance leave during which the paper was actually written. I would also like to thank, for their comments on previous (and in some cases very different) versions of the paper, Wolfgang Mann and another reader for Archiv für Geschichte der Philosophie, Katy Abramson, Chris Bobonich, Tad Brennan, Sarah Broadie, Dan Devereux, Alice van Harten, Richard Holton, Rae Langton, Stephen Menn, Yaseen Noorani, Terry Penner, Sara Rappe, Naomi Reshotko, George Rudebusch, Gerasimos Santas, Rachel Singpurwalla, David Velleman, and Roslyn Weiss, as well as audiences at the University of California at Berkeley, McGill University, University of Edinburgh, University of Arizona, Ohio State University, Indiana University, Stanford University and the University of North Carolina Chapel Hill Philosophy Colloquium.}


