Routledge
Taylor & Francis Group

The verdictive organization of desire

Derek Baker

Department of Philosophy, Lingnan University, Hong Kong

ABSTRACT
Deliberation often begins with the question 'What do I want to do?' rather than a question about what one ought to do. This paper takes that question at face value, as a question about which of one's desires is strongest, which sometimes guides action. The paper aims to explain which properties of a desire make that desire strong, in the sense of 'strength' relevant to this deliberative question. The paper argues that one's judgment about one wants most will sometimes play a verdictive role, partially determining what the agent most wants, and so making itself true.

1. Introduction

I sometimes find myself in a bind: I don't know what I want to do. In the two years following college I didn't know what I wanted to do with my life. This produced a decent amount of anxiety. Less melodramatically, I sometimes find myself at the record store, CD in hand, wondering whether I want to buy it. Coming to a conclusion can take several minutes. And, in general, while my deliberation about what to do next often begins with the question, 'What *should* I do now?' it just as frequently begins 'What do I *want* to do now?' Answering the second question can be as difficult as answering the first.

Of course, 'Do I want this CD?' is elliptical. I must have some desire for it. I have the CD in hand; I'm thinking about buying it. The deliberative bind arises because I also want to keep my money. So we should read the question 'Do I want this CD?' as shorthand for 'Do I want this CD enough?' or 'Do I want this CD more than I want to keep my fifteen dollars?'

But what am I trying to figure out when I ask myself if I want the CD more? There are several reasons to want a philosophical account of what it is to want one thing more than another. Most ambitiously, such an account will be of great value to standard Humean theories of reasons. Sometimes desires conflict,

and so Humeanism will ascribe conflicting reasons to the agent in question. The theory, then, must also assign relative weights to these conflicting reasons. It is natural to think that in these cases, the rational agent will act on her *strongest* desire, and that the relative weights of reasons will correspond to the relative strengths of the desires providing them.[1] But then we need an account of strength.

More modestly, it seems that we try to figure out what we want the most, because sometimes that will be the most reasonable thing for us to do. There are many worthwhile careers I could pursue, and there are many esthetically valuable pieces of music I could listen to (Chang 1998). Plausibly, when faced with these situations, I should choose that career which I most want of the acceptable options, and I should listen to the music I most want to hear. But then what is this aspect of a desire that makes its object the object I *most* want? What is it for one desire to be stronger than another?

One might object that it is enjoyment or happiness rather than wanting that provides me with reason to listen to music (Chang 2004). This answer is also plausible. But fully assessing the relative merits of 'do what you want' compared to 'do what you would enjoy' requires a better understanding of what wanting is and what enjoyment is. So for the purposes of this paper I will assume that one of the following is true. Either enjoyment is something that results from the satisfaction of a desire, and hence deserves to be identified as part of a desire's *phenomenological intensity*, something which on my account is a component of a desire's strength; or else one's desires are independent of feelings of pleasure or enjoyment, except to the extent that these are the objects of one's desires.

I will also assume that desires are not identical to evaluative beliefs, and that their strength can vary independently of the degree to which the agent believes their objects as valuable.[2] As Harry Frankfurt (1987) points out, caring about one's own children more than other people's children does not depend on the belief that one's own children are more valuable than other children. I care about philosophy considerably more than painting, sculpture, physics, or macroeconomics, but not because I believe philosophy to be an intrinsically more worthwhile pursuit. I just find that I care about it more.[3]

For these reasons, the paper will take the question, 'What do I want to do?' at face value, as a question about one's psychology, which we ask in deliberation. 'Strength' and 'wanting more' in the sense at interest in this paper are non-normative properties of an agent's psychological states that explain certain normative facts, such as the weight of the reason provided by a desire, or which action would be most rational for an agent to perform.

In the next section, I will introduce what I take to be the most natural account of what it is for one desire to be stronger than another: for that desire to motivate action more strongly than the other, and for the associated phenomenology to be more intense. I will argue (Section 2) that this natural picture leads, however, to massive indeterminacy in the strength of one's desires, to a degree that would

render them unfit for rationalizing an agent's choice or settling what to do. I then argue (Section 3) that this indeterminacy is mitigated because the agent's answer to her deliberative question, her judgment about what she wants most, can play a role in establishing, definitively, what it is she wants the most.

The paper, then (in Section 4) will argue for an instance of the familiar, if highly contentious, claim that a person's self-understanding has a self-fulfilling aspect, a claim advocated most notably by Velleman (2000, 2006, 2008). The version offered here, however, is not based on any of the explanatory assumptions used by Velleman; rather it will turn out to follow from very widely accepted platitudes about desires and intentional action explanation, platitudes to which many are already committed. So, the paper will show that we always should have expected self-conceptions to have a self-fulfilling aspect, that this was already implicit in a standard understanding of the belief-desire model.

2. The indeterminacy of desire

Here is the most natural thing to think about the strength of desire: one intrinsic desire is stronger than another if it motivates more strongly than the other, and if it has a more intense phenomenology or felt aspect.[4] Motivational strength can be characterized dispositionally: in the simplest case, in which it is directly within the power of the agent to bring about the desired ends, the agent desires p more than q just in case she is disposed to choose p rather than q when she believes those to be her alternatives. In more complex cases, in which the ends are not immediately attainable, the desire's motivational strength will be modulated by how likely the agent believes her option will bring about the desired end. Phenomenological intensity is harder to define, but I assume readers are familiar.

This is primarily a story of the strength of intrinsic desires. While the strength of an instrumental desire can perhaps come apart from the strength of the intrinsic desires on which it is based, when the agent is instrumentally irrational, from the point of view of an agent asking the deliberative question 'What do I want most?' the strength of an instrumental desire will be *in general* a function of the strengths of the agent's intrinsic desires. It will be determined by how much the instrumentally desired object contributes to the satisfaction and frustration of the agent's intrinsic desires, weighted by how strong each intrinsic desire is, and by how probable satisfying the instrumental desire makes their satisfaction and frustration (Cases in which, from the deliberator's point of view, the strength of an instrumental desire is independent of the strength of the intrinsic desires on which it is based will be discussed in Section 3).

The first problem with this account is that motivational strength and phenomenological intensity can come apart. Humberstone (1990) points out that we can characterize at least one form of laziness this way: when the force with which a desire motivates is significantly weaker than the desire's felt intensity. In the other direction, one may find oneself, especially while tired, strongly

motivated to repeatedly check one's email even though it seems to do nothing to bring satisfaction or alleviate frustration.[5] So when motivation and phenomenology come apart, which desire is strongest?

One response is to put all the weight on the phenomenology. There are of course senses of 'wanting more' and 'strength of desire' that refer to the desire's motivational properties. But the sense relevant to the deliberative question 'What do I want to do most?' refers to the phenomenological properties of the desire. My strongest desire, in the sense relevant to the deliberative question, is the desire with the most intense phenomenology, or so the answer would go.

But if we look at a wider range of cases, this position seems implausible. Consider an agent who is lazy in Humberstone's sense: she would be very happy if she discovered that her office was cleaned up, but has no motivation at all to clean it. Now imagine another agent who would be *equally happy* to have a clean office, but she is actually motivated to clean it. I am perfectly willing to grant that the first agent wants a clean office, perhaps even strongly desires it; but it seems obvious that the second agent wants a clean office even more. We might put the point this way, the first agent would like a clean room, but she doesn't care about it as much as the second. In general, if one *feels* strongly about certain possibilities but is unwilling to take steps to realize them, one's concern seems more superficial than one with the same feelings who is also motivated to act.

Let's return to the agent who would be very happy to discover that she had a clean office but isn't motivated to clean it. Let's add that she has a second desire for a cup of coffee. We can stipulate that while she gets only a tiny amount of satisfaction from drinking another cup of coffee (she is already good and caffeinated), she is actually motivated to go get it; she even braves the rain. What does she want more – to have a clean room or drink a cup of coffee? I lack clear intuitions on the matter.

There are two lessons we can draw. First, motivational force contributes *something* to a desire's overall strength. Second, how much it contributes relative to phenomenology is unclear. This is my first point in favor of indeterminacy in the relative strengths of our desires. If strength is to have precise values, phenomenological intensity and motivational force must add up on a common scale, but there seems not to be such a scale.

Even if one finds this argument unconvincing, because one believes that the phenomenology (or the motivational force) should bear all the normative weight, the case for widespread indeterminacy remains strong. Motivational force can change with differences in the environment, sometimes quite trivial ones like the presence of advertisements (that's why advertisements work). And motivation is unstable not just across situations, but over time. I can be strongly disposed to say something unkind for a moment or two, but once that time has passed be strongly disposed to avoid saying unkind things. This sort of instability is clearly relevant to what I want most in the deliberative sense. If I am more motivated to buy the CD now, because of advertisements, but will be

more motivated to try to return it later when the advertisements are gone, that counts against my strongly wanting the CD *in the sense relevant to the deliberative question*, all else being equal.[6]

The phenomenology of our desires is, if anything, even less stable; they vary with mood, attention, vividness, and our distance from their fulfillment or frustration (cf. Bykvist 2009).[7] What's more, our feelings of satisfaction and frustration come in different 'flavors' – the satisfaction of a particular desire might take the forms of happiness, relief, sensual pleasure, the cool sense of achieving intellectual clarity, and so on; frustration can come as sadness, pain, irritation, anger, despair, and so on. Comparing these different flavors on any scale of intensity is hard. As Krister Bykvist (2009, 26–27) notes in a different context, we have no idea what it would mean to be as proud that p as one was pleased that q. Consider: 'I am as proud of running the marathon as I am pleased by the chance to watch TV.' Or imagine the mild excitement one would get from speeding on the highway. Then imagine what it would feel like to be slightly lonelier than that. Perhaps what can be compared in at least some cases is the degree of pleasure produced by satisfying the desire. But pleasures seem to come in different 'flavors' as well. Compare the pleasure that comes from finally seeing the solution to a difficult intellectual problem with the pleasure one gets from drinking a cold beverage on a hot day, or either to the pleasure that comes from listening to a Tom Waits song.[8] Each of these pleasures feels different, and consequently it is hard to compare them with any precision.

The last point should be emphasized: what is difficult is comparing the intensity of different kinds of feelings *with precision*. Compare the pride one might take in completing a marathon for the first time with the very mild pleasure one would derive from watching a mediocre sitcom. It seems that one could know that the former clearly outweighs the latter in its intensity. The claim that sounded strange was one that asserted the intensities of the two feelings were equal: 'I am as proud of running the marathon as I am pleased by the chance to watch TV.' Similarly, it is hard to imagine what it is like to be *just slightly* lonelier than mildly excited. I take this to be evidence that the intensity of phenomenology does not come in precise degrees. Comparisons, then, are possible between different kinds of experience when the difference in intensity is large. But when two experiences (pride and being pleased, say) are close in intensity, there is often no fact of the matter which is most intense (or if there is equality). This amounts to indeterminacy in how phenomenologically intense our desires are.

What's more, phenomenology does not simply arise upon desire's satisfaction or disappointment. There is prospective phenomenology as well, which also comes in different flavors.

Again, this isn't to deny that we have a rough idea of which desires are more intense when the flavors differ – in fact, it is necessary for my account that we do. But we have good grounds for thinking phenomenological intensity imprecise;

and so it will often be indeterminate which of a set of desires has the most intense satisfactions and frustrations.

A final source of indeterminacy may come from a distinction between those wants that should be categorized as values and those that are mere desires. An agent may be more strongly motivated to continue smoking than to quit, and the craving for another cigarette may come with much more intense feelings than the desire for health. Nonetheless, there may be reasons to think she *values* her health more than she values smoking. Is the fact that she values health more reason to think that she *wants* health more?[9]

For my purposes I can plead neutrality: whether valuing is a state distinct from being motivated or disposed to have certain experiences, and whether that distinct state contributes to how much one wants something, will depend on larger claims about moral psychology – on one's particular theory about what valuing is. However, if valuing does contribute to how much one wants a given end, this simply introduces further grounds for indeterminacy in the relative strength of desires. Not only must we compare motivational facts with phenomenological facts, but these must be compared with facts about valuing as well.

To summarize, the obvious view is that the strength of a desire is determined by its motivational and phenomenological properties (and possibly by more complex states as well, such valuing the relevant end). But motivation and phenomenology are complex phenomena, and each complexity seems potentially relevant. We have motivational force and its variation across situations and over time. We have the range of emotional responses which satisfaction or frustration (and their anticipation) might bring, both across situations and times. Each of these differences can give reason to call a desire stronger or weaker. But how important is each element relative to others? How can we construct a scale on which all of these elements add up?

Occasional indeterminacy would likely be theoretically unproblematic (however undesirable for the agent who finds herself in such a situation). But the preceding argument suggests that it is potentially widespread. Indeterminacy is virtually guaranteed to occur in any of the larger dilemmas in which the question of what one wants arises – such as 'What do I want to do with my life?' or 'Do I want children?' (The flavors differ; toward these questions my feelings and inclinations vacillate; if I am honest I know that my motivation and enthusiasm for carrying out the decision will vary over time and context). Indeterminacy is also likely to arise in those smaller conflicts where personal pleasures are at issue, such as 'Do I want to listen to music or go on a hike?'

Perhaps we should accept this. But the case against settling on such widespread indeterminacy is straightforward: it is precisely when confronting problems like the career to pursue or whether to have children that people begin deliberation with the question 'What do I want?' It would be strange to discover that we are inclined to ask ourselves, and treat as normatively relevant, a question that could have no answer in exactly those situations most likely to inspire us to ask.

3. Verdictive self-reflection and the legal analogy[10]

I will argue here that the agent's own judgment about what she most wants to do will often mitigate indeterminacy. The judgment can do this, because, I will argue, there is reason to think that such judgments can function as *verdictives* – that is, as reports on some matter of fact, which at least sometimes succeed in making it the case that that very fact reported obtains (Austin 1962). Some care is needed here: a verdictive is a type of speech act – the classic example being the referee's judgment that one team has scored a point against another, a judgment which makes it the case that the team has one more point than previously. The care is needed because I do not wish to argue that self-reflective judgments are a speech act; they are instead a form of belief. What's more, verdictives are identified as a type of action; hence there is a question whether one must intend for one's speech act to be a verdictive if one is genuinely to issue one, or whether the speech act can be performed unintentionally. This leads us to a question of whether speech acts should be categorized on the basis of communicative intention or communicative effect, a problem that would take us far off course in this paper.[11] Fortunately, we can put it to one side and say that, however we choose to categorize them, there is a class of speech acts that will have a *verdictive function* even if the speaker intends only to issue a garden-variety report. The explanation of how these reports could come to possess a verdictive function, regardless of speaker intention, will shed light on how a belief could come to have a self-fulfilling aspect, without positing any desire, drive, or intention to vindicate the belief.[12]

I will use HLA Hart's (1961) discussion of *the law's open texture* and the role that judicial verdicts play with respect to it as an example that sheds the most light on how self-reflective judgments could function as verdictives. For Hart, the correct interpretation of the written law is often indeterminate (124–136). No law could specify its correct application in every possible scenario, and so there will inevitably be cases in which a range of reasonable interpretations exists. How bad must driving be, before it counts as reckless? Does a ban on vehicles in the park apply to skateboards? Does it apply to strollers? (126). So, 'at the margin of rules… the courts perform a rule-producing function …'(135). In normal cases, their verdicts simply report on what the law mandates. But in cases where the law itself is indeterminate, the judge's verdict is partially determinative of the facts about the law.[13]

Whether or not this is correct about the law, it's what Hart should say, given his commitment to legal positivism. For Hart, legal facts are a species of social fact – facts about what is enforced by state authorities, what is deferred to as authoritative by the population, and so on. If a set of prescriptions count as *law* in a given society, it is because they guide actual behavior; they play an explanatory role in the society. We should expect, then, that judicial verdicts

would sometimes be determinative of the law: they are themselves objects of social reality, and so potentially part of the law's grounds or basis of reduction.

The judicial verdict is a description of a legal prescription, made by a figure others take to be authoritative. This ruling on what the law says will thus determine the behavior of legal officials. Bailiffs, police, and prosecutors will use it as a guide in enforcement; judges will cite it as precedent in their own rulings; other lawyers will cite it in their arguments and in their advice to clients. In other words, the verdict plays the causal-explanatory role of law: it supports the right complex of dispositions in the society. The legal fact exists in virtue of facts about people's behavior; the judge's verdict changes people's beliefs about the law, causing their behavior to change. So the fact the verdict describes will sometimes be explained by the verdict itself.

Admittedly, judges may sometimes come to the conclusion that the law is indeterminate, and so *decide* to consciously exercise their rule-producing power – that is, they may intend to issue a verdict that is also a verdictive. Likewise, the relevant legal authorities may believe a particular verdict to have produced a *new* rule. In these cases the verdict will seem much more analogous to a decision or stipulation than a report or belief. But the important point is that the verdict *can be* intended and understood simply as a report on the law, and still serve a rule-producing function. Given the structure laid out above, all that is required is that people change their behavior to conform with the ruling, and they may well do that because they take it to be an accurate report on the content of the law (which it is, thanks to their compliance). And what's more, there are obvious reasons why judicial verdicts may sometimes play their verdictive role more effectively if the parties involved understand them as responses to evidence which attempt to get things right.

Returning to the theory offered here, it is a commonplace that belief-desire psychology is both normative and explanatory.[14] At the very least, *strength* of desire has been assumed to be a property with both explanatory and normative roles: one desire is stronger than another in virtue of a certain counterfactual profile; but at the same time, judgments about the strength of one's desires are, or at least often function as, implicit judgments about what one has most reason to do, or which action one's attitudes would best rationalize. This judgment will, in successful cases, regulate self-conscious deliberation, choice, and the exercise of will; in other words, it will instantiate the sorts of psychological dispositions to which the property of strength reduces or which ground it.

The agent asks herself 'What do I want to do?' This question involves reflection on her psychology. Sometimes, it is already decided by the facts in place. Even if the degrees of strength are fuzzy, only a limited number of desires will be at stake in a given choice situation, and of those, one may be simply stronger than all competitors. 'How much stronger?' may have no answer of course – but in a simple decision problem, where I can immediately effect the relevant end, I don't need to answer *that* to resolve deliberation. Even if uncertainty is an issue,

I will need some idea of how much stronger one desire is than another, but how precise this needs to be will vary with the decision problem I face.[15] And again, in many cases, one desire is simply strong enough, relative to competitors and the likelihood of the various outcomes that could follow from taking each option, that one of these options is determinately superior. One makes a judgment, then, about which desire is stronger on the basis of one's evidence – the evidence being one's evidence about the desires' motivational and phenomenological profiles. On the basis of that judgment one will conclude one's deliberation, ending, assuming there is no weakness of will, in a choice. If the judgment is accurate the choice will be reasonable (assuming it was already reasonable to pursue what one wanted the most); otherwise it will be unreasonable.

Sometimes, however, there is no prior fact of the matter. But even if there is no prior fact of the matter, the evidence available to the agent can seem (at least to that agent) to favor a conclusion. This will lead her to a belief about what it is she most wants, and that belief will often function like a verdictive: the belief that she wants one option more than another will make it the case.

After all, if I believe myself to want something, this allows me to consciously decide to pursue the object, or consciously plan to achieve the goal. A desire or goal that I don't realize I have is not one that can figure in conscious deliberation. If I believe I desire some goal more than its rival, then in conscious decision-making I will, in ordinary cases, plan to bring the goal about even at the expense of the rival. And it will make more sense to me to exercise willpower (that is, consciously direct attention, or exhort myself) on behalf of the desire when it conflicts with its rival. These various aspects of conscious deliberation will then influence my behavior, making my behavior more likely to conform to the belief.

In short, a judgment about a desire's strength is an implicit judgment about the desire's normative significance, and so will tend to play a role in directing choice (that tendency, of course, moderated by tendencies to akrasia, to abandoning conscious plans and decisions, and other familiar defects of practical rationality). Because of her belief about the desire's strength, the agent will tend to have dispositions to act; and these are the dispositions that (partially) determine the strength of a desire. Notice that, just as a judicial verdict can play a verdictive role without the judge intending to issue a verdictive but simply to report on the law's content, the relevant belief can determine the strength of a desire while being a *belief* – an attitude that represents how things are rather than how they are to be.

This last point is important, because the argument that our beliefs about our desires sometimes alter their strength is not meant to follow from introspective evidence. I fully admit that when I inquire into what it is I want, it feels to me like I am discovering, not creating, something about myself. But this is what the verdictive model predicts. Just as a judicial verdict can play its verdictive role without seeming to participants in the legal system like a decision to change the law, so a self-reflective belief can play a verdictive role in my psychology, even

while it seems to me to simply be a report on an independently existing reality. Insofar as there is familiar evidence favoring the verdictive account, it is simply the fact that I can better and more consistently pursue my desires when I know that I have them (which, if a desire is a state with a functional role of disposing an agent to pursue its satisfaction, means that knowing what I want tends to improve the ability of my want to play its functional role).

But the central argument for the verdictive model is simply that it follows naturally from standard theoretical commitments. Like the law on Hart's positivist conception, an agent's attitudinal psychology has normative and explanatory dimensions: desires both causally explain further instrumental motives and actions, and rationalize them. The judgment that one has a desire of a certain strength therefore is an implicit judgment about which motivational dispositions it would be rational for one to possess – namely, the very same motives which would be explained by a desire of the strength one believes oneself to have. The agent making the judgment, as a rational agent, will thus have a general disposition to acquire those instrumental motives, making her psychology more closely resemble the one she takes herself to have.

One worry with this account might be that it allows beliefs to alter an agent's intrinsic desires (or at least how strong those desires are). But how can a belief do that? The answer is similar to Michael Smith's account of the practicality (or motivating role) of moral beliefs (2004b, 'Defence,' 273–274). Self-reflective beliefs about one's desires are implicit judgments about which instrumental desires or intentions it would be rationally coherent to possess. Failure to possess attitudes one believes that it would be more rationally coherent for one to possess is a way of being at odds with oneself, and hence is a failure of rational coherence. What's more, rational agents possess a general disposition to update their attitudes in the direction of greater rationality (ibid.). Given this, and the fact that a belief about the strength of a desire is an implicit judgment on which instrumental motives it is rational to have, the agent's general disposition to greater rationality will tend to result in the belief playing a verdictive role.[16]

Note that the verdictive role of self-reflective beliefs is explained by a general psychological tendency toward greater rational coherence. However, because we are imperfectly rational, there will be cases in which this tendency fails to manifest. In these cases, self-reflective judgments will fail to play their verdictive role. (Thanks to imperfect rationality any attitude-type can fail on some occasions to play some of their functional roles.) The theory also leaves open the possibility of more complex cases – in which the belief results in a general motivational disposition, but fails to motivate on this occasion. These will be cases in which the belief plays its verdictive role, although on this occasion the agent displays weak will (Whether the agent has a desire of the relevant strength depends on what she is disposed to do – and what she does on this occasion may be excellent evidence about her dispositions, but it does not determine the matter).

There are several additional aspects of the verdictive account that deserve clarification. First, is the belief a partial determinant of the desire's strength in a constitutive or a causal sense?[17] The basic verdictive model is compatible with either answer. My preferred answer is that it is constitutive in the ideal case, though it may be causal in cases that depart from the ideal. In the legal case, the judicial verdict will often be a constituent of the legal fact it makes true. Given the surrounding legal infrastructure (systems of disseminating the ruling, state officials, lawyers, dispositions of deference among the public) the ruling will itself instantiate the relevant dispositional facts – such as being disposed to cause judges to rule in certain ways on certain cases – that make it true that the law prescribes as the verdict says it does. Similarly, given the agent's background rationality, the belief about the strength of one's desires will realize the relevant functional properties. It is only in less ideal cases, in which the agent's background rationality initially fails, that the determination will be causal. These will be cases in which the judgments are initially false, but still do enough work guiding one's action, deliberation, imagining, and self-declarations that they start to shift one's conative profile in the direction of their truth.

In cases where the relation is causal, the judgment will initially be false, though it will become true more or less quickly. More generally, it should be obvious that the model does not in any way entail infallibility about the strength of one's desires. The judgments exercise a verdictive function because of their implicit normative import – but normative import sometime fails to move us.

Nor does the model imply that we are always rationally obliged to act on our acknowledged desires. Of course I can acknowledge myself to have a desire but reject acting on it, and nothing about the verdictive model says otherwise. Humeans who accept the model can still say this rejection is rationalized by an even stronger acknowledged desire, while for those who reject Humeanism it may be rationalized by something else – most likely one's normative beliefs.

It is also worth noting that nothing in the model suggests that self-reflective verdicts fully eliminate indeterminacy. In fact, there is good reason to suspect they could not. Unless I am extraordinarily self-controlled and mindful there will still be possible circumstances in which I would act contrary to my verdict; and the phenomenology associated with two rival desires may remain unaffected by that verdict as well.[18] But the motivations of the theory are compatible with some indeterminacy even after the verdict is reached. All that is needed to resolve the deliberative question is that the verdicts mitigate indeterminacy, so that the agent definitively prefers one of her options: the claim here is that often it enough to determine such a preference that the verdict would lead the agent to systematically choose pursuit of X instead of Y in a wide range of options.

Another question is whether the verdictive account here applies only to intrinsic desires, or to instrumental desires as well?[19] The explanatory story told here suggests that it should work with both. The story, remember, is that the belief that I have a desire with a certain strength puts rational pressure on me to

act in a manner consistent with a desire of that strength. Given that instrumental desires can rationalize further instrumental aims, a judgment that I desire some means (money, say) to a certain degree should put rational pressure on me to make choices (investing more, working extra hours) consistent with that instrumental desire.

Now in many cases I do not need to come to an independent assessment of how much I want an instrumental good, because the strength of my instrumental desire simply follows from how much I want the ends that they promote, especially from the point of view of deliberation. If I judge that what I want to do most now is go on a hike, this will put rational pressure on me to form an instrumental desire to put on my shoes, which will rationalize a further desire to walk to where my shoes are. Forming an independent judgment about how much I want to put my shoes on would be unnecessary. But in some cases an independent belief about how much I desire some means may serve the purpose of bringing consistency to my deliberation.

Consider again the desire for money. Money is instrumental to a number of things I desire intrinsically – security, health, happiness, the happiness of my children, their education, and so on. Actually enumerating all of the ends served by having more money, coming to assessments of my relative strength of concern for each of these ends, and then assessing how likely it is that a given increase in my net worth will allow me to achieve these ends might be too complex a problem for me to solve. Coming to an independent judgment of how strongly I desire a certain amount of money may allow me to resolve dilemmas such as whether to buy a new CD by simplifying them. Then I don't need to come to a judgment about how much I want each of the possible things on which I could spend that 15 dollars, let alone the uses to which I could put it after allowing it to sit in an investment account for 20 years.

Admittedly, this could lead to instrumental desires that are inconsistent in strength with the intrinsic desires on which they are based. But this is just to note that agents will not always be perfectly instrumentally rational, something which is true of actual human agents anyway. It is not unique to the verdictive account.

While other advantages of the verdictive model will become apparent in the following section, the most obvious is this: it allows us to reconcile the commonsense picture of strength of desire, in which the desire's strength is determined by its dispositional profile, including both dispositions to act and dispositions to experience feelings of certain intensities, with the fact that we often frame deliberation in terms of what we most want, and are able to resolve deliberation so framed.

4. Alternatives and objections

This section addresses alternative accounts of how agents sort and organize their desires and considers objections to the verdictive account.

4.1. Comparison with Velleman's theory

Because of obvious similarities, it is worth contrasting the account here with David Velleman's well-known theory. (A full critique of Velleman's account is beyond the scope of this paper; the comparison is presented here for clarification.) Velleman also argues that self-reflective beliefs can make themselves true, or determine the facts about their objects, again largely by the way such beliefs will influence subsequent choice. For Velleman the explanatory mechanism is different, however. In the earlier (2000) work it is a desire for self-understanding; in the later, it is a sub-personal drive to know the truth. In either version, this desire or drive will be satisfied if the agent's behavior conforms to her predictions about what she will do. This provides significant motivation to bring her behavior in line with her predictions – which consequently function as self-fulfilling prophecies. Intentions, for Velleman, just are beliefs that function as self-fulfilling prophecies.[20] Ideal agency is consequently a matter of forming a relatively comprehensive and detailed self-conception which the agent is motivated to comply with in order to have a more accurate picture of the world.

The verdictive judgments proposed here are not predictions at all (though various predictions will feed into them as evidence). When I judge that one desire is stronger than another, I am not implicitly predicting that I will actually act on that desire in the next relevant situation. I must believe that I would act on the desire in some host of more favorable circumstances, but that is compatible with the belief that current circumstances are unfavorable, that, for example, I will give in to temptation. The judgment is an implicit commitment about which action would *be most rational for me.* But this is to say that it directs my behavior because there is an implicitly normative component to the judgment, which determines choice in the standard way that normative assessments direct our deliberation – not by serving as a self-fulfilling prophecy, and not because I have any desire or drive to achieve an accurate picture of the world.

Velleman's view has been developed over the course of over two decades, and it would be beyond the scope of this paper to offer a decisive argument against it here. Nonetheless, several reasons can be given here to hope the verdictive account will ultimately be more promising.[21] Most important is the weakness of the theoretical commitments needed for the account. Velleman must posit a motivation, universally at work in normal adult human beings, to understand oneself. The model here depends only on the standard platitude that desires play both an explanatory and a rationalizing role. A fundamental drive for knowledge is thus unnecessary, because the idea that self-conceptions

would have a self-fulfilling aspect was always implicit all along in the commitments of the standard belief-desire model.

The verdictive model, moreover, avoids commitment to the very contentious thesis that intentions are actually a species of belief. The simplest case against this thesis is that it seems possible to intend some action without believing that one will perform the action; one can intend but be unsure one will follow through (Holton 2009; Ross 2009).

One can object to Velleman's theory that the formation of self-fulfilling predictions must be epistemically irrational (Langton 2004). Imagine that one is trying to determine whether it will rain tomorrow or not. You are informed, by God, that it will rain tomorrow if you believe it will and won't if you believe it won't. On what basis could you possibly reach one conclusion or another? Similarly, once someone comes to know the truth of Velleman's theory, how does she continue to form intentions?

It may seem that the verdictive account is equally vulnerable to this objection. After all, once someone knows his desire is likely to be indeterminate, how can he continue to conclude there is something he most wants? But the verdictive model can offer a straightforward answer: by believing irrationally, something eminently possible for us.

Theoretical and practical reason may sometimes make incompatible demands. The evil demon may torture me if I fail to believe an additional unsupported falsehood; it may be easier to jump across the gorge if I believe I will be successful at jumping across the gorge. The objection points out that the verdictive theory generates similar conflicts. Agents in a certain evidential situation cannot engage a standard capacity for achieving greater practical rationality without committing epistemic irrationality. The verdictive account can simply allow that this is true and deny it is a significant objection.[22] In these scenarios some will remain epistemically rational, and so unable to resolve the indeterminacy; others will trade epistemic irrationality for greater practical unity, and thus ignore or partition evidence of indeterminacy in their own case.[23]

Velleman would be hard pressed to similarly deny the force of the objection. Reasons for action, on his account, are reasons to form the relevant belief about oneself. The objection denies that one can be aware of the self-fulfilling nature of these beliefs and continue to have sufficient reason to justify forming any of them. The theory cannot present itself, then, as a theory of practical reason that may sometimes be at odds with what's epistemically required. That answer requires that reasons for belief and reasons for action be distinct.

4.2. Organization through attitudes other than belief

This discussion of the previous subsection raises another question, though, of why we should insist that a determinate answer to the deliberative question 'What do I want to do?' is settled in cases of prior indeterminacy through a *belief*.

There are many theories allowing for the possibility of self-creation, reflective self-governance, or acts of defining commitment. They propose that these are achieved through long-term intentions (Bratman 2004), such intentions working in tandem with higher order desires (Tiberius 2000), or through an act of decisive identification (Frankfurt 1987). Connie Rosati (2003) takes it that what is in the agent's personal good might itself be indeterminate, allowing for the possibility of the agent choosing to make it more determinate. Ruth Chang (2009) argues that incompleteness of reasons can be resolved through an act of will that creates new reasons for the agent in question. These theories all have implications about how we shape our motivational profiles, and if they do not directly address how this would allow us to answer the deliberative question 'What do I want to do?' they or a similar theory could plausibly be modified to do so.[24] So why not allow one of these, unbelief-like attitudes or acts to do the work of imposing determinacy, thereby sparing us worries about epistemic rationality?

My argument at this point will be modest. One or several of these theories may be correct; they may describe actual capacities we possess. In typical cases, though, it could not be through the exercise of any of these capacities that actual agents determine what it is that they want. Despite their differences, all of these theories involve the idea that the relevant endorsement or commitment is volition-like, act-like, or decision-like; it is something the agent actively forms. This gets the experience of wanting wrong, at least in most cases. In general, we experience ourselves as passive with respect to desire. Desire is something that happens to us, something we discover, not something we choose.

Romantic love is frequently felt this way, as is love for one's children; and romantic comedies often use the trope of the protagonist realizing him- or herself to be in love through an epiphany. This is fiction, of course, but seems to correspond to something about how it is typically experienced; it isn't something one consciously decides to commit to; one experiences oneself as *falling* in love.[25]

One's commitments to various projects are often experienced in the same way. We do sometimes choose our commitments, of course; but deliberating on what one wants is often resolved through what *feels like* a recognition rather than an act.

The verdictive model shares with many of the theories above the sense that agents – especially through their reflective powers – play a role in shaping their motives and concerns; nonetheless, it is compatible with the standard experience of passivity. Because beliefs are not formed voluntarily, but are instead a result of the world impressing itself upon us, this capacity can be operative, even while we experience our desires as things that largely happen to us. But though we might *experience* ourselves as passive with respect to our beliefs, they are still products of our rational activity, something we play a role in producing.[26]

4.3. *The epistemology of self-knowledge and desire*

Another objection to the verdictive account of desire offered here is that it is incompatible with a popular view about the nature of self-knowledge. According to these views, we know our own mind, not by attending to ourselves, trying to detect some psychological state, but by attending to the outside world. As Gareth Evans argued, one does not answer the question 'Do you believe there will be a World war III?' by introspecting for a belief. Rather, one considers the evidence for and against the occurrence of a World War III (Evans 1982). In short, one looks out at how the world seems to be, and knows one's belief on that basis.

This point has inspired a wide range of *transparency accounts* of self-knowledge (Moran 2001; Fernández 2007; Byrne 2011a, 'Thinking,' and 2011b, 'Want'; Valaris 2014).[27] These differ on the details, but all agree that questions about one's own attitudes are transparent to questions about how the world is. This is directly contrary to my descriptions of how we typically answer the deliberative question 'What do I want?' Here, I've treated the question as one about the agent's own subjective state, rather than some state in the world. I've had the agent respond to evidence – constituted by her feelings and motivational history – to come to a judgment about what that subjective state is. So the verdictive account of deliberation depends on a rejection of the transparency account of self-knowledge.

Fortunately, there are alternatives to the transparency account of self-knowledge, and there is reason to think that, whatever its merits as an account, say, of knowledge of one's beliefs, it fares poorly as an account of knowledge of one's desires. The problem with these accounts is simply the one put forward at the beginning of this paper: it is part of common sense that the state of the world underdetermines what it would make most sense for me to want – that is, even after specifying how the world is, or how I take the world to be, there is no set of desires that constitute the *uniquely* appropriate response to that specification (Way 2007).

To make this clearer, we can ask to which question about the world my question 'What do I want?' is supposedly transparent. The standard answer is that I know what it is I want by looking out on the world and determining what is worth wanting – what is good, desirable, what I should want (Moran 2001; Fernández 2007; Byrne 2011b, 'Want'; Valaris 2014). But this leads to two problems. First, there are many more things that I recognize as valuable that I nonetheless do not want: I do not wish to pursue every valuable career and some music I acknowledge as valuable is not to my tastes. A transparency account of self-knowledge fails to explain how I could know this about myself.[28] Second, I often know myself to want one thing most, even in cases where I do not regard any of the possible ends as best. But again, the transparency accounts wrongly suggest that questions about what I want most are transparent to questions about what is most desirable.

Fortunately, Krista Lawlor ([2009]) account of knowing one's own desires is more congenial to the verdictive account.[29] Lawlor argues that knowledge of one's own desires comes from various *internal promptings*: one feels various sensations in the case of simpler desires, or 'imagined natural language sentences and images' in the case of more complex desires (49). Lawlor also argues that we sometimes self-ascribe a desire, but provisionally, and wait to see if the self-ascription feels accurate as time goes on (57). In some cases, we might find ourselves spontaneously resistant to the self-ascription (ibid.). On the basis of these kinds of internal promptings and feelings, we infer what kinds of desires we likely have, which desires would best explain these promptings.

This account seems accurate as a story of how we come to know what it is we want, especially with large and complex decisions such as which career one wants to pursue, or whether one wants children. It is also compatible with the verdictive story offered here – the verdictive account simply adds that judgments about one's desires will sometimes possess additional accuracy, because the facts will shift to accommodate the judgment.

4.4. Is the theory still compatible with Humeanism about reasons?

Before concluding, I will address one final potential misunderstanding: the worry that this theory could not serve as a contribution to Humean normative theory, because it explicitly requires that beliefs can change one's intrinsic desires.[30] Keep in mind that, even if this objection were correct, the verdictive model could still play one of the other, less ambitious theoretical roles discussed in Section 1 (the theory is meant to be compatible with Humeanism, not to assume it). But the worry is misguided. For certain theorists, the *empirical claim* that beliefs do not alter intrinsic motives is an important part of their Humeanism (Sinhababu [2009]; Arpaly and Schroeder [2012]); however, this has hardly been a standard commitment for normative Humeans (who, after all, are offering a theory about reasons, not necessarily a theory of the etiology of desires). Bernard Williams, for example, argues that aspects of deliberation such as the exercise of imagination can generate new desires:

> In his unaided deliberative reason, or encouraged by the persuasions of others, [the agent] may come to have a concrete sense of what would be involved, and lose his desire for it, just as, positively, the imagination can create new possibilities and new desires. …
>
> We should not, then, think of S [the agent's motivational set] as statically given. The process of deliberation can have all sorts of effects on S, and this is a fact which a theory of internal reasons should be very happy to accommodate.
>
> ([1981], 104–105)

Similarly, Donald Hubin, in his ([1999]) elaboration of the commitments of Humeanism about practical reason, writes of the possibility of deliberation bringing into existence new desires for ends:

…I think that we are often [when deliberating] engaged not in practical reasoning but in a precondition for practical reasoning. And this precondition is not the discovery of a preexisting conative structure with which to define our evaluative point of view. We are not trying to discover our evaluative point of view; we are creating it. Just as we entertain a hypothesis to draw out its implications and (sometimes) to see the world as we would if we believed it, we at times "try on" normative stances. If I am right, this can be more than a process of gaining knowledge of our values; it is sometimes an act of self-creation.

(41–42)

The theory offered here is different in that it is not imagination – imagining possibilities or 'trying on' evaluative perspectives – that alters one's desires. Rather, it is a kind of belief. Nonetheless, what the view here shares with these pictures is a commitment that in deliberation about one's final ends, there is no rationally mandatory or correct answer about which ends to have, independent of the answer the deliberating agent comes to. This, moreover, is the only commitment about the role of deliberation in producing motives that a normative Humean would have to accept (otherwise passion would in some cases be subject to reason). It is easily captured by a theory that models such conclusions off judicial interpretations of indeterminate law. Rather than a rejection of normative Humeanism, the verdictive model is simply one particular way for the Humean to make sense of the real phenomenon of deliberation about one's final ends.

5. Conclusion

It is worth noting again that the verdictive role of self-reflection seems to follow very naturally from standard platitudes about the belief-desire model. This is perhaps the account's most notable theoretical virtue: its explanations come compliments of the theoretical house where many of us already reside. It posits no additional capacities, attitude-types, or aims. There may well be more theoretical commitment in trying to justify the decision to send it back, than in simply accepting the account.[31]

Accepting this complementary theoretical machinery allows us a straightforward explanation of how we are able to resolve deliberation framed in terms of our wants, despite the fact that the most natural picture of how desires are ranked implies widespread imprecision in that ranking. It offers a promising account of weighting for a Humean theory of reasons, and additionally suggests for the Humean a more exact interpretation of what is going on when agents deliberate about their final ends. It provides a mechanism by which agents can determine their personal projects and other idiosyncratic concerns, but which also explain why we rarely experience ourselves as choosing what to care about. In the course of doing so it makes room for the possibility of self-fashioning experienced as passive.

This last point – that self-reflection on this model is heavily tied to a form of self-fashioning – suggests another point at which the legal analogy might

be appropriate. The verdictive model may point to the possibility of kind of autonomy not dependent on acts of self-legislation, but instead evolving out of a case law or common law of the self.

Notes

1. Mark Schroeder (2007) identifies this position as Proportionalism; but note that he offers an alternative Humean answer to the problem of weighting reasons, which he calls Hypotheticalism.
2. Thanks to a referee for helping me make this point more clearly.
3. Also see (Bratman 2012, especially p. 93).
4. See (Sinhababu 2009) for a picture like this. There are many other theories of desire; however, the aim of this paper is simply to introduce a new theory and address some of the more obvious objections. Summary and critique of other views would make it unacceptably long. Readers should consider, however, views which seem to identify the rational force of a desire with its motivational force after cognitive enhancement (cf. Williams 1981); views which identify its rational force with its role as a quasi-perceptual state representing value (Oddie 2005; Schafer 2013); and Timothy Schroeder's empirically informed account which identifies a desire's rational force with the amount of *reward* its satisfaction would bring (see his 2004; 2010; Arpaly and Schroeder 2012). For an alternate account of how reward and desire interact, see Richard Holton's (2009, 101–111) discussion of empirical work on addiction, also summarized in (Berridge 2004). Finally, note that all of these other theories may in fact be consistent with the primary conclusion of this section – that often the relative rational import of a desire is indeterminate. But discussing this for each possible theory would be beyond the scope of this paper.
5. See (Holton 2009, 101–111) for discussion of empirical studies showing that motivation and feelings can come apart; for an example of a recent study, see (Berridge 2009, especially §2.2).
6. See (Sinhababu 2011) for discussion of similar cases and the argument that they actually increase the plausibility of Humean theories of reasons, by allowing Humeans space to discuss and explain irrationality; also see (Hubin 2001, 2003; Sobel 2011, 66–70).
7. I have here assumed that the felt aspect of desires has to do primarily with their satisfaction or frustration. But a referee worries that there may be cases in which I would be happier if some counterfactual were true, but I don't desire the situation at all – perhaps because I never considered the possibility. If this is so, *anticipatory* phenomenology might be more significant. Feelings of excitement, fear, longing, and so on may play an especially important role in determining how much I want some end. This may be: but the reasons for thinking the phenomenological intensity of satisfactions is in many cases indeterminate are also reasons to think the same about the intensity of anticipations.
8. Thanks to Nate Sharadin for help with making this point.
9. Thanks to an anonymous referee for making this point.
10. My ideas here have been heavily influenced by the work and insights of (Taylor 1976; McGeer 1996; Moran 2001; Helm 2002; Rosati 2003; Velleman 2000, 2006; 2008; Wood's 1990 discussion of the relation between freedom, happiness, and desire in Hegel, especially pp. 58–71).
11. Thanks to Louise Anthony for calling this problem to my attention.

12. The idea of understanding self-reflective beliefs via an analogy with speech acts is inspired by (McGeer 1996), in which such beliefs are compared with comissives.

13. For a taxonomy of theories of the objectivity of the law and the role verdicts play in establishing the law, see (Coleman and Leiter 1993). Using their terminology, it is unclear to me if Hart is better read as holding legal facts to be *minimally objective*; or if he instead occupies an intermediate position between *minimal* and *modest* objectivity. Also see (Rosati 2004) for an overview of the problems of objectivity in the law.

14. Borrowing ideas from legal positivism in order to better understand the structure of normativity in desire-based accounts of reason is also pursued in (Hubin 2001).

15. Thanks to a referee for pushing me on this point.

16. Some readers may be concerned that Smith's picture of practical reasoning is anti-Humean (2004a, 'Internal Reasons'), and hence adopting his assumptions here is inconsistent with a Humean picture of practical reason. However, Smith's anti-Humean conclusions depend crucially on his commitment that ideally rational agents will converge in their intrinsic desires (ibid.). If we abandon this commitment, we can accept his points that agents desires will evolve according to their assessments of what would be most rational for them, without the anti-Humean consequences.

17. Thanks to Tristram McPherson for raising this issue.

18. On the other hand, my interpretation of my phenomenology may change it, or it may change the rational weight that different feelings contribute. This theory is compatible with either answer.

19. Thanks to an anonymous referee for pushing me on this point.

20. At least, this is clearly the earlier position. In the later work (2008) it is unclear whether Velleman thinks that intentions are self-fulfilling prophecies. Nonetheless, he continues to hold that it is a drive to maintain the truth of one's self-conception that allows these self-conceptions to be self-fulfilling.

21. A referee questions whether Velleman's theory and the verdictive account are even incompatible. I believe that they are: Velleman argues that facts about which attitudes it would make most (normative) sense to have is ultimately explained in terms of what attitudes would make most explanatory sense to the agent, given her self-conception (2008). Part of the assumption of the verdictive model is that facts about which attitude the agent's psychology would best rationalize can come apart from what we would predict that the psychology is likely to cause, because systematic and predictable irrationality is possible. However, if I am wrong on this point, the following can be read as simply pointing out that certain objections to Velleman do not apply to the verdictive account.

22. The relevant evidential scenario may also be rare. It requires that agents know about indeterminacy of strength of desires (an arcane philosophical thesis) and that they have compelling evidence of such indeterminacy among the desires at issue in a given case.

23. Does this introduce a significant rift between practical and theoretical rationality? It seems only to suggest that epistemically ideal agents will sometimes lack standard capacities for managing their practical attitudes. But this is an instance of a familiar fact. A mysophobe ('germaphobe,' colloquially) may, for example, employ self-deception about the ubiquity of microbes to prevent panic. Perfect epistemic rationality would likely hobble such an agent.

24. Some care is required here. These theories do not address questions of what it is for one desire to be stronger than another. But some presumably could be revised to do so, while others should plausibly be understood as interpreting

the question 'What do I want to do?' not as a question about what one wants most, but as a question about which motive makes up the 'agent's perspective' (as in Watson 1975), or has authority 'to speak for the agent' (as in Bratman 2000).

25. An exception to this general tendency to see self-creation in volitional terms are Frankfurt's discussions of the nature of caring and love (1999a, 'Autonomy'; and 1999b, 'Caring') as volitional *necessities* – concerns that the agent cannot help but have. Nomy Arpaly's discussion of *romantic incompatibilism* (2006, Chap. 2) also addresses this feature of love, but in the service of undermining the normative significance of self-creation.

26. While the theory here is ultimately significantly different from the ones they offer, (Taylor 1976) and (Helm 2002) both offer compelling and insightful cases in which self-reflection could plausibly bring about something between self-discovery and self-invention. Helm especially offers interesting discussion of how this fact would leave us in some respects active and in other respects receptive.

27. Also see (McGeer 1996) for an account of self-knowledge which does not explicitly endorse transparency, but has similarities with certain transparency accounts, especially with (Moran 2001).

28. But see (Byrne 2011b, 'Want'; Ashwell 2013) for ways in which a transparency theorist might account for our ability to know when our desires and our assessments of value come apart. Byrne's story depends crucially on transparent knowledge of our own *intentions*. See (Baker 2015) for criticisms of Byrne's account of transparent knowledge of one's own intentions. Ashwell's story depends on knowledge of what *appears* valuable to us.

29. Thanks to an anonymous referee for bringing this paper to my attention.

30. Thanks to an anonymous referee of an earlier version of this paper for raising this problem.

31. A referee helpfully notes that desires for happiness, or a higher order desire for desire satisfaction, or a desire for rational coherence, could also explain cases in which my beliefs about what I desire affect my choice, without those beliefs playing any sort of verdictive role. I have no disagreement here: the effects of our beliefs about desires on our choices could easily be overdetermined. To repeat, the case for the verdictive role of self-reflective judgments is the fact that judgments about what I want put rational pressure on me to adopt motives consistent with that judgment; this combined with a general disposition toward rational coherence will lead self-reflective judgments to play a verdictive role.

Notes on contributor

Derek Baker is an Associate Professor at Lingnan University, Hong Kong SAR. His work in primarily in the areas of metaethics and moral psychology.

References

Arpaly, Nomy. 2006. *Merit, Meaning, and Human Bondage: An Essay on Free Will*. Princeton: Princeton University Press.

Arpaly, Nomy, and Timothy Schroeder. 2012. *In Praise of Desire*. Oxford: Oxford University Press.

Ashwell, Lauren. 2013. "Deep, dark … or transparent? Knowing our desires." *Philosophical Studies* 165 (1): 245–256. doi:10.1007/s11098-012-9950-3.

Austin, J. L. 1962. *How to Do Things with Words*. Oxford: Oxford University Press.

Baker, Derek. 2015. "Why Transparency Undermines Economy." *Synthese* 192 (9): 3037–3050. doi:10.1007/s11229-015-0700-x.

Berridge, Kent. 2004. "Pleasure, Unfelt Affect and Irrational Desire." In *Feelings and Emotions: The Amsterdam Symposium*, edited by Anthony S. R. Manstead, Nico Frijda, and Agneta Fischer, 243–262. Cambridge: Cambridge University Press.

Berridge, Kent. 2009. "'Liking' and 'wanting' Food Rewards: Brain Substrates and Roles in Eating Disorders." *Physiology and Behavior* 97 (5): 537–550. doi:10.1016/j.physbeh.2009.02.044.

Bratman, Michael. 2000. "Reflection, Planning, and Temporally Extended Agency." *Philosophical Review* 109 (1): 35–61. doi:10.1215/00318108-109-1-35.

Bratman, Michael. 2004. "Three Theories of Self-Governance." *Philosophical Topics* 32 (1/2): 21–46. doi:10.5840/philtopics2004321/25.

Bratman, Michael. 2012. "Constructivism, Agency, and the Problem of Alignment." In *Constructivism in Practical Philosophy*, edited by Jimmy Lenman and Yonatan Shemmer, 81–98. Oxford: Oxford University Press.

Bykvist, Krister. 2009. "No Good Fit: Why the Fitting Attitude Analysis of Value Fails." *Mind* 118 (469): 1–30. doi:10.1093/mind/fzn151.

Byrne, Alex. 2011a. "Knowing That I Am Thinking." In *Self-Knowledge*, edited by Anthony Hatzimoysis, 105–124. Oxford: Oxford University Press.

Byrne, Alex. 2011b. "Knowing What I Want." In *Consciousness of the Self: New Essays*, edited by Jee Loo Liu and John Perry, 165–183. Cambridge: Cambridge University Press.

Chang, Ruth, ed. 1998. *Incommensurability, Incomparability, and Practical Reason*. Cambridge: Harvard University Press.

Chang, Ruth. 2004. "Can Desires Provide Reasons for Action?" In *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*, edited by Jay Wallace, Philip Pettit, Samuel Scheffler and Michael Smith, 56–90. Oxford: Oxford University Press.

Chang, Ruth. 2009. "Voluntarist Reasons and the Sources of Normativity." In *Reasons for Action*, edited by David Sobel and Steven Wall, 243–271. Cambridge: Cambridge University Press.

Coleman, Jules, and Brian Leiter. 1993. "Determinacy, Objectivity, and Authority." *University of Pennsylvania Law Review* 142 (2): 549–637. doi:10.2307/3312546.

Evans, Gareth. 1982. *The Varieties of Reference*. Oxford: Clarendon Press.

Fernández, Jordi. 2007. "Desire and Self-Knowledge." *Australasian Journal of Philosophy* 85 (4): 517–536. doi:10.1080/00048400701676419.

Frankfurt, Harry. 1987. "Identification and Wholeheartedness." In *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*, edited by Ferdinand David Shoeman, 27–45. Cambridge: Cambridge University Press.

Frankfurt, Harry. 1999a. "Autonomy, Necessity, and Love." In *Necessity, Volition, and Love*, chap. 11, 129–141. Cambridge: Cambridge University Press.

Frankfurt, Harry. 1999b. "On Caring." In *Necessity, Volition, and Love*, chap. 14, 259–298. Cambridge: Cambridge University Press.

Hart, H. L. A. 1961. *The Concept of Law*. Oxford: Oxford University Press.

Helm, Bennett W. 2002. "Emotional Reason: How to Deliberate About Value." *American Philosophical Quarterly* 37 (1): 1–22.

Holton, Richard. 2009. *Willing, Wanting, Waiting*. Oxford: Clarendon Press.

Hubin, Donald C. 1999. "What's Special About Humeanism." *Nous* 33 (1): 30–45. doi:10.1111/0029-4624.00141.

Hubin, Donald C. 2001. "The Groundless Normativity of Instrumental Rationality." *The Journal of Philosophy* 98 (9): 445–468. doi:10.2307/2678494.

Hubin, Donald C. 2003. "Desires, Whims, and Values." *The Journal of Ethics* 7 (3): 315–335. doi:10.1023/A:1024691303625.

Humberstone, I. Lloyd. 1990. "Getting, Wanting, Having." *Philosophical Papers* 19 (2): 99–118. doi:10.1080/05568649009506332.

Langton, Rae. 2004. "Intention as Faith." *Royal Institute of Philosophy Supplement* 55: 243–258. doi:10.1017/S1358246100008705.

Lawlor, Krista. 2009. "Knowing What One Wants." *Philosophy and Phenomenological Research* 79 (1): 47–75. doi:10.1111/j.1933-1592.2009.00266.x.

McGeer, Victoria. 1996. "Is 'Self-Knowledge' an Empirical Problem?" *Journal of Philosophy* 93 (10): 483–515. doi:10.2307/2940837.

Moran, Richard. 2001. *Authority and Estrangement: An Essay in Self-Knowledge*. Princeton: Princeton University Press.

Oddie, Graham. 2005. *Value, Reality, and Desire*. Oxford: Clarendon Press.

Rosati, Connie. 2003. "Agency and the Open Question Argument." *Ethics* 113 (3): 490–527. doi:10.1086/345625.

Rosati, Connie. 2004. "Some Puzzles about the Objectivity of the Law." *Law and Philosophy* 23 (3): 273–323. doi:10.1023/B:LAPH.0000014571.10545.6e.

Ross, Jacob. 2009. "How to Be a Cognitivist about Practical Reason." *Oxford Studies in Metaethics, vol. 4*, edited by Russ Shafer-Landau, 243–281. Oxford: Oxford University Press.

Schafer, Karl. 2013. "Perception and the Rational Force of Desire." *Journal of Philosophy* 110 (5): 258–281. doi:10.5840/jphil2013110528.

Schroeder, Mark. 2007. *Slaves of the Passions*. Oxford: Oxford University Press.

Schroeder, Timothy. 2004. *Three Faces of Desire*. Oxford: Oxford University Press.

Schroeder, Timothy. 2010. "Irrational Action and Addiction." In *What Is Addiction?*, edited by Don Ross, Harold Kincaid, David Spurrett, and Peter Collins, 391–408. Cambridge: MIT Press.

Sinhababu, Neil. 2009. "The Humean Theory of Motivation Reformulated and Defended." *Philosophical Review* 118 (4): 465–500. doi:10.1215/00318108-2009-015.

Sinhababu, Neil. 2011. "The Humean Theory of Practical Irrationality." *The Journal of Ethics and Social Philosophy* 6 (1): 1–13.

Sobel, David. 2011. "Parfit's Case against Subjectivism." In *Oxford Studies in Metaethics, volume 6*, edited by Russ Shafer-Landau, 52–78. Oxford: Oxford University Press.

Smith, Michael. 2004a "Internal Reasons." In *Ethics and the A Priori*, chap. 1, 17–42. Cambridge: Cambridge University Press.

Smith, Michael. 2004b. "In Defence of the Moral Problem: A Reply to Brink, Copp, and Sayre-McCord." In *Ethics and the A Priori,* chap. 13, 258–298. Cambridge: Cambridge University Press.

Taylor, Charles. 1976. "Responsibility for the Self." In *The Identities of Persons*, edited by Amelie Oksenberg Rorty. Berkeley: University of California Press.

Tiberius, Valerie. 2000. "Humean Heroism: Value Commitments and the Source of Normativity." *Pacific Philosophical Quarterly* 81 (4): 426–446. doi:10.1111/1468-0114.00113.

Valaris, Markos. 2014. "Self-Knowledge and the Phenomenological Transparency of Belief." *Philosophers' Imprint* 14 (8): 1–17.

Velleman, J. David. 2000. *The Possibility of Practical Reason*. Oxford: Oxford University Press.

Velleman, J. David. 2006. *Self to Self: Selected Essays*. Cambridge: Cambridge University Press.

Velleman, J. David. 2008. "A Theory of Value." *Ethics* 118 (3): 410–436. doi:10.1086/528782.

Watson, Gary. 1975. "Free Agency." *The Journal of Philosophy* 72 (8): 205–220. doi:10.2307/2024703.

Way, Jonathan. 2007. "Self-knowledge and the Limits of Transparency." *Analysis* 67 (3): 223–230. doi:10.1093/analys/67.3.223.

Williams, Bernard. 1981. "Internal and External Reasons." In *Moral Luck*, chap. 8, 101–113. Cambridge: Cambridge University Press.

Wood, Allen. 1990. *Hegel's Ethical Thought*. Cambridge: Cambridge University Press.