

Articles

**Reasoning with the Irrational:
Moral Psychology in the *Protagoras***

Rachel Singpurwalla

We all, on occasion, experience motivational conflict and weakness of the will. When we diet, we desire to eat sweets, even while striving to check our impulses. Sometimes we succeed, and sometimes we succumb. Other times, we make decisions to improve our work habits, but then permit ourselves to be distracted by passing pleasures. Later, if we have allowed ourselves to indulge, we rebuke ourselves and feel remorse or regret. We describe ourselves as having been weak-willed, as having been overcome by our desires. What conception of the mental states that are capable of playing a role in motivating our actions explains the fact that we can have these sorts of experiences? In particular, do we need to posit non-rational or irrational desires—motivational states that arise independently of and so can conflict with our reasoned conception of the good—to explain the phenomena?¹ And if so, exactly what kind of mental state is an irrational desire?

It is widely held by commentators that in the *Protagoras*, Socrates attempts to give an account of the phenomena of psychic conflict and weakness of the will without positing irrational desires (see, e.g., Frede 1992, xxix-xxx; Irwin 1977, 78, and 1995, 209; Penner 1990, 1996, 1997; Reshotko 1992). This interpretation of Socrates' position is largely due to the fact that he denies that we ever pursue actions that we either know or believe to be bad; instead, we always pursue what we think is best. On this account, if we do take an action that is in fact bad for us, it must be due to ignorance (358c-d). Penner 1990, 68, for example, infers from this Socratic position that 'there are no desires that are not desires for what is the best option available in the circumstances'. And Irwin 1977, 78 states that according to Socrates, 'if A wants x, he wants it for its contribution to some good y, and ultimately to the final good, and if he ceases to believe that x contributes to the final good, he will cease to want x'.

These commentators argue, then, that Socrates denies that we ever pursue actions that we think are bad *because* he denies the existence of irrational desires. Given this explanation of Socrates' position it is not surprising that many philosophers have found his conception of our moral psychology excessively

¹ I use the terms non-rational and irrational desire somewhat interchangeably throughout the article, but I conceive of the distinction between them as follows: a non-rational desire is a desire that arises independently of reasoning and so has the potential to come into conflict with our reasoned conception of the good; an irrational desire is one that arises independently of our reason *and* that we consciously recognize as conflicting with our reasoned conception of the good.

rational and thus implausible. Indeed, many commentators note approvingly that the moral psychology put forth in the *Republic*—where Plato explicitly appeals to rational and irrational desires to explain mental conflict and weakness of the will—is an explicit rejection of the Socratic picture. Frede 1992, xxx states: ‘if we find this highly intellectualist account...implausible, we should keep in mind that it is only Plato, in the *Republic* (IV, 437bff.), who, precisely to explain how one can act against the judgment of one’s reason, for the first time introduces different parts of the soul, each with its own desires, allowing us to understand how irrational desire may overcome the dictates of desire and reason’.²

I aim to challenge this commonly held line of thought: I argue that Socrates accepts the existence of distinct sources of motivation, and in particular, irrational desires, and uses them to explain mental conflict and weakness of the will in the *Protagoras*. The resulting picture is both psychologically plausible and philosophically distinctive. Recently, Devereux 1995, and following him, Brickhouse and Smith 2000, 179-181 and 2002 and 2005, have also argued that Socrates posits the existence of irrational desires in the *Protagoras*. These commentators have not, however, given a clear and explicit account of the nature of an irrational desire, but have implicitly assumed a very standard conception, namely, that an irrational desire is simply a strong feeling of attraction or aversion to a certain object or state of affairs. But in adopting this conception of an irrational desire these commentators have failed to explain certain interesting aspects of Socrates’ theory of our moral psychology. In particular, they have failed to explain sufficiently (i) how Socrates can both deny that we ever act against our belief about what is best and accept the existence of irrational desires, and (ii) why Socrates thinks that knowledge quells the efficacy of irrational desires. I argue instead that Socrates holds that an irrational desire is a certain kind of evaluation. This alternative conception of an irrational desire opens up the interesting possibility that irrational desires, while resistant to reason, are capable of both affecting and being affected by our reasoned judgments. Socrates holds the distinctive view, then, that irrational desires are not *in principle* immune to rational revision. I conclude by briefly suggesting that the *Republic* also defends the view that an irrational desire is an evaluation that is resistant yet amenable to rational considerations. And I then suggest that the fact that irrational desires in both the *Protagoras* and the *Republic* are characterized as mental states that are resistant but not immune to rational revision suggests that we should understand the division of the soul in both dialogues as a division of reason itself.³

² Irwin 1995, 209; Penner 1971, 1990; and Reeve 1988, 134-135 also argue that the moral psychology of the *Republic* differs from that of the *Protagoras* in so far as the *Republic* acknowledges the existence of irrational desires.

³ My aim is to argue that the character ‘Socrates’ countenanced irrational desires in the *Protagoras*. I do not address here the important but separate question of whether or not the historical Socrates countenanced irrational desires. Thus, I do not consider other sources of evidence for the philosophy of the historical Socrates. I do not, e.g., consider the evidence in Xenophon’s *Memorabilia* iv 5 that

I. Mental Conflict and Weakness of the Will in the *Protagoras*

In the *Protagoras*, Socrates defends the striking claim that knowledge (ἐπιστήμη) is sufficient for right action (352c1-7). That is, Socrates thinks that if we know what is good, we will do what is good; there is nothing that can prevent us from acting on our knowledge of the good. The phenomenon of weakness of the will poses a *prima facie* challenge to this claim, for weakness of the will is the experience of failing to do what we know to be best (352d1-7). Moreover, the common explanation for this is that we are overcome (ἡττωμένους) by pleasure or pain or are being ruled (κρατουμένους) by any number of other passions (352d7-e1). All of this suggests that there is a source of motivation that is independent of our reason and that can move us to act contrary to our knowledge of the good.

To defend his claim that knowledge is sufficient for virtue, Socrates provides a different account of mental conflict and weakness of the will. Socrates claims that no one who either knows (εἰδώς) or believes (οἴομενος) that one action is better (βελτίω) than another will choose the worse action. Thus, weakness of will (ἡττω αὐτοῦ) is nothing other than ignorance (ἀμαθία), and self-control (κρείττω ἑαυτοῦ) is nothing other than wisdom (σοφία, 358b6-c1). In order, however, to provide an account that respects the undeniable fact that it surely looks as if people often fail to do what they believe to be best, Socrates must concede that at least prior and posterior to the action, the person does recognize which course of action is truly best. Thus, the ignorance must be a passing condition of the person.⁴ And if this is the case, then Socrates owes an explanation for why the person becomes ignorant about the right course at the time of action.

Socrates argues that we are prone to make mistakes about which course of action is most valuable (which, for the purposes of this argument, he equates with the most pleasant, 353c-355c),⁵ because of the fact that things can appear differently to us in relation to the varying temporal perspectives we can have of them. At 356c5-8 he says:

Do the same magnitudes appear (φαίνεται) to our sight larger when close at hand, smaller when farther away, or not? ...And similarly for thicknesses and pluralities? And equal sounds seem greater when close at hand, and lesser when farther away?⁶

suggests that Socrates countenanced irrational desires. Nor do I canvas Plato's so-called 'early dialogues', which are often thought to represent the views of the historical Socrates, for evidence that might suggest that the historical Socrates countenanced irrational desires, e.g., *Gorgias* 491d-494b.

⁴ In the *Nicomachean Ethics*, Aristotle also argues that if Socrates' account is to reflect the manifest evidence, namely, that the person is not ignorant before the time of action, then he must assume that the ignorance is a passing condition of the person (1145b28-32). Devereux 1995, 389-396 also makes this argument.

⁵ I do not here take a position on whether or not Socrates is actually committed to the hedonism he puts forward in the *Protagoras*. See Santas 1966 for a defense of the view that Socrates' argument against weakness of the will need not rely on hedonism.

⁶ Translations are my own unless otherwise noted.

Socrates thinks that just as things that are far away appear smaller than they truly are, pleasures or pains that are in the distant future appear worth less than they really are; and just as objects that are closer to us appear larger than they truly are, pleasures or pains that are immediate appear to be worth more than they truly are. In short, the temporal proximity of pleasures or pains distorts our estimation of their worth.

In sum, then, Socrates offers the following account of weakness of the will: at some time prior to the moment of action, a person correctly believes that A is better for him or her than B. Closer to the moment of action, however, presumably when certain pleasures or pains are nearer at hand, the person falls prey to the power of appearances, and comes to think that B is better than A, and so takes action B. Afterwards, he or she regains the correct belief that A is better than B and regrets the action.⁷ Thus, no one acts against what is believed at the moment of action to be best.

Socrates insists, however, that people who have the measuring art (ἡ μετρητικὴ τέχνη), which he characterizes as a kind of knowledge (ἐπιστήμη, 357b4), will never have this experience, since their knowledge will prevent them from ever mistaking appearance for reality:

If our doing well (εὖ πράττειν) consisted in doing and taking large quantities, and avoiding and not doing small ones, what would be the salvation (σωτηρία) for our life? The measuring art (ἡ μετρητικὴ τέχνη) or the power of appearance (τοῦ φαινομένου δύναμις)? The power of appearance makes us wander (ἐπλάνα) and exchange (μεταλαμβάνειν) the same things up and down (ἄνω τε καὶ κάτω) many times, and regret (μεταμέλειν) our actions and choices with respect to the large and the small. But the art of measurement would make the appearance powerless (ἄκυρον) by showing the truth (τὸ ἀληθές), and remaining (μένουσαν) in the truth we would have peace (ἡσυχίαν) of mind (τὴν ψυχὴν), and so the measuring art would save our lives. (356c8-e1)

Socrates thinks, then, that knowledge guarantees that we hold on to the correct belief about value in the face of appearances which conflict with that belief. Thus, the individual with knowledge will never have the experience of losing—at the moment of action—the correct belief about what to do.

If this account is defensible, then Socrates has removed a challenge to the claim that knowledge is sufficient for virtue, for Socrates has explained the phenomena without positing motivations that are independent of, and may be stronger than, our evaluative beliefs. On Socrates' account, we always act in accordance with our belief about what is best. Thus, for one who has knowledge, which Socrates believes to be immune from the power of appearances, there is

⁷ Brickhouse and Smith 2002, 2005, Devereux 1995, and Penner 1990, 1996, 1997 also defend this interpretation of Socrates' re-description of weakness of the will.

nothing to prevent the person from acting on this knowledge.

As mentioned at the outset, commentators such as Frede, Irwin, and Penner have inferred from Socrates' re-description of mental conflict and weakness of the will that he does not countenance irrational desires in the *Protagoras*. After all, Socrates certainly never says that we have irrational desires. In addition, being ignorant about the comparative values of courses of action need not suggest that an irrational element is at play, especially when the source of this ignorance, namely, the power of appearances, seems to have nothing to do with irrational desires. On the Socratic account, they argue, we simply become confused by the way things appear in relation to varying temporal perspectives. Similarly, acting on a misperception need not be attributed to an irrational element; acting on a misperception may be a simple cognitive mistake. Thus, these commentators infer from the fact that Socrates attributes the source of error to ignorance that he does not think that there are irrational desires.

While this interpretation of Socrates' position is not without textual motivation, I now wish to argue that there are problems with attributing this view to Socrates. My argumentative strategy is as follows: I shall first show that there is a major philosophical problem with trying to explain mental conflict and weakness of the will without appealing to some kind of irrational element. I will then argue that the text does not in any way force us to saddle Socrates with such a poor explanation. Finally I will argue that there is evidence in the *Protagoras* that suggests that irrational desires are at play in Socrates' account of motivation.⁸

There is a philosophical problem with trying to explain motivational conflict without appealing to irrational desires: this type of explanation fails to account for all of the phenomenological data that characterize the experience of conflict and weakness of the will. For example, in many cases of psychic conflict, we are angry or frustrated with ourselves for having motivations that persist and threaten to move us to action despite our belief that we should not act upon or even experience them. In addition, if we do take the action that is in fact bad, we feel conflicted about it at the time of action; we feel, that is, an internal resistance or shame with regard to taking the action. Finally, if we do take the wrong action, we hold ourselves culpable; we blame ourselves.⁹

The phenomenological data tell us something about the experiences in question. They tell us that in some sense, we *are* aware of what the right thing to do is. For in the first case, we would not be angry or frustrated at *specific* desires or beliefs unless we believed that those desires were pulling us towards the wrong action. Similarly, we would not feel conflict and shame at the time of action, unless we felt, at the time of action, that there was something wrong with how we were acting. Finally, we hold ourselves culpable for our failure to do the right thing because we feel that we did realize what the right thing to do was. When we

⁸ See Brickhouse and Smith 2002, 2005, and Devereux 1995, for related arguments.

⁹ Ferrari 1990 appeals to similar phenomenological considerations in giving his distinctive analysis of the argument.

make a simple mistake, we do not blame ourselves for the mistake in the same way that we do when we are weak. When we are angry with ourselves for making a straightforward mistake we might say, 'I should have known better', but in cases of mistakes due to weakness, we say, 'I did know better'. But if the phenomenological data suggest that we are aware of the right thing to do, then we should not understand conflict as a case of straightforward confusion, nor should we understand the ignorance involved as straightforward ignorance. The data suggest instead that there is something in us that is resistant to acting on our beliefs about goodness, to being guided by reason in our actions, and that we are aware of this element in us. It seems, then, that there is some kind of irrational element at play in cases of mental conflict and weakness of the will, and thus a good explanation of the phenomena should make reference to this irrational element.

We should not attribute to Socrates a poor explanation of motivational conflict and weakness of will unless the text forces us to do so. But the text does *not* force us to adopt the view that all motivation is rational, or that mental conflict is a case of confusion rather than conflict. In the first place, Socrates never denies that there is an irrational element in the soul. Second, nowhere in the *Protagoras* does Socrates explicitly characterize motivational conflict as a case of straightforward confusion. Penner 1990, 1996, and 1997 argues that on the Socratic account, psychic conflict must be a case of purely intellectual confusion; specifically, he argues that according to Socrates, when decisions are sufficiently complex we change our minds as different 'gestalts' of the value inherent in the situation appear to us. This conception of mental conflict follows, of course, from the view that there are no irrational desires. But again, Socrates leaves the phenomenon of motivational conflict radically under-described. The closest thing to a characterization of the phenomenon is Socrates' claim that the appearances make 'us wander (ἐπλάννα) and exchange (μεταλαμβάνειν) the same things up and down (ἄνω τε καὶ κάτω) many times, and regret (μεταμέλειν) our actions and choices with respect to the large and the small' (356d4-7). But there is nothing in this description that necessarily implies that mental conflict is a case of straightforward intellectual confusion, for there is no explicit reference to what causes us to fluctuate between one action and another, nor is there any reference to our attitude towards this experience. In short, this picture is perfectly compatible with the view that mental conflict involves desire-induced struggle where we endorse more fully one of the attitudes engaged in the conflict. Thus, unless we are confident that for Socrates there are no irrational desires, we should not conclude that mental conflict is a case of pure cognitive confusion. Finally, we should not infer from the fact that Socrates argues that we always do what we believe to be best and that error in action is the result of ignorance that there are no irrational desires. All we are entitled to infer from this claim is that if there are irrational desires, then they must bear some relation to our beliefs.

In addition, a little reflection on the text itself should give us reason to doubt the claim that irrational desires are not at play in Socrates' account of the phe-

nomena. Recall that Socrates attributes the source of error to the power of appearances. Now certainly Socrates is correct to note that pleasures or pains that are immediate can appear to us to be worth more than they truly are, and that pleasures or pains that are in the distant future can appear to us to be worth less than they truly are. It is a fact that our judgments of value can be erroneously affected by the relative proximity of the goods or bads in question. But now *this* fact demands an explanation. In the visual case, we can explain why things that are nearer look bigger than they truly are and things that are further look smaller than they in fact are by appealing to facts about the nature of human visual perception. But obviously we cannot appeal to facts about the nature of human visual perception in order to understand why pleasures or pains that are near at hand appear greater or lesser than they truly are—pleasures and pains do not, after all, literally *look* a certain way. But then what causes the false appearances?

One likely explanation for the appearances is that how good or bad a certain pleasure or pain appears to us is a reflection of the strength of our current attraction or aversion for it. Moreover, the strength of our attraction or aversion to certain pleasures or pains is, no doubt, often affected by the relative proximity of the pleasure or pain in question. But the strength of our attraction or aversion to certain pleasures or pains can arise independently of reason and can conflict with our reasoned judgment of the worth of the pleasure or pain. Thus, just as facts about our visual apparatus can cause us to see things differently from how they are, facts about the way we desire can cause us to ‘see’ things as worth more or less than they are. Moreover, just as we can realize that our sight is deceiving us in the visual case, we can realize that our desire is deceiving us in the case of determining what to do. Thus, it seems that underlying Socrates’ discussion of appearances is the notion of a non-rational desire—a desire that arises independently of, and so can conflict with, our pre-existent reasoned beliefs about value.¹⁰

In sum, then, a phenomenologically sound account of mental conflict and weakness of the will requires that we posit desires that we recognize as conflicting with our conception of what is best, and it is likely that Socrates is aware of this requirement. Minimally, Socrates never explicitly denies in the *Protagoras* that we have irrational desires, nor does he explicitly characterize mental conflict as a case of straightforward confusion. More importantly, Socrates’ discussion of appearances is most naturally understood as involving in some sense the notion of a non-rational desire. For all of these reasons, we should not conclude that Socrates denies the existence of irrational desires in the *Protagoras*.

II. Irrational Desires in the *Protagoras*

All of the agreement about Socrates’ denial of irrational desires has cut short any attempt to understand exactly what irrational desires are and what role they

¹⁰ Brickhouse and Smith 2002, 24, and 2005, and Devereux 1995, 395 appeal to similar considerations to argue for the view that irrational desires play a role in the Socratic account of motivation.

play in the Socratic account of motivation. We know that an irrational desire is a state that conflicts with our reasoned conception of what is best, but any one of several different kinds of mental states could play this role, and which one Socrates has in mind may affect the soundness of his account. The relevant alternatives are:

- (1) A feeling of attraction or craving for an object or state of affairs with no necessarily associated belief about its goodness.
- (2) An appearance that an object or state of affairs is good.
- (3) A belief that an object or state of affairs is good that is based on an appearance that the object or state of affairs is good.

Devereux and Brickhouse and Smith, who have also argued that Socrates countenances irrational desires in the *Protagoras*, imply, I think that Socrates holds that an irrational desire is a feeling of attraction or craving for an object or state of affairs with no necessarily associated belief about goodness.¹¹ Could this be the conception of an irrational desire that Socrates has in mind?

I do not think so. In the first place, Socrates makes no mention of pure feeling states in his account of motivational conflict and weakness of the will. The many certainly refer to passions, such as the desire for pleasure, pain, anger, fear, and love in explaining the phenomena, but they do not seem to rely on any particular analysis of what these states are (352b-c). Socrates does away with any such expressions in his own re-description of the phenomenon; the only states he mentions are appearances and beliefs. Second, a phenomenologically sound account of the phenomena does not require that we posit irrational desires in the sense of feelings, for these are not the only sorts of states that could arise independently of reasoning and conflict with our reasoned judgments; appearances, and indeed beliefs themselves, could play this role as well.

The fact that Socrates makes no mention of pure feeling states in his discussion, along with the fact that positing pure feeling states is not necessary for explaining the phenomenological data, suggests that Socrates' strategy in refuting the many's account is to give an analysis of the passions that both shows that they will not overcome a person's reason if he or she has knowledge and does justice to the phenomena. But this brings us to perhaps a more important reason for rejecting (1) as an account of irrational desires in Socrates: if irrational desires as Devereux and Brickhouse and Smith conceive of them exist, then why should we be so sure that knowledge would prevail over them in cases of conflict? How could knowledge quell the efficacy of states with no propositional content? Devereux explicitly acknowledges that positing irrational desires as he conceives of them poses this problem for his account. He replies by arguing that there is evidence in Aristotle (*NE* 1145b23-24) that Socrates found the idea that

¹¹ Neither Devereux nor Brickhouse and Smith explicitly characterize the sort of state that constitutes an irrational desire. Devereux 1995, 393 describes an irrational desire as a 'craving'. Moreover, he seems to think that irrational desires cause appearances, but that they are two separate states (395). Brickhouse and Smith explicitly claim to be following Devereux's account (2002, 25, n5; 2005).

nothing could overcome knowledge of the good ‘intrinsically plausible’ (1995, 394). But on this reading Socrates assumes the very issue at hand in his discussion of weakness of the will and in the *Protagoras* as a whole, namely, the issue of whether or not knowledge is sufficient for right action. Brickhouse and Smith 2005 get around this problem by arguing that the individual with knowledge simply will not have strong irrational desires, because these desires, since they distort our ability to discern reality clearly, are incompatible with the acquisition of knowledge. This may be correct, but Devereux and Brickhouse and Smith still face a related problem: since Socrates holds that we always act in accordance with our beliefs, Devereux and Brickhouse and Smith argue that irrational desires in the sense they describe motivate our actions by affecting our beliefs. But, again, how could such a state affect our beliefs? Why, that is, should we think that a pure feeling state would directly affect our beliefs about value? Thus, even if Devereux and Brickhouse and Smith can escape the problem of explaining why the individual with knowledge will never be overcome by irrational desires, they face at least a *prima facie* difficulty explaining how desires as they conceive of them can affect our beliefs, and thus motivate our actions.

Now Devereux and Brickhouse and Smith might object here that even though they must concede that it is a mystery how desires of the sort they posit could affect or be affected by our beliefs, we must posit them, since their existence is the only thing that could explain why things appear to us to be worth more than they are. But this assumes, without warrant, that a feeling of attraction and an appearance of value are two distinct states. That is, while these commentators are correct in thinking that we need to appeal feelings of attraction in order to understand why things can appear to us to be worth more than we have reasoned them out to be, it does not follow that we need to posit two distinct states: an attraction and an appearance of value. Rather, we might think that being attracted to something just is its appearing good. Moreover, there is no textual reason to think that Socrates holds that an attraction and an appearance of value are distinct states, since the only sorts of motivational states he describes are appearances and beliefs. Finally, once one assumes, as at least Brickhouse and Smith seem to think Socrates does, that any time we are attracted to something it appears good to us, then there is no strong philosophical reason to think that an attraction and an appearance are not one and the same state that includes both phenomenological and representational content. Thus, the thought that Socrates posits pure feelings states in order to explain appearances is unfounded. In sum, then, since we cannot explain how pure feelings states could affect and be affected by our beliefs, and since we do not need to posit pure feelings states to explain appearances of value, we ought to reject the idea that Socrates thinks that irrational desires are pure feeling states. Instead, we should conclude that Socrates’ aim in the *Protagoras* is to give a conception of passions and desires such that they are susceptible to certain kinds of rationalization. In this respect, the thought that an irrational desire is an appearance that some object or state of affairs is good fares better, and thus we should move to option (2).

Is an irrational desire an appearance that some object or state of affairs is good? The view that appearances play a significant role in Socrates' moral psychology is textually motivated. In addition, we have seen that it is reasonable to conclude that Socrates thinks that appearances are at the least desire-like states: an attraction just is an appearance of value. Finally, the relationship between appearances and beliefs is clear: if something appears good, then, barring countervailing evidence, we will believe that that thing is good. Indeed, in many cases the slide from an appearance to a belief is immediate.

There is, however, a problem with arguing that irrational desires are appearances, and this is that Socrates argues that knowledge renders the appearances powerless (*ἄκυρον*), but does not necessarily make the appearances disappear (356d8). We can conclude, then, that the individual with knowledge still experiences appearances. At the same time, Socrates describes the individual with knowledge as being in a condition of psychic peace (356e1). It is, however, unlikely that Socrates thinks that having strong irrational desires is compatible with psychic peace. Thus, we ought to conclude that while Socrates thinks that appearances bear a significant relationship to desires, they are not themselves full-fledged desires. In other words, an attraction is not yet a desire.

Now, one might object to this claim on the grounds that there is no distinction between an attraction (an appearance) and a full-fledged desire. But there seem to be many cases that illustrate that there is such a distinction. I might, for example, in some moments find a lifestyle of jet-setting and party-hopping attractive, but not really desire it, since I realize it is incompatible with fulfilling desires or goals that I believe make an essential contribution to living the good life. Or, I may feel a dislike for attending a loved one's sporting events, but it may not occur to me at all to act on this feeling of dislike, since it does not fit into my conception of caring for this loved one, which I take to be vitally important for realizing a maximally good life. Such feelings of attraction or aversion are qualitatively distinct from states that represent their objects as potential reasons for action. Thus, it seems reasonable for Socrates to think that feeling an attraction or aversion to something and actually desiring to act on that feeling are distinct and that even the individual with knowledge may feel attractions or aversions to some things without actually being motivated by these states, or have them ruin his or her peace of mind.

What, then, does it take for an appearance (an attraction) to lead to a desire? The obvious answer is that the appearance must be endorsed as true, and so give rise to a belief. Again, since the move from appearance to belief can be instantaneous, appearances can, especially for one whose views on value are undeveloped and uncertain, bear an almost direct relation to belief and thus to motivation. And so it seems, then, that (3) is the best option from the relevant alternatives. Non-rational desires are evaluative beliefs that are based on appearances and so may conflict with our other, well-reasoned beliefs about goodness.¹²

¹² In claiming that desires are beliefs about value (whether based on appearances or reason),

It is the fact that these beliefs are based on the way things appear to us that explains why they can be resistant to reasoning and persist despite the fact that we have contrary and well-reasoned beliefs about the object in question, for appearances can persist in the face of the truth, and thus unless we are *sure* that an appearance of value does not reflect reality, it is difficult to prevent ourselves from forming a belief on its basis.

This interpretation of Socrates' account has the advantage of enabling him to capture a good deal of the phenomenology of motivational conflict with his explanation. Motivational conflict is a conflict between our well-reasoned conception of the good, and a belief based on the way things appear to us. Although we might arrive at a conception of what we ought to do on the basis of reason, an alternative course of action might still appear to us as good, and thus we may endorse the appearance and come to believe that it is good. And indeed, we can recognize this belief as conflicting with our reasoned conception of the good, and threatening to move us towards another course of action. This conception of psychic conflict explains why we can be angry with some of our motivations, for we trust our evaluations that are based on reasoning more than our evaluations that are based on the way things appear to us. All the same, if these appearances are vivid, and if they persist, then our attention will be dominated by the evaluative belief that is based on the appearances, and we will act on that belief. It is the fact that we have both beliefs (although one is more dominant, or occupies more of our attention at the time of action) that explains the feelings of conflict and shame that often accompany wrong action. Finally, if we do act in accordance with the false belief, we can now see why we hold ourselves responsible for taking the wrong action. We blame ourselves because we had attained, through reason, the information necessary to realize what we should have done, and thus, we should have recognized at the moment of action that the appearances were only appearances, and did not reflect reality.

The fact that Socrates conceives of irrational desires as evaluative *beliefs*, as opposed to desires that are independent of any evaluation of the object of desire or appearances, opens up an interesting possibility, namely, for Socrates irrational desires are resistant to reason but not invariably or essentially so. Irrational desires are still beliefs, and the aim of beliefs is to represent the world; thus any evaluative belief is sensitive to evidence. It is just, perhaps, that irrational beliefs, because they are based on intense or vivid appearances, need more evidence and convincing than a false belief typically should. Thus, on my account of an irrational motivation, Socrates can give an account of how increased rational beliefs, culminating in knowledge, prevents us from having irrational desires.

On the Socratic account, we prevent our likings or appearances from giving rise to beliefs through an act of reason, through convincing ourselves that the appearances do not reflect reality. The more information we have about the

nature of value in the world, what sorts of things tend to appear good to us in what circumstances, and the veridicality of these appearances, the more effective we will be at preventing our appearances from eventuating in beliefs, and thus quelling their motivational efficacy.¹³ And indeed, for the person with knowledge, this could be quite effortless. We know from the *Meno* that Socrates thinks that knowledge is stable because it is secured by reasons, or is completely justified (97e-98a). Thus, when something that is not good appears good to the person with knowledge, he or she will have the means to see quickly that it is not good, since he or she knows what is good, why it is good, and why other things are not good. There is, for the person with knowledge, simply no room for the appearances to eventuate in beliefs.¹⁴

Indeed, on my interpretation of an irrational desire, it becomes obvious why the individual with knowledge would never be threatened by irrational desires. For by all accounts, knowledge for Socrates requires consistent beliefs. Thus, persons with knowledge will not *by definition* have any beliefs which conflict with that knowledge, and thus, since irrational desires are beliefs, they will not have any irrational desires. This explains Socrates' claims about the peace of mind of the knowledgeable: they have no false beliefs and thus no motivations that conflict with their rational conception of what is best.

If these observations are correct, then a dominant component of the standard interpretation is mistaken: we should not think that Socrates attempts to explain mental conflict and weakness of the will without positing irrational desires in the *Protagoras*. Instead, we should conclude that Socrates has a conception of an irrational desire that both does justice to the phenomenological data and explains his faith in the power of knowledge.

III. Moral Psychology in the *Protagoras* and the *Republic*

As previously mentioned, many commentators have held that the moral psychology of the *Republic* which posits irrational desires to explain mental conflict and weakness of the will marks a radical departure from the *Protagoras*. On the view that I am defending, however, Socrates does posit irrational desires to explain the phenomena in the *Protagoras*. If I am right, then we ought to doubt the claim that the moral psychology of the *Protagoras* and the *Republic* sharply diverge. We might still ask, however, whether or not the *Protagoras* and the *Republic* put forward the same conception of an irrational desire. While I do not here attempt a full exposition of the labyrinthine moral psychology of the *Repub-*

¹³ My account of the non-rational part of the soul invites comparison with Aristotle's account of the non-rational part of the soul in *NE* i 13.1102a15-1103a10. Here, Aristotle argues that the non-rational part of the soul is 'non-rational' in the sense that it conflicts with and struggles against reason in the continent and incontinent person, but that it also has reason in the sense that it listens to and is persuaded by reason.

¹⁴ See Penner 1996 for a persuasive account of why knowledge cannot be jarred by the power of appearances. This aspect of Penner's account of Socrates' moral psychology is, I believe, compatible with my interpretation of the *Protagoras*.

lic, I would like to conclude by suggesting that the moral psychology of the *Republic* is importantly similar to that of the *Protagoras* in that (i) irrational desires in both dialogues are beliefs based on appearances, and (ii) these beliefs are resistant to, but not immune from, rational revision. This interpretation is most clearly suggested in *Republic x*.

In *Republic x*, Socrates is concerned to show that art has a bad effect on its audience because it appeals to and strengthens the irrational (ἀλόγιστόν) part of the soul (604d9).¹⁵ In language that is immediately reminiscent of the *Protagoras*, Socrates notes that how we see things is affected by the varying conditions under which we view them. He says:

The same magnitude seen from close at hand and from far away does not appear (φαίνεται) equal to us...and the same thing appears crooked when looked at in water and straight when out of water...and it is clear that every wandering (πλάνην) of this sort is in our souls. (602c7-d1)

Socrates goes on to say that measuring (μετρέϊν), counting, and weighing help us in these cases, so that we are not ruled (ἄρχειν) by the way things appear (φαινόμενον), but by reasoning (τὸ λογισάμενον) and measurement, which are the work of the rational part of the soul (τοῦ λογιστικοῦ, 602d6-e2). Nonetheless, sometimes, even when the reasoning part has shown us the truth about certain objects, the opposite still appears to us. According to Socrates, this shows that there are two parts of the soul:

But many times when this part has measured (μετρήσαντι) and indicated that some things are larger or smaller or the same size as others, the opposite appears (φαίνεται) to it at the same time concerning these same things... Didn't we say that it is impossible for the same thing to believe (δοξάζειν) opposites about the same things at the same time... Then the part of the soul that believes (δοξάζον) contrary to the measurements could not be the same as the part of the soul that believes in accordance with the measurement... Now the part that trusts in measurement and calculation would be the best part of the soul...thus, what opposes it would be something inferior in us. (602e4-603a8)

Socrates proceeds to compare this conflict between beliefs about how things look to the case of conflict between beliefs about value: 'Just as in matters of sight (a person) was at war with himself and held opposite beliefs (δόξας) at the same time concerning the same things, so in matters of action he is at war and fights against himself' (603d1-3). So, we have a prima facie reason at least for thinking that in both the *Protagoras* and the *Republic* the irrational part of the soul is home to beliefs based on appearances, and that motivational conflict is a conflict

¹⁵ I do not address here whether the irrational part that Socrates is discussing in *Republic x* is the appetitive part, or the spirited part, or both.

between two beliefs—one based on reasoning and the other based on the way things appear. When these beliefs are about the value of a certain object or course of action, then no matter which belief we act upon we will be acting in accordance with our beliefs about value.¹⁶ This explains Socrates' claim in *Republic* vi (again, highly reminiscent of the *Protagoras*) that every soul pursues (διώκει) the good and does everything it does for its sake (505d11-e1).

Moreover, there is evidence that suggests that Socrates does think that reasoning can affect our irrational desires. In *Republic* viii, for example, Socrates criticizes the oligarch for the way in which he deals with his irrational desires. Socrates says that the oligarch does not prevent himself from acting on his irrational desires by persuading (πειθων) them that it is better (ἄμεινον) not to act on them, nor does he prevent himself from acting on them by taming them with argument (λόγῳ). Rather he controls himself through fear of losing his other possessions (554d1-3). The fact that Socrates criticizes the oligarch for not taming his desires by reason suggests, of course, that it is possible for a person to quell his or her desires through reasoning.¹⁷

Recently, Nehamas 1999, 264-266 has argued that the psychic division presented in *Republic* x is different from the one drawn in *Republic* iv. He states that while *Republic* iv argues for the view that the soul is divided into three parts: rea-

¹⁶ It might be objected here that this cannot be the correct account of *motivational* conflict since Socrates does not think that the appetitive part of the soul is related to our beliefs about value. I take it, however, that the objections to this view have been well-rehearsed in the literature (see, e.g., Moline 1978, Lesses 1987, Carone 2001, and Bobonich 2002, 216-292), so I will just briefly review the most important of these objections here. The primary evidence for the view that the desires arising from the appetitive part of the soul are independent of our beliefs about value comes from *Republic* iv, where Socrates attempts to distinguish appetites from other sorts of desires by claiming that they are solely for their object (437d-e); he proceeds to say that we should not be dissuaded from this claim by the view that everyone desires (ἐπιθυμοῦσιν) good things (τῶν ἀγαθῶν, 438a). Some commentators have taken this to be a direct repudiation of the view that all desires are for what we believe to be good (see, e.g., Irwin 1995, 209; Penner 1971, 1990; and Reeve 1988, 134-135). This passage, however, hardly provides conclusive evidence that the *Republic* rejects the view that all desires are for what is believed to be good. In the first place, Socrates is cautioning his audience against the idea that the claim that all desires are for good things constitutes an objection to the claim that appetites are distinguished from other sorts of desires by being solely for their object. This leaves wide open the possibility that Socrates thinks that the claim that all desires are for good things is true, but just does not provide an objection to the thesis on the table; and he would be right to think so: for even if we all desire good things, it does not follow that there are not different kinds of desires. For example, we may have some desires—appetitive desires—that arise independently of reason, but bring beliefs about value along with them; other desires—rational desires—arise as a result of reasoning about value. (See Lesses 1987 and Carone 2001 for alternative interpretations of the *Republic* iv passage.) Moreover, there is every reason to think that we should look for an alternative interpretation of this passage. First, in claiming that everyone pursues the good (505d-e), Socrates does not seem to depart from the view that we are always motivated by our beliefs about the good; this suggests, of course, that even when we are motivated by the appetitive part of the soul, we are acting in accordance with our beliefs about goodness. And second, Socrates in numerous places explicitly characterizes the appetitive part of the soul as capable of having beliefs about value (see, e.g., 442c, 562b-c).

¹⁷ See also *Rep.* 443c-d, 549b, and 586d-e for passages that suggest that our non-rational desires can be affected by our reason.

son, spirit, and the appetites, *Republic x* argues for the view that one part of the soul, namely, reason, is further divided into two parts: a rational and an irrational part. Nehamas' primary argument for this interpretation of *Republic x* is that since Socrates describes the rational part of the soul as the subject of the opposite beliefs, it is the rational part of the soul that must be divided, and not the soul as a whole. I close by suggesting that we agree with Nehamas that the division presented in *Republic x* is a division of reason, but disagree with him that this is a different division from the one drawn in *Republic iv*.¹⁸ Rather, we ought to entertain the possibility that Socrates considers the soul as a whole to be reason, but that reason can have both rational and irrational—in the sense I have articulated—desires. This claim makes perfect sense in light of the fact that Socrates conceives of irrational desires as resistant, but not in principle immune, to rational considerations.¹⁹

Department of Philosophy
Southern Illinois University, Edwardsville
Edwardsville IL 62026-1433

BIBLIOGRAPHY

- Bobonich, C. 2002. *Plato's Utopia Recast*. Oxford: Oxford University Press.
- Brickhouse, T. and N. Smith. 1999. *The Philosophy of Socrates*. Boulder, CO: Westview Press.
- Brickhouse, T. and N. Smith. 2002. 'Incurable Souls in Socratic Psychology' *Ancient Philosophy* 22: 21-36.
- Brickhouse, T. and N. Smith. 2005. 'The Socratic Paradoxes' forthcoming in H. Benson ed. *Blackwell Companion to Plato*.
- Carone, G. 2001. 'Akrasia in the *Republic*: Does Plato Change His Mind?' *Oxford Studies in Ancient Philosophy* 20: 107-148.
- Devereux, D. 1995. 'Socrates' Kantian Conception of Virtue' *Journal of the History of Philosophy* 33: 381-408.
- Ferrari, G. 1990. 'Akrasia as neurosis in Plato's *Protagoras*' *Proceedings of the Boston Area Colloquium in Ancient Philosophy* 6: 115-139.
- Frede, M. 1992. 'Introduction' in *Plato. Protagoras*. S. Lombardo and K. Bell trans. Indianapolis: Hackett.
- Irwin, T. 1997. *Plato's Moral Theory*. Oxford: Oxford University Press.
- Irwin, T. 1995. *Plato's Ethics*. Oxford: Oxford University Press.
- Moline, J. 1978. 'Plato on the Complexity of the Psyche' *Archiv für Geschichte der Philosophie* 60: 1-26.

¹⁸ There is no reason to think that Socrates is discussing a further division in the soul, since he explicitly identifies the part of the soul that believes in visual appearances with the part that is strengthened by imitative poetry (605b-c); and he says that imitative poetry nurtures the desires, pleasures and pains that accompany our actions (606d), which must, I think, be a reference to the desires that arise from the appetitive part of the soul.

¹⁹ I am grateful to many friends and colleagues for comments on earlier versions of this article, including Julia Annas, Elizabeth Asmis, Tom Blackson, Eric Brown, Matt Evans, John Ferrari, Monte Johnson, Rachana Kamtekar, Chris Kelly, Ronald Polansky, and an anonymous referee for *Ancient Philosophy*. Gerasimos Santas, Clerk Shaw, Christopher Shields, Nick Smith, Paul Studtmann, Dan Sturgis, Shelley Wilcox, audience members at the Eighth Annual Arizona Colloquium in Ancient Philosophy, and at the group meeting of the Society for Ancient Greek Philosophy at the Central Division APA.

- Lesses, G. 1987. 'The Divided Soul in Plato's *Republic*' *History of Philosophy Quarterly* 4: 147-161.
- Penner, T. 1971. 'Thought and Desire in Plato' 96-118 in G. Vlastos ed. *Plato*. vol. 2. New York: Anchor Books.
- Penner, Terry. 'Plato and Davidson: Parts of the Soul and Weakness of the Will' 35-72 in D. Copp ed. *Canadian Philosophers: Celebrating Twenty Years of the Canadian Journal of Philosophy*. Supplementary Volume 16 of the *Canadian Journal of Philosophy*.
- Penner, T. 1996. 'Knowledge vs. True Belief in the Socratic Psychology of Action' *Apeiron* 29: 199-229.
- Penner, T. 1997. 'Socrates on the Strength of Knowledge: *Protagoras* 351b-357e' *Archiv für Geschichte der Philosophie* 79: 117-149.
- Reeve, C.D.C. 1988. *Philosopher-Kings*. Princeton: Princeton University Press.
- Reshotko, Naomi. 1992. 'The Socratic Theory of Motivation' *Apeiron* 25: 145-170.
- Santas, Gerasimos. 1966. 'Plato's *Protagoras* and Explanations of Weakness' *Philosophical Review* 75: 3-33.