

# **Akratic Action under the Guise of the Good**

Eugene Chislenko

Department of Philosophy, Temple University, Philadelphia, PA, USA

Email: [chislenko@temple.edu](mailto:chislenko@temple.edu)

(final version—now published in the *Canadian Journal of Philosophy*)

*Abstract:* Many philosophers have thought that human beings do or pursue only what we see as good. These “guise-of-the-good” views face powerful challenges and counterexamples, such as akratic action, in which we do what we ourselves believe we ought not do. I propose a new way for guise-of-the-good views to address this central counterexample by appealing to conflicting beliefs. I then answer concerns that this appeal is insufficiently explanatory, attributes too much conflict, leaves out an essential asymmetry in action against one’s ‘better’ judgment, attributes systematic error about one’s own beliefs, and is too implausible.

*Keywords:* guise of the good; akrasia; intentional action; intention; belief; Donald Davidson

## 1. The Problem

Some ancient philosophers saw an essential connection between motivation and evaluation. They understood action as a pursuit of something conceived as good. We read in Plato that “Every soul pursues the good and does its utmost for its sake.” Or as Aristotle put it: “Though in any case it is the object of appetite which originates movement, this object may be either the real or the apparent good.” The idea that we act or desire only “under the guise of the good” is a familiar one in Plato and Aristotle, medieval scholasticism, Kant, and, more recently, Davidson, Korsgaard, Raz, and others. Kant called one version of it an “old formula of the schools.”<sup>1</sup>

---

<sup>1</sup> These quotations are taken from Plato (1997), p. 1126 (*The Republic*, 505e); Aristotle (1984), vol. I, p. 689 (*On the Soul*, 433a27-29 in the standard Bekker pagination); and Kant, (1997), p. 5:59 in the standard Akademie pagination. For other classic statements, see the *Protagoras*, *Gorgias*, and *Meno* in Plato (1997); the opening lines of Aristotle (1999); Kant, (1997), p. 5:29 and Kant (1998), p. 4:446-7 in the standard Akademie pagination; Anscombe, (1957), pp. 70ff; Davidson (1980a), discussed below; and Davidson (1980b), esp. pp. 96-102. The full ‘formula’ mentioned by Kant reads: *nihil appetimus, nisi sub ratione boni; nihil aversamus, nisi sub ratione mali*, i.e.: “We desire nothing except under a conception of the good; we avoid nothing except under a conception of the bad” (my translation). Influential rejections of guise-of-the-good views in the contemporary literature include Stocker (1979), Velleman (2000), and Setiya (2007).

There are many motivations for the idea. First, it offers a way to explain what is distinctive of action and agency. Guidance by evaluation has been thought to be essential to what makes a movement *mine*—what makes it attributable to me as a person, rather than the product of mechanistically conceived desires operating in or through me. Second, it offers a unified picture of action and belief, on which both are trying to get things right—belief with respect to the true, action with respect to the good. Third, such a view lets a single explanation account for evaluation and motivation at the same time. If we act under the guise of the good, one can explain why someone does something *by* saying what she saw as good about it. Or if we do not see how to explain a perplexing action, we can start by looking for the unobvious evaluative element. Fourth, it offers a hope for a foundation for ethics. If we all pursue what we see as good, a conclusion about the nature of the good will have an inescapable claim on us. Lastly, a guise-of-the-good view is an extension of ordinary generosity or charity of interpretation, systematized into a view about acting beings. It says that even when we are conflicted, confused, or exhausted, we pursue what we see as good.

These theoretical and practical motivations are not what *this* paper is about. I will not make an argument here about their legitimacy, or about their uniqueness to a guise-of-the-good view. I mention them just to give a sense of why the view is worth considering. Instead, I will defend the view itself, on one of the key issues on which it needs defending.

It can seem obvious that we often do and want what we do *not* see as good, and

---

More recently, there has been a growing sense that these views were on to something after all; see Korsgaard (1996, 2009), Tenenbaum (2007), Boyle and Lavin (2010), and Raz (2010). I take this paper to be in line with this resurgence of the traditional view, though none of its proponents have made the argument I am making here.

sometimes even what we see as bad. We can be malicious, perverse, compulsive, self-destructive, indifferent, or in conflict with our own better judgment. Arguably, some creatures who act are not even old enough, or of the right species, to have or use evaluative concepts. The classical view then seems to paint an impoverished picture, leaving out these widespread and important forms of activity. It seems to describe only, as Velleman (2000, 99) put it, “a particular species of agent, and a particularly bland species of agent, at that.” If this is right, the view cannot capture anything general about personhood or motivation. And clinging to a guise-of-the-good view would be an extension, not of ordinary, reasonable generosity of interpretation, but of a blind optimism that fails to see the person it is trying to be kind to.

I believe that a guise-of-the-good view, properly understood, *can* account for the wide variety of agency while doing justice to the motivations I mentioned. In this paper, I want to propose a new way for the view to handle what is often taken to be the central counterexample: *akrasia*, or acting against one’s own better judgment. Recent defenders have tended to retreat to fairly modest guise-of-the-good views, partly in an attempt to address challenges like *akrasia*.<sup>2</sup> I think these modest views concede too much, and do less justice to the initial motivations for guise-of-the-good views. If I am right, even an especially ambitious version of the view can accommodate *akrasia*, without the need for a retreat. Seeing how an ambitious guise-of-the-good view allows for akratic action can also offer new resources for defending

---

<sup>2</sup> See especially Tenenbaum (2007) and Raz (2010). Other counterexamples, such as cases like that of Buridan’s Ass, can also be thought to call for a retreat from ambitious guise-of-the-good views. For a response to Buridan’s Ass and similar cases that avoids this retreat, see Chislenko (2016).

guise-of-the-good views more generally.<sup>3</sup>

Let us consider one particularly ambitious guise-of-the-good view:

*The Normative Belief Requirement:* If A *x*'s intentionally, A must believe she ought to *x*.

This requirement is especially demanding. A more modest version might require that A believe there is *something* good in *x*-ing; or, less modestly, that she believe *x*-ing is a good thing to do overall; or, less modestly still, that she believe *x*-ing is the *best* thing to do, rather than one good thing among others. All of these views are consistent with A's treating *x*-ing as optional. They can allow that A believes what she is doing to be good, or best, without believing that she ought to do it. Putting the requirement in terms of 'ought' rules out this possibility. It requires A to believe that she *ought* to act as she does. Nor does it insist that *x*-ing must *appear* to her to be good. On this requirement, A must *believe* that she ought to *x*.<sup>4</sup> This requirement is so

---

<sup>3</sup> There are many more guise-of-the-good views than have ever been considered. Some can be summarized in the following schematic template: "An (action / intention / desire) (is / requires / requires a capacity to have) a (belief / judgment / appearance) that the (action / end / outcome) is (what one ought to do or bring about / good / in some way good)." Multiplying the numbers of options within each parenthesis gives us  $3^5 = 243$  logically possible guise-of-the-good views. Even these are only a starting point, leaving room for many further specifications and additions. Because there are so many possible views, I do not claim that the explanation I give in this paper will apply to all of them, though it may be available to a range of more modest views.

<sup>4</sup> I will occasionally use "should" as a synonym for "ought"; I consider a discussion that focuses on judgment rather than belief, and return to differences between conceptions of belief, in §2 below. For

demanding that it is likely to seem implausible. On the other hand, if even this requirement, formulated in terms of ‘ought’ and belief, can be shown to allow the possibility of akratic action, more modest guise-of-the-good views are likely to allow *akrasia* as well. With this in mind, I will focus in what follows on the Normative Belief Requirement. I will argue that there is conceptual room for akratic action, even if this requirement is correct.

Here is the problem. It is hard to deny that we sometimes intentionally do what we believe we ought not do. These actions are often called akratic, or sometimes weak-willed, or “incontinent,” or against one’s own better judgment.<sup>5</sup> For A to *x akratically* is for A to *x intentionally*, while believing she ought not *x*. We believe, for example, that we should not eat dessert, or stay on the computer, or insult our friend. But we do it anyway. This seems to violate the Normative Belief Requirement, which requires normative endorsement, not rejection or indifference. If we act akratically, how can any such requirement be correct? How can evaluation and motivation line up so neatly if they so obviously come apart?

I will answer this question in two parts. In §2, I argue that the Normative Belief Requirement can allow that someone believes she ought not do what she is doing, as long as she *also* believes that she ought to do it. Rather than ruling out the possibility of akratic action directly, it requires akratic agents to have conflicting normative beliefs. I use a discussion of Donald Davidson’s classic account of *akrasia* to show how easy it is to overlook the possibility

---

quick dismissals of the view I go on to defend, see, for example, Bratman (1979), pp. 157 and 171n13; Mele (1987), p. 30; and Buss (1987), p. 22.

<sup>5</sup> These terms may not all be interchangeable. For recent work on the relation between akratic action and “weakness of will” in the ordinary sense, see Mele (2010), May and Holton (2012), and Beebe (2013). I leave this debate aside, avoiding the term “weakness of will” to keep the focus on akratic action.

of conflicting beliefs in an explanation of *akrasia*. I then describe the nature of the conflict, and argue that such conflict is possible in principle. In §3, I turn to defending the more controversial view that *all* akratic actions are cases of conflicting normative beliefs. This view faces several concerns: that it fails to explain the possibility of *akrasia*; that it attributes a degree of conflict that is not possible in a single person; that it leaves out an essential asymmetry in action against one's 'better' judgment; that it attributes widespread error to people about their own beliefs, and that it is simply too implausible. On the contrary, I will argue, appealing to conflicting beliefs is genuinely explanatory; attributes a significant but recognizable degree of conflict that helps explain why akratic action is puzzling; helps to identify various forms of asymmetry; attributes a small but plausible amount of self-ignorance; and leads doubts about plausibility into a more fruitful, systematic investigation of guise-of-the-good views. Even a demanding view like the Normative Belief Requirement can allow for and illuminate akratic action. This suggests that guise-of-the-good views more generally need not be seen as describing only a narrow, boring kind of agent. Instead, they bring out the full extent of the inner conflict which is so characteristic of human life and activity: a conflict not only between evaluation and motivation, but also within evaluation itself.

## **2. The Possibility of Conflicting Beliefs**

It seems clear that we can believe we should not do something, and still do it anyway. Sometimes, we believe we should not watch one more episode, or eat one more slice, but we still go ahead and do it. How can anyone deny that? The key initial point is that the Normative

Belief Requirement makes no such denial. Instead, it attributes a *positive* evaluation—a belief that we ought to act as we do. We should distinguish *lacking* a normative belief, on the one hand, from believing in a prohibition, on the other. The Normative Belief Requirement requires a belief that we ought to act as we do. When we act akratically, we believe we *ought not* act as we do. So strictly speaking, there is no problem about believing we ought not do something and doing it anyway. There is always the possibility of believing that one ought to *and* believing that one ought not. *Akrasia* is possible when we have conflicting normative beliefs.

To see how easy it is to miss this possibility, it helps to turn briefly to Donald Davidson’s classic account of *akrasia*. As we will see, considering conflicting beliefs in turn sheds light on that account, and undermines a long popular understanding of it.

In “How is Weakness of the Will Possible?”, Davidson describes the problem as a conflict between three principles that all “seem self-evident”:

P1. If an agent wants to do x more than y and believes himself free to do either, he will intentionally do x if he does either intentionally.

P2. If an agent judges it better to do x than y, he wants to do x more than to do y.

P3. There are incontinent actions.

P3 simply restates what it seems we already knew: that it is possible to act intentionally against one’s own better judgment. In Davidson’s terms,

In doing x an agent acts incontinently if and only if: (a) the agent does x intentionally; (b) the agent believes there is an alternative action y open to him; and (c) the agent judges that, all things considered, it would be better to do y than to do x.<sup>6</sup>

---

<sup>6</sup> These principles and definition are taken from Davidson (1980a), pp. 22-23. One complication in them is that Davidson, and others following him, define *akrasia* in terms of available alternatives. As



P3 simply tells us that we do sometimes act this way. And yet the existence of such actions seems to fly in the face of “another doctrine that has an air of self-evidence: that, in so far as a person acts intentionally he acts...in the light of some imagined good.” P1 and P2 together reflect that other doctrine. Davidson thinks the conflict cannot be resolved: “No amount of tinkering with P1-P3 will remove the underlying problem.” He then goes on to develop a conception of *akrasia* that attempts to preserve all three principles.

Among Davidson’s three principles, P2 is the principle that most directly connects evaluation with motivation. It has been common to resist guise-of-the-good views by rejecting Davidson’s first two principles, and especially P2. Resisting Davidson’s view, Gary Watson argues that “there is no univocal interpretation of the key phrases of P1 and P2 on which these

---

Tenenbaum (2007), p. 257 puts it, “An akratic agent will think that A is better than B yet pursue B.” Here I temporarily go along with Davidson’s formulation, though I think a non-comparative one is both simpler and more precise. You may, for example, believe you should not smoke, but not have thought about how else you would spend your next five minutes. If someone does not have a particular better alternative in mind, but believes she should not do B, her doing (or ‘pursuing’) B is still akratic in the central sense of being against her better judgment, or something she believes she ought not do.

If simply refraining—not doing B, not smoking—counts as an alternative action, the comparative definition and the non-comparative one may apply to at least close to the same cases. But there would still be two differences. First, the comparative definition has the person thinking in terms of what is better rather than what she ought to do. Second, even without that difference, the comparative definition has the person believing she should instead refrain, rather than, more directly, that she should not do B. These differences can be distracting. Rather than asking how significant the differences are, I use the simpler formulation: we act akratically when we intentionally do what we believe we ought not do.

principles turn out to be true, or even very plausible.” Robert Audi, noting that “the most common way out is to deny P2,” himself argues that “It would be at least unreasonable to maintain P2 or any equally strong internalist principle.” Alfred Mele argues that “(P2) is false. The connection between unconditional judgments and the balance of an agent’s motivation is more complex than Davidson thinks.”<sup>7</sup> Reactions like these tend to accept Davidson’s view that there is a deep conflict between his P1-P2 and the existence of akratic action, which no amount of ‘tinkering’ will remove.

But Davidson’s P2 is a misleading way of presenting a guise-of-the-good view, and rejecting it is a misguided way of resisting these views. This resistance is misguided, because a guise-of-the-good view should itself *reject* P2. P2 says: “If an agent judges it better to do x than y, he wants to do x more than to do y.” By the same token, if an agent judges it better to do y than x, she wants to do y more than to do x. So if her judgments conflict—if she judges both that it is better to do x than to do y, and that it is better to do y than to do x—then she must both want to do x more than to do y, and want to do y more than to do x. That is not just irrational; it is impossible. The strength of one desire cannot be both greater and less than the strength of another.<sup>8</sup> So if P2 is right, it is impossible both to judge it better to do x than y, and to judge the

---

<sup>7</sup> See Watson (1977), p. 321; Audi (1979), pp. 189-90; and Mele (1983), p. 366. For a helpful classification of ways of resisting P2, see Walker (1989), pp. 668-9. These are typically ways of resisting guise-of-the-good views, whereas the argument I go on to give is meant to bolster defenses of those views.

<sup>8</sup> Here I assume that “wants more” is more naturally read in terms of desire strength, rather than as synonymous with “prefers.” On the latter reading, the problem is somewhat different. Many people believe that preferences can be intransitive; and it is natural to believe that a single individual can have

contrary. P2 is committed to denying the possibility of conflicting judgments. In other words, rather than simply expressing a guise-of-the-good view, P2 combines this expression with a ruling out of all conflicting judgments of a certain kind.<sup>9</sup>

This combination threatens to confuse the issue. Both defenders and deniers of guise-of-the-good views may accept that our evaluative judgments can conflict with each other in important ways. I will now argue, against P2, that such conflict is indeed possible, and that a conception of akratic action should allow it. Doing so will also help to make the conflict recognizable in typical cases of *akrasia*.

Returning to the ‘ought’ language of the Normative Belief Requirement, we can formulate a principle that constrains conceptions of *akrasia*:

---

conflicting preferences as well. But P1, with “wants more” read as “prefers,” rules out conflicting preferences in cases of intentional action. Guise-of-the-good views should then reject P1. Alternatively, if we deny the possibility of conflicting preferences, P2 still rules out the possibility of conflicting judgments in the way I discuss in the text.

<sup>9</sup> Davidson is not the only one to rule out the possibility of such conflict. Mele (1983), pp. 357-8 argues that since we can intend against our own better judgment, intention cannot itself be such a judgment—a conclusion that follows only if we rule out the possibility of conflicting judgments. Bratman (1979), p. 157, gives an example of Sam, discussed below, who drinks after deciding it would be best not to. As Bratman sees him, “Sam surely does not also conclude that it would be best to drink; though guilty of some form of irrationality, Sam is not guilty of such blatant inconsistency.” Bratman then adds, in a somewhat different context: “I assume that the agent does not hold logically inconsistent views”(p. 171n13). On the view I am developing, such conflict—though perhaps not properly called “logical”—is central to akratic action, and its possibility undermines any theory that depends on ignoring it.

*The Conflict Constraint:* A person can both believe she ought to do something, and believe she ought not do it.

The Conflict Constraint does not claim that conflicting beliefs of this kind are present in all akratic action. It merely paves the way for this more ambitious claim, to which I return in §3. But since its significance is easily overlooked, and since it offers crucial resources for defending the Normative Belief Requirement, we can begin with a brief consideration of the nature and justification of the Conflict Constraint.

The Conflict Constraint insists that we allow the possibility of conflicting normative beliefs: beliefs that together require both performing an action, and not performing it. It is a constraint, because accepting it leads us to reject views, including conceptions of *akrasia*, that deny the possibility of this kind of conflict. It is a “Conflict Constraint,” rather than a “Contradiction Constraint,” because what it requires us to allow is not contradictory beliefs, of the form “p” and “not-p”, but normative conflict, of the form “I ought to *x*” and “I ought not *x*.” This conflict can easily lead to contradiction; we naturally infer from “I ought to *x*” that “It is not true that I ought not *x*.” But it is nevertheless distinct from contradiction. We may not always infer the contradiction, even if it is logically entailed. Moreover, the conflict is in one way even starker than contradiction. The contradiction “I ought not *x*” and “It is not true that I ought not *x*” both forbids and permits an action. The conflict “I ought not *x*” and “I ought to *x*” both forbids and *requires* an action. This is a distinctive and particularly striking conflict.

There are two available senses of ‘conflict in belief’ here: conflict *within* a belief, and conflict *between* beliefs. A single belief can show normative conflict, if the belief is of the form: “I ought to *x* and I ought not *x*.” For normative conflict *between* beliefs, it is enough to have a

belief of the form: “I ought to x,” and a belief of the form “I ought not x.” These beliefs are in normative conflict with each other, even without a belief in the conjunction: “I ought to x and I ought not x.”

The Conflict Constraint requires only that a conception of *akrasia* allow the possibility of conflict *between* beliefs. But distinguishing these two kinds of normative conflict also allows us to distinguish two arguments for the Conflict Constraint.

One argument begins with examples of apparent within-belief conflict. A sharp conflict of this kind can be seen in Thomas Nagel’s essay “War and Massacre.” Nagel describes some especially stark moral dilemmas related to war, such as the question of whether to torture a terrorist. He suggests that situations like these can put us into what he calls a “moral blind alley,” in which, for example, torturing the terrorist would be wrong, and not torturing the terrorist would be wrong. He writes:

The idea of a moral blind alley is a perfectly intelligible one. It is possible to get into such a situation by one's own fault, and people do it all the time. If, for example, one makes two incompatible promises or commitments—becomes engaged to two people, for example—then there is no course one can take which is not wrong, for one must break one's promise to at least one of them. Making a clean breast of the whole thing will not be enough to remove one's reprehensibility. The existence of such cases is not morally disturbing, however, because we feel that the situation was not unavoidable: one had to do something wrong in the first place to get into it. But what if the world itself, or someone else's actions, could face a previously innocent person with a choice between morally abominable courses of action, and leave him no

way to escape with his honor? Our intuitions rebel at the idea, for we feel that the constructibility of such a case must show a contradiction in our moral views. But it is not in itself a contradiction to say that someone can do X or not do X, and that for him to take either course would be wrong. It merely contradicts the supposition that *ought* implies *can*— since presumably one ought to refrain from what is wrong, and in such a case it is impossible to do so.<sup>10</sup>

Assuming that Nagel is describing his own beliefs, this passage offers an example of within-belief conflict. Nagel suggests that it is intelligible to think that one ought not do X *and* ought not *not* do X. He thinks people get into such situations “all the time.” And so, it seems, Nagel believes that in some situations, for some X, we ought to X and ought not X. About particular situations, he may have beliefs like: “I ought to keep my promise and I ought not keep it,” or: “I ought to torture the terrorist and I ought not torture him.”

---

<sup>10</sup> Nagel (1972), pp. 143-4. Nagel’s self-attribution of within-belief conflict is analogous to self-attributions of within-belief contradiction by some dialetheists, who believe that some statements are both true and false. Priest (2006), pp. 96-7 writes:

There are many cases where people consciously believe an explicit contradiction (and with no real doubt)...I, for example, believe that the Russell set is both a member of itself and not a member of itself. I do not deny that it was difficult to convince myself of this, that is, to get myself to believe it. It seemed, after all, so unlikely. But many arguments, most of which appear in this book, convinced me of it.

Like Nagel, Priest seems to believe what he says he believes, although his beliefs may puzzle us.

Here it is important to be clear about the sense of ‘ought’ that is at issue. On one interpretation of these cases, every course of action is *morally* wrong, or one that morality requires us to avoid; but like the ‘ought’ of etiquette, this moral ‘ought’ need not be overriding. There is room for disagreement about whether what we ought morally to do is also what we ought to do, full stop. But for our purposes, there is no need to insist that the moral ‘ought’ is overriding. Indeed, we need not insist that Nagel holds this view. It is enough that *some* people can believe themselves to be in “normative blind alleys,” in which every possible course of action is one they normatively or deliberately ought not take. Such people can see moral ‘oughts’ as overriding, or they can ignore moral ‘oughts’ entirely. In either case, they might believe that they ought—deliberatively, normatively, or “full stop”—to torture a terrorist and ought not torture him. To endorse the possibility of a normative blind alley is to endorse the possibility that a belief with a within-belief normative conflict can be true. Although we may not entirely understand what it is like to have such a conflict, seeing someone insist, and argue, that she is in one may lead us to admit that it is possible. The very existence of someone who avows such views is evidence that they can be believed.

The possibility of within-belief conflict plausibly entails the possibility of between-belief conflict. It is natural to think that belief distributes over conjuncts: that someone who believes “p and q” already believes “p” and believes “q.” And even if belief does not distribute over conjuncts, someone is normally able to infer each conjunct from the conjunction. Someone who believes herself to be in a normative blind alley believes in, or at any rate can infer, each of its components. For someone who believes “I should torture this terrorist and I should not torture him,” both “I should torture this terrorist” and “I should not torture him” seem to be, at the very least, easy conclusions to draw from the conjunctive belief. According to this line of

thought, within-belief conflict is possible; if within-belief conflict is possible, between-belief conflict is possible; so between-belief conflict is possible. We can call this the *Argument from Within-Belief Conflict*.

The inference in the other direction can be more difficult. An anorexic who believes “I should lose weight” and believes “I should not lose weight” may find it far from trivial to form the conjunctive belief “I should lose weight and I should not lose weight.”<sup>11</sup> Within-belief normative conflict can be harder to maintain than between-belief conflict, and on a more restrictive view, it may not be possible. Some may think that the notion of a normative blind alley makes so little sense that no one, including Nagel, could understand it well enough to ever believe herself to be in one. If this is true, we should not grant that someone who claims to be in a normative blind alley is accurately reporting her own beliefs.

For those who accept the possibility of within-belief conflict, the argument from within-belief conflict can show the possibility of conflicting beliefs. But it is also worth noticing that within-belief conflict is not essential. We can, instead, attribute between-belief conflict directly, by an *Argument from Belief Attribution*. I will not try to consider every view of the nature of belief here. But when it comes to attributing a belief to someone, we tend to look to a cluster of typical marks of belief. We look for sensitivity to evidence, to apparent evidence; recall in relevant circumstances; felt conviction; a disposition to report the belief to others; and use of a belief in further reasoning. In, for example, conflict about weight loss, each of two beliefs can manifest these characteristic features. Each may be sensitive to apparent evidence—one to an apparent horrible overabundance of fat seen in the mirror, the other to a doctor’s warnings of

---

<sup>11</sup> For resistance to attributing some closely related kinds of conflicting beliefs in cases such as anorexia, see Adler (2002).



malnutrition. Each may be reported sincerely in various contexts and to different people, recalled often and with conviction, and so on. And it is not always easy to settle on a single belief attribution statistically, based on a greater frequency or duration of dominance of one belief over another. Someone's thoughts and behavior can be in such stark conflict that it can be described only by attributing two conflicting normative beliefs to him. This can be true even though it can be difficult for him to combine these beliefs into one. "I should lose weight and I should not lose weight" is much harder to support with evidence, to hold with any conviction, to sincerely report, and so on. But between-belief conflict can be easier to maintain, and more resistant to resolution. So although accepting the possibility of within-belief conflict is one way to accept the possibility of between-belief conflict, it is not the only way. The usual characteristics by which we recognize belief can themselves conflict, and can call for the attribution of two conflicting beliefs.

The Argument from Belief Attribution will go somewhat differently on different views about proper belief attribution, with their different underlying views about the nature of belief. Williamson (2000, p.99) suggests that, "Intuitively, one believes  $p$  outright when one is willing to use  $p$  as a premise in practical reasoning." One might instead hold, with De Sousa (1971, p. 64), that " $Bp$  [belief that  $p$ ] is a *disposition to assent*"—or, with Cohen (1992, p.5), that "Belief is a disposition to feel." Assent, reasoning, and feelings are all notoriously capable of conflict. Indeed, they are all capable of simultaneous conflict; although there are some cases of a rapid and repeated change of mind, in other cases we can find complex and conflicting patterns of assent, thought, and feeling in a single person at the same time. Which conflict is crucial for the attribution of conflict between beliefs will depend on which marks of belief we take as central for belief's proper attribution. Not all views of belief, of course, will allow that such conflict is

possible even in principle. But many do, and even without a developed view about the nature of belief, it is plausible to think that the mind can sometimes be conflicted in these ways. Both the Argument from Within-Belief Conflict and the Argument from Belief Attribution give us independent reason to accept the Conflict Constraint.

The Conflict Constraint can be accepted regardless what one thinks of a guise-of-the-good view of action. But it also helps make a guise-of-the-good view believable. Respecting the constraint helps the view allow an important kind of inconsistency, and so helps it avoid limiting itself to describing “a particularly bland species of agent,” as Velleman puts it. And yet Davidson’s formulation of his principles in “How is Weakness of the Will Possible?” can lead us to unwittingly accept a view that violates the constraint. His P2 is close to a blanket denial of the possibility of conflicting beliefs about what one ought to do. Once we let go of *that* denial, a guise-of-the-good view becomes defensible.

### **3. The Conflicting Belief View**

The problem of akratic action is: how can one believe that one ought not do something, and still intentionally do it? In one way, this problem has an easy answer, even if we accept the Normative Belief Requirement. We intentionally do it—take an extra helping of ice cream, for example—because we do believe we ought to, and we act on that belief. Believing we ought to do something does not stop us from also believing we ought not. Unfortunately, believing we ought not does not stop us from believing we ought to, either. We can act akratically, because our beliefs can conflict in this way.

The basic idea, once again, is that *akrasia* itself involves belief that conflicts with our ‘better judgment’. Reaching for dessert has an evaluative structure.<sup>12</sup> Typically, the dessert suddenly strikes us as delicious and as something we ought to have, even if we disapprove of our own motivation and believe we ought to skip dessert. The possibility of conflicting beliefs takes some of the bite out of the counterexample. It prevents *akrasia* from providing a direct refutation of the Normative Belief Requirement.

The possibility of conflict makes accounting for *akrasia* difficult instead of impossible. But *akrasia* still presents a challenge. The Normative Belief Requirement can allow that we can do what we believe we ought not do. But it does deny that we can do what we do *not* believe we *ought* to. It then must hold that *whenever* we intentionally do what we believe we ought not do, we must *also* believe that we ought. It has to hold, in other words, that *akrasia* always involves conflicting normative beliefs. We can call this view

*The Conflicting Belief View.* Akratic action requires conflicting normative beliefs.

---

<sup>12</sup> Aristotle (1999, p. 1143b18; cf. 1149b14-18) famously makes a brief mention of the possibility that “The incontinent or base person will use rational calculation.” The view I propose is not that all akratic action is calculating; it leaves open the possibility that an akratic action might involve normative belief without any means-end reasoning or other calculation. But since acting on an evaluative belief usually does involve some inference about means, I am, in effect, treating calculating *akrasia* as the paradigm case.

Though this view is rarely taken seriously<sup>13</sup>, I believe it is defensible. But it does take some defending. We might wonder: can this view help us understand akratic action? Is so much conflict possible? Does the view explain the sense in which an akratic action defies one's "better" judgment? When someone denies that she is conflicted in this way, on what grounds can we say she is mistaken about her own beliefs? Is the view really plausible? These questions are all worth answering. To illustrate and defend the view, I will take each of them in turn.

### 3.1 Explanatory Failure

Pointing to conflicting beliefs, one might think, could not in principle provide an account of *akrasia*. It would not explain *how* such conflict is possible; so how would it explain how action can be akratic?

I think this objection is right in wanting further explanation, but wrong to want it here. The Conflicting Belief View is not meant as a conception of akratic action. It is a particular thought about akratic action, which follows from the Normative Belief Requirement and can seem to be a problematic consequence of it. The Conflicting Belief View treats akratic action as one species of a broader genus: conflict between beliefs about what one ought to do. Its explanatory ambitions are limited. My goal is not to explain everything one wants to know about akratic action, but to defend a necessary condition on it. The challenge is to explain how, given that we act akratically, the Normative Belief Requirement could be true.

---

<sup>13</sup> More precisely: the view is rarely taken seriously by modern writers. For discussion of earlier thinkers who endorsed closely related views, see Gosling (1990), chapter 5, and Saarinen (2011), esp. pp. 36-42. On modern dismissals of the view, see note 4 above.

On the other hand, if one is wondering how an action can be both akratic and under the guise of the good, the Conflicting Belief View offers part of the answer. It points to an underlying mechanism: the operation of conflicting normative beliefs. I have argued that this conflict is a phenomenon whose existence we should accept on independent grounds. We might rightly find it puzzling, and seek a better understanding of it. If we are also puzzled about the possibility of akratic action, the former puzzlement can help explain the latter. We act akratically because our normative beliefs conflict, and akratic action is possible because that conflict is. This, I think, is genuinely explanatory.

Most importantly, the explanation is enough to address the problem *akrasia* poses for the Normative Belief Requirement. The problem was that the requirement seems to rule out the possibility of *akrasia*. The solution is that one can do what one believes one ought not do, while also believing one ought to do it. There is good reason to agree that such conflict between beliefs is possible, though we may not yet fully understand how. Noticing the possibility of such conflict, or accepting the Conflict Constraint, is a significant development. The constraint is minimal enough to accept, powerful enough to dispel doubts about *akrasia*, and interesting enough to get us thinking about how it can be true.

### 3.2 *Conflicted Agency*

We often act akratically. If we were riddled with conflicting beliefs every time, would we be rational beings at all?

A concern about widespread conflict in normative belief is partly a concern about the possibility of *any* conflict. We lose our picture of fully unified agency as soon as we accept the

Conflict Constraint. Beyond this point, increasing disunity can also be increasingly disconcerting. As Davidson (1982, p. 303) puts it, “It is a matter of degree. We have no trouble understanding small perturbations against a background with which we are largely in sympathy, but large deviations from reality or consistency begin to undermine our ability to describe and explain what is going on in mental terms.” The deviations may look somewhat larger when we think of *akrasia* as involving conflicting normative beliefs. But the difference is not in allowing that we do have such conflicting beliefs; any view that satisfies the Conflict Constraint does that. Any plausible view of the mind will have to allow some conflict between beliefs. Nor is the difference in the number of akratic cases, which are already natural to describe as involving some kind of disunity. The difference is in seeing a conflict between beliefs more often than other views do. This does sharpen the disunity, but there is so far no reason to think that it makes the difference between someone who is recognizably an agent and something that is not. On the contrary: once we reject a ‘bland’ conception of agency as perfectly consistent, we can more easily recognize that a striking degree of conflict is a part of life, and, more to the point here, part of agency.<sup>14</sup>

Here, I think bringing out the extent of the disunity is an *advantage* of explaining *akrasia* in terms of conflicting beliefs. If we accept the possibility of conflicting normative beliefs in general, and start to see them here, we have helped to explain why akratic action is so puzzling in the first place. It is puzzling, at least in part in the same way that conflicting beliefs are puzzling. It is puzzling that one can be so disunified, and still be a single person. This disunity is rightly puzzling, although the puzzle need not lead us to deny the phenomenon.

---

<sup>14</sup> For a helpful discussion of the extent of inner conflict, see the ‘mosaic’ conception in Sripada (2016), pp. 1225-1230.

### 3.3 *The Asymmetry Problem*

We often describe *akrasia* as action against one’s “better” judgment. But as I have described it, it seems to include intentional action against *any* judgment—or, to put it differently, *any* intentional action in a case of conflicting beliefs about what we ought to do. If we believe we ought to go to the beach, and believe we ought not go, anything we do will conflict with something we believe. To put it yet another way: neither belief seems in any way singled out as the “better” one. Both are treated as on the same footing. Understanding *akrasia* in terms of conflicting beliefs can then seem to leave out an essential *asymmetry* in an akratic’s relations to her ‘better’ belief and her conflicting belief that she ought to act as she does. We can call this the Asymmetry Problem.

As we saw in §3.1, the Conflicting Belief View is not meant as a complete account of akratic action. It aims to describe a necessary condition, rather than a full set of central features of *akrasia*. Since it can leave some features of akratic action without discussion, it does not clearly face an Asymmetry Problem. Nevertheless, it is worth asking whether the Conflicting Belief View is consistent with a satisfying response to the Asymmetry Problem, and whether it offers any resources for responding to the problem.

By itself, pointing to conflicting beliefs says little about a symmetry or asymmetry. It does rule out one obvious possible asymmetry: that one potential action is believed to be what we ought to do, and another is not. But otherwise it leaves the issue open to the other features we might have in mind when we talk about a ‘better’ judgment. We often speak of the ‘better’ judgment as a more considered or “all things considered” judgment, which takes into account

the full range of relevant considerations. I might see that, overall, I ought to stop eating; when I take a dessert anyway, my reasoning is likely to take less into account. We may also have in mind that a ‘better’ judgment is the one that we, on reflection, more strongly endorse or identify with. We may also sometimes mean the judgment that is more reasonable, or more likely to be correct, or actually better. Though the range of evidence taken into consideration is *one* central dimension of asymmetry, it is worth noticing that the Conflicting Belief View can allow for any of these other asymmetries.<sup>15</sup>

In practice, the degree of ‘asymmetry’ in any of these senses can vary. Consider James. James is opposed to same-sex relationships. His considered belief is that we should “do what’s good for our species, and not go against nature.” He is also vegetarian, but has not thought much about why, and is taken aback when the connection to his other views is pointed out to him. Because he believes that eating meat is natural and good for our species, he decides to give up vegetarianism. But that evening, he cannot bring himself to try meat. He has decided that he ought to eat meat, but he akratically refrains.

---

<sup>15</sup> On all-things-considered judgment, see Davidson, (1980a, 1982). Though I leave aside the irrationality of *akrasia*, it is worth noting that the Conflicting Belief View offers a straightforward description of one way in which akratic action is irrational. On the other hand, the view makes no claim to describe *akrasia*’s main or only irrationality, or to capture its distinctively practical irrationality. Allowing the various asymmetries mentioned in the text also makes the Conflicting Belief View consistent with the view that akratic action can be more rational than not performing the action would be. On rational *akrasia*, see McIntyre (1990) and Arpaly (2000). For a recent attempt to explain the irrationality of *akrasia* without appeal to a guise-of-the-good view, see Setiya (2016).



How should we understand his case? Perhaps the first thing to say is that it is complicated. His stated speciesist belief may itself be akratic. He might have a deeper belief in the sanctity of life, and in the importance of letting every living creature live and flourish without interference, which comes out vividly in many other contexts and conflicts with disregard for animal life. In that case his stated principle would be more local than he thinks, and may be little more than a rationalization of his homophobic feelings. We do not know exactly how much he has considered or where his deeper commitments lie. But if I am right, imagining James as genuinely akratic at his vegetarian dinner is imagining him with two conflicting beliefs—that he ought to eat meat, and that he ought not—to which his relation is probably asymmetrical in various ways. The relevant beliefs can vary in their responsiveness to evidence, their reasonableness, depth or strength of identification, and other features.

The Conflicting Belief View can allow for any of these various kinds of asymmetry. Moreover, its focus on belief encourages us to distinguish them, and to explain *akrasia*'s characteristic asymmetry in terms of them. This explanation offers an especially direct answer to the question: "Better than what?" On the Conflicting Belief View, we can say, akratic action accords with a second belief—one that is, at least typically, less responsive to reasons, less firmly held, or "worse" in one of the other ways in which beliefs are thought to be "worse."<sup>16</sup>

---

<sup>16</sup> If James acts on a strongly held or better considered belief, and against a weakly held belief, we might hesitate to call his actions "akratic." We may then want to redefine *akrasia* as action against one's *better* judgment, in some specified sense of 'better'. But this is a hesitation about the attribution or demarcation of *akrasia*, not about the presence of conflicting beliefs. It offers no objection to the Conflicting Belief View. Similarly, if we have doubts about using the word 'better' in the first place, these are doubts about

### 3.4 The Error Attribution Problem

According to Aristotle (1984, vol. I, p. 689), what “originates movement” is “either the real or the apparent good.” But according to *us*, what we ourselves do is often not good, and not what we ought to do. *Akrasia* seems puzzling largely because, when we act akratically, it seems clear to each of us that we do what we do *not* believe we ought to. It then seems that the Normative Belief Requirement flies in the face of our own experience. The Requirement seems forced to attribute widespread, systematic error to people about their own beliefs. I call this the Error Attribution Problem.

Does pointing to a conflicting belief avoid the problem? People seem to often experience themselves intentionally doing what they clearly *do not believe* they ought to. People even say so. They say: “I don’t think I should be doing this.” The Conflicting Belief View says they *do* think they should be doing it. Why should we believe the view, and not the person?

As a partial response, it helps to notice that the surface grammar of these self-reports is often misleading. Many of them are examples of a familiar syntactic phenomenon known as “neg-raising,” in which a negation is “raised” from an embedded clause into the main clause of a sentence.<sup>17</sup> “I don’t believe she’s coming” typically means: “I believe she’s not coming.” “I didn’t think that movie was very good” is a way of saying “I thought that movie was not very good.” “I don’t think he’s being a good friend” does not usually report an *absence* of belief. It

---

the thoughts that give rise to the Asymmetry Problem, rather than doubts about the Conflicting Belief View.

<sup>17</sup> For a recent systematic overview, see Collins and Postal (2014).

reports a belief. Normally, the speaker *does* believe that the “he” in the statement is *not* being a good friend, and uses these words to report that belief. Similarly, “I don’t think I should be doing this” often means “I think I shouldn’t be doing this.”<sup>18</sup> As before, the prohibitive belief expressed in the latter assertion is consistent with the Conflicting Belief View. The requirement only holds that the akratic person must *also* have a contrary belief. The Conflicting Belief, in other words, attributes fewer mistakes than it may seem to. Neg-raising makes the Error Attribution Problem seem worse than it is.

We can imagine someone saying, “Trust me, I don’t believe at all that I should be doing this.” It is a point in favor of the Conflicting Belief View that such assertions are rare. As Aristotle (1999, p. 1146b36) put it, such evaluative clarity without a practical change of mind “seems extraordinary.” Even when someone does say this, it can often be insincere. It can also be a grammatically misleading expression of strong disapproval, rather than lack of approval. In other cases, the person might be right, because the action is not intentional; it may be a bare reflex, or the involuntary cursing of someone with Tourette’s syndrome, or a more complex compulsive movement, as in alien hand syndrome, of which the person is an alienated observer. Saying, “Trust me, I don’t believe at all that I should be doing this” *can* be a lie, or a joke, or an expression of self-blame, or it can be straightforwardly true and unproblematic because the action is not intentional.

I think we should accept that it might not be any of these. In an example of Michael Bratman’s (1979, p. 156), a man named Sam says: “I don’t think it would be best to drink. Do

---

<sup>18</sup> Gregory (2017) discusses a related ambiguity in “I don’t want to...” in expressions of desire, warning that it can mislead theoretical discussions of desire and, in particular, “lead us...to understate the relationship between normative judgment and desire”(p. 247).

you think I'm stupid enough to think that, given how strong my reasons for abstaining are? I think it would be best to abstain. Still, I'm drinking." Sam explicitly denies a neg-raising interpretation of his self-report, by insisting that he is not stupid enough to have the belief he denies having. He also seems sincere; there is no need to imagine him as lying or joking. The Conflicting Belief View may be forced to see him as mistaken about his beliefs. I will soon turn to the thought that the attribution is implausible in this particular case. But we can first finish considering whether the Conflicting Belief View attributes a troublingly widespread, systematic error to people about their own beliefs.

The Conflicting Belief View is threatened if it is forced to see people as systematically mistaken about their own beliefs. It does not have to say that we are *never* mistaken in this way. Such infallibility is unlikely, and a view that denies it is not at a disadvantage. So I think we can and should say that, in some special cases, we can be mistaken about what we believe.

Here, acknowledging the possibility of conflicting beliefs helps in two ways. First, it helps us see that the Conflicting Belief View attributes error not in ascribing a belief to oneself, but only in *denying* that we have a belief. If we seem to ourselves to obviously have some normative belief, the view will never disagree with us. This makes the range of attributed error smaller, by limiting the attributed error to self-ignorance, or error in denial.

Second, such mistakes are especially likely in the presence of conflicting beliefs, either of which can make the other easier to miss or deny.<sup>19</sup> Some of us may be assuming when we act akratically that, since we believe we ought not do something, we do not believe that we ought to do it. Like Sam, we might be unwilling or unable to consider ourselves "stupid enough"

---

<sup>19</sup> Enoch (2011, pp. 228-9) makes a similar point in the case of contradictory beliefs about reasons for acting.

to maintain a conflicting belief. Even those of us who think conflicting beliefs are possible tend to ignore the possibility in our own activity. Pointing to the possibility of conflict offers an explanation of our mistake.<sup>20</sup>

In sum, the Error Attribution Problem is powerful when the attributed error is extremely widespread. The Conflicting Belief View attributes error in a small range of special cases in which we deny having a belief. This, I think, properly attributes to us a recognizable degree of self-ignorance.<sup>21</sup> We are neither systematically mistaken, nor always aware of what we ourselves believe.

### 3.5 *Intuitive Plausibility*

Even if so much conflict is possible, is it really plausible to think that *all akrasia* involves conflicting normative beliefs?

When one doubts that *akrasia* must involve conflicting normative beliefs, the alternative must be that we can be akratic without having both of those conflicting beliefs. At least one of the two conflicting beliefs must then be lacking. The doubt does not imagine the ‘better’ or prohibitive belief to be lacking. That belief is essential to the phenomenon of acting against

---

<sup>20</sup> For a response to the error attribution problem more generally, not limited to akratic action, see Raz (2010), pp. 112-4. I avoid Raz’s view that the normative concepts in question are technical, and “not to be confused with the concepts that are normally expressed by ordinary use of these terms” (p. 114), as well as his resistance to the view that “agents capable of intentional action must have the concepts used in stating the [guise-of-the-good] Thesis” (p. 114).

<sup>21</sup> Perhaps it even attributes too little; see Schwitzgebel (2012).

one's better judgment, or acting as we believe we ought not act. So the doubt must be about the 'worse' belief. One doubts whether the person must believe she ought to act as she does.

This doubt is not specific to *akrasia*. To be specific to *akrasia*, its guiding idea would have to be that an action may not be motivated or accompanied by a normative belief, *when the person has a conflicting belief*. Why would this conflicting belief make the difference? The thought may be that one cannot believe one ought to perform an action that one believes one ought not perform. But to think this is to violate the Conflict Constraint. It is to rule out the possibility of conflicting normative beliefs.

Still, the doubt might apply to particular cases of akratic action. As I ate that extra ice cream after lunch today, did I really believe I ought to do it? Is it plausible to insist that Sam believes he should drink, even while he insists that he does not believe it? One can find the attribution of the belief intuitively implausible in many particular cases. This is a lingering doubt about the Conflicting Belief View that none of the answers to the previous objections can fully resolve.

I think they cannot resolve the doubt, precisely because the doubt is not specific to *akrasia*. If the objection does not essentially depend on the presence of a 'better' belief, in ways that violate the Conflict Constraint, it is not an objection to the Normative Belief Requirement's view of *akrasia*. It is a general doubt about the requirement. Rather than threatening the account of akratic action, or rendering *akrasia* distinctively problematic, it expresses skepticism about the requirement as a whole. This skepticism can be extended more generally to the guise-of-the-good views with which we started, though it may arise more intensely in the context of the especially ambitious Normative Belief Requirement. It is proper for a doubt to linger at this point, perhaps in a generalized form, perhaps in the context of many particular cases, both

akratic and non-akratic. The way to address it is to take on the larger issue of which this paper treats one smaller part: to think through the motivations and difficulties of guise-of-the-good views, and see whether there is a substantive and defensible one. Far from offering a knockdown counterexample, akratic action leads instead to a more systematic consideration of the motivations and weaknesses of guise-of-the-good views. The examples are inconclusive. We must return to thinking about what it takes for a movement to be recognizable as intentional, or as mine, in the way paradigmatic human actions are.<sup>22</sup> There is no distinctive problem about *akrasia* here, even for a view as ambitious as the Normative Belief Requirement.

This conclusion can seem too quick. If the Normative Belief Requirement is true, one might think, normative beliefs are the conclusions of practical reasoning. They play a key role in intentional action in part by bringing deliberation to a close. But on the Conflicting Belief View, it seems, normative beliefs may not bring deliberation to a close. Conflicting beliefs present a dire need for further reasoning to resolve the conflict. And if both conflicting beliefs are present, what explains why someone acts on one of them? By raising these difficulties, akratic action seems to present a distinctive challenge to the plausibility of the Normative Belief Requirement after all. The Conflicting Belief View can make it harder to see how normative belief brings deliberation to a close, or plays a role in the explanation of intentional action.

These concerns are worth addressing. We can take explanation of action first. The evaluation in cases of akratic action is a belief with a particular content: “I ought to *x*,” or “I ought not *x*.” But as we saw in considering the Asymmetry Problem, beliefs also have other

---

<sup>22</sup> Here I follow Setiya (2007) and Boyle and Lavin (2010) in their conclusion of a ‘stalemate’ with respect to particular examples—though I doubt they would agree that there is a stalemate with respect to views as ambitious as the Normative Belief Requirement.

important features. They vary in strength, responsiveness to evidence, reasonableness, depth or strength of identification, and in other ways. An explanation of akratic overeating can be that, for example, the belief ‘I ought to take an extra helping’ was at that moment more strongly held, and the apparent reasons for it more salient. Moreover, even if we insist that the mere presence of a belief should explain an action—that I ate because I believed I ought to eat—we can still say that the belief explains the action, other things being equal. It is just that, in cases of conflicting beliefs, other things are not equal. A powerful opposing force presents an obstacle to the normal operations of our normative beliefs. In these special cases, the explanation of an action must consider the interaction between the two beliefs in more detail. Normative belief then still plays a key role in the explanation of intentional action—but the belief’s various features, as well as other factors besides the belief that one ought to act, play a role as well.

Moreover, in the relevant sense, even conflicting normative beliefs can bring deliberation to a close. As we saw in considering the Conflict Constraint, a person can have multiple, conflicting normative beliefs at the same time. I argued that we should accept this possibility and more fully appreciate its significance. By the same token, it seems, we should also accept that we can have multiple, sometimes conflicting lines of deliberation at the same time. An akratic anorexic can formulate and execute a complex and demanding weight loss regimen while also thinking through and taking steps to avoid weight loss. We can see multiple, completed deliberations here, rather than one incomplete one. Normative belief need not bring deliberation to a close in the sense of resolving all normative conflict. Expecting such resolution is yet another way to violate the Conflict Constraint, and to view evaluation as more coherent than it is.



In rejecting guise-of-the-good views, Velleman (2000, p. 99) writes that he hopes “for a moral psychology that can make room for the whole motley crew” of acting creatures. Stocker (1979, p. 739) writes: “Philosophical theories...have depicted the psyche, especially the interrelations between motivation and evaluation, as far too simple, far too unified, and far too rational.” I think that some of our theories have depicted evaluation itself as far too simple, far too unified, and far too rational. Even the theories that explicitly allow for the possibility of conflicts in evaluation rarely do justice to the details or the significance of those conflicts. I have tried to make more room for the motley crew of acting creatures by insisting on the complexities of evaluation, and especially of normative belief. And I argued that it is naïveté about the coherence of our normative beliefs, not about their connection to intentional action, that we should be exorcising from our understanding of ourselves. Once we do that, even an ambitious view like the Normative Belief Requirement becomes defensible. I think the lesson is more general: that a necessary connection between intentional action and normative belief can allow for and unify *all* of the phenomena of intentional action. What I have given here is one piece of the argument. I have argued that an understanding of the conflicts in our normative beliefs can make room for akratic action under the guise of the good.

*Acknowledgments:* This paper has benefited from discussion with Dylan Bianchi, Benjamin Brast-McKie, Jeremy Carey, Lee-Ann Chae, Colin Chamberlain, Julia Chislenko, Joshua Eisenthal, Nick French, Elena Garadja, Hannah Ginsborg, August Gorman, Ulf Hlobil, David

Hunter, Gregory Khasin, Niko Kolodny, Christine Korsgaard, Arthur Krieger, Dylan Murray, Derek Parfit, Kirsten Pickering, T. M. Scanlon, Daniel Sharp, Barry Stroud, Benjamin Wald, R. Jay Wallace, Yuan Wu, and audiences at the American Philosophical Association (Central Division), Indiana University—Bloomington, the Pittsburgh Area Philosophy Colloquium, UC Berkeley, and UT Knoxville. I gratefully acknowledge all of these people, and, most of all, Charles Goldhaber for constant encouragement and detailed feedback on many drafts. Research on this project was supported by a Mabelle McLeod Lewis Memorial Fund fellowship, a Temple University Summer Research Award, and a Temple University College of Liberal Arts Research Award.

## References

- Adler, J. (2002). "Akratic believing?" *Philosophical Studies*, 110(1), 1-27.
- Anscombe, G.E.M. (1957). *Intention*. Oxford: Blackwell.
- Aristotle. (1984). *The collected works of Aristotle*. Ed. J. Barnes. Princeton, NJ: Princeton University Press.
- Aristotle. (1999). *Nicomachean ethics*. Tr. T. Irwin. Indianapolis, IN: Hackett.
- Arpaly, N. (2000). "On acting rationally against one's best judgment." *Ethics*, 110(3), 488-513.
- Audi, R. (1979). "Weakness of will and practical judgment." *Noûs*, 13(2), 173-196.
- Beebe, J.R. (2013). "Weakness of will, reasonability, and compulsion." *Synthese*, 190(18), 4077-4093.
- Boyle, M., and Lavin, D. (2010). "Goodness and desire." In S. Tenenbaum (Ed.), *Desire, practical reason, and the good* (pp. 161-201). New York: Oxford University Press.
- Bratman, M. (1979). "Practical reasoning and weakness of the will." *Noûs*, 13(2), 153-171.
- Buss, S. (1987). "Weakness of will." *Pacific philosophical quarterly*, 78(1), 13-44.
- Chislenko, E. (2016). "A solution for Buridan's Ass." *Ethics*, 126(1), 283-310.
- Cohen, J. (1992). *An essay on belief and acceptance*. Oxford: Oxford University Press.
- Collins, C., and Postal, P.M. (2014). *Classical NEG raising: An essay on the syntax of negation*. Cambridge, MA: MIT Press.
- Davidson, D. (1980a). "How is weakness of the will possible?" In D. Davidson, *Essays on actions and events* (pp. 21-42). Oxford: Clarendon Press.
- Davidson, D. (1980b). "Intending." In D. Davidson, *Essays on actions and events* (pp. 83-102). Oxford: Clarendon Press.

- Davidson, D. (1982). "Paradoxes of irrationality." In R. Wollheim (Ed.), *Philosophical essays on Freud* (pp. 289-305). Cambridge: Cambridge University Press.
- De Sousa, R. (1971). "How to give a piece of your mind: Or, the logic of belief and assent." *The review of metaphysics*, 25(1), 52-79.
- Enoch, D. (2011). *Taking Morality Seriously*. New York: Oxford University Press.
- Gosling, J. (1990). *Weakness of the will*. New York: Routledge.
- Gregory, A. (2017). "How Verbal Reports of Desire May Mislead." *Thought*, 6, 241-249.
- Kant, I. (1997). *Critique of practical reason*. Tr. Mary Gregor. Cambridge: Cambridge University Press.
- Kant, I. (1998). *Groundwork of the metaphysics of morals*. Tr. Mary Gregor. Cambridge: Cambridge University Press.
- Korsgaard, C. M. (1996). *The sources of normativity*. Cambridge: Cambridge University Press.
- Korsgaard, C. M. (2009). *Self-constitution: Agency, identity, integrity*. Oxford: Oxford University Press.
- May, J., and Holton, R. (2012). "What in the world is weakness of will?" *Philosophical studies*, 157(3), 341-360.
- McIntyre, A. "Is akratic action always irrational?" In O. Flanagan and A. Rorty (Eds.), *Identity, character, and morality* (pp. 379-400). Cambridge, MA: MIT Press.
- Mele, A. (1983). "'Akrasia,' reasons, and causes." *Philosophical studies*, 44(3), 345-358.
- Mele, A. (1987). *Irrationality: An essay on akrasia, self-deception, and self-control*. New York: Oxford University Press.
- Mele, A. (2010). "Weakness of will and akrasia." *Philosophical studies*, 150, 391-404.
- Nagel, T. (1972). "War and massacre." *Philosophy and public affairs*, 1(2), 123-144.

- Plato. (1997). *Plato: Complete works*. Ed. John M. Cooper. Indianapolis, IN: Hackett.
- Priest, G. (2006). *In contradiction*, 2<sup>nd</sup> edition. Oxford: Clarendon Press.
- Raz, J. (2010). "On the guise of the good." In S. Tenenbaum (Ed.), *Desire, practical reason, and the good* (pp. 111-137). New York: Oxford University Press.
- Saarinen, R. (2011). *Weakness of will in Renaissance and Reformation thought*. New York: Oxford University Press.
- Schwitzgebel, E. (2012). "Self-ignorance." In J. Liu and J. Perry (Eds.), *Consciousness and the self: New essays* (pp. 184-197). Cambridge: Cambridge University Press.
- Setiya, K. (2007). *Reasons without rationalism*. Princeton, NJ: Princeton University Press.
- Setiya, K. (2016). "Akrasia and the constitution of agency." In K. Setiya, *Practical knowledge: Selected essays* (pp. 253-271). Oxford: Oxford University Press.
- Sripada, C. (2016). "Self-expression: A deep self theory of moral responsibility." *Philosophical studies*, 173(5), 1203-1232.
- Stocker, M. (1979). "Desiring the bad." *The journal of philosophy*, 76(12), 738-753.
- Tenenbaum, S. (2007). *Appearances of the good: An essay on the nature of practical reason*. New York: Cambridge University Press.
- Velleman, J.D. (2000). "The guise of the good." In J. D. Velleman, *The possibility of practical reason* (pp. 99-122). Oxford: Oxford University Press.
- Walker, A.F. (1989). "The problem of weakness of will." *Nous* 23(5), 653-676.
- Watson, G. (1977). "Skepticism about weakness of will." *The philosophical review*, 86(3), 316-339.
- Williamson, T. (2000). *Knowledge and its limits*. Oxford: Oxford University Press.