



Assessing Recent Agent-Based Accounts of Right Action (and More)

Graham Renz¹

Accepted: 1 March 2020 / Published online: 5 March 2020
© Springer Nature B.V. 2020

Abstract

Agent-based virtue ethical theories must deal with the problem of right action: if an action is right just in case it expresses a virtuous motive, then how can an agent perform the right action but for the wrong reason, or from a vicious motive? Some recent agent-based accounts purport to answer this challenge and two other related problems. Here I assess these accounts and show them to be inadequate answers to the problem of right action (and one of the other problems for agent-basing). Overall, it is shown that the most recent and promising attempts at squaring agent-based virtue ethics with commonsense morality are flawed, and so, the case for agent-basing in general that much dimmer.

Keywords Virtue ethics · Agent-basing · Right action · Daniel Doviak · J.P. Walsh

1 Introduction

Agent-based virtue ethical theories may be distinguished from other virtue ethical theories, such as Aristotelian virtue ethical theories (for example, Hursthouse 1999), by the fact that they take the normative weight of virtues to be primitive or foundational. Rather than ground the normative ‘oomph’ of virtue in the notions of flourishing or well-being, agent-based views take it that the expression of praiseworthy motives is what matters most. As Michael Slote (2001: 5), the premier contemporary defender of agent-based virtue ethics, puts it, the view is one that “treats the moral or ethical status of acts as entirely derivative from independent aretaic (as opposed to deontic) ethical characterizations of motives, character traits, or individuals.”¹ So, on such a view, Jane’s action is right just in case it flows from or expresses, say,

¹For further discussion, see Hursthouse and Pettigrove (2016: §2.2). My concern here is only with Slote’s flavor of agent-basing, where a *particular* agent’s motives are what matters, not what a *hypothetical virtuous agent’s* motives are (cf. Zagzebski 2004).

✉ Graham Renz
renzgraham@gmail.com

¹ Department of Philosophy, Washington University in St. Louis, Wilson Hall, One Brookings Drive, St. Louis, MO 63130, USA

concern for Jim's well-being (benevolence), and not whether it leads to Jane living a good life or something else entirely, such as certain consequences. What's right is what comes from primitively good aspects of an agent's inner-life (Slote 2001: ch. 1.3.).

Agent-based theories can do quite a bit of work, for example, by handling the problem of moral luck and furnishing a plausible metaethics.² But they suffer from some serious issues. Perhaps the most pressing issue such theories face is how to square their account of right action with fairly intuitive or commonsense judgments about whether a certain action is right or wrong. If all that matters in determining whether some action is right or wrong is whether it came from or expresses virtue or vice, then so long as an agent's motives are virtuous, they cannot act wrongly. But it seems rather clear that agents can perform the right action while expressing vicious motives, and conversely, that agents can perform the wrong action while expressing virtuous motives. Consider Ramon Das's (2003: 326) example of the poolside savior:

A man dating a woman with a young child dives into a swimming pool to save the child from drowning. He cares not at all for the child, and is motivated exclusively by a desire to impress the woman as a means, let us suppose, to sleeping with her.

Intuitively, the poolside savior performs the right action—commonsense suggests that saving the child is the right thing to do. But, despite this, he acts from deplorable motives, expressing an unsavory inner-psychological life. So if, per agent-based theories, the expression of virtues is what grounds the ethical status of actions, how do we accommodate the fact that the poolside savior acts rightly?

Das (2003) and other critics of agent-based theories have shown convincingly that Slote's response to this and similar cases is inadequate.³ But recent proponents of agent-basing have put forward rather ingenious ways to get around this challenge, strengthening the overall case for agent-basing. In particular, recent accounts of right action proposed by Doviak (2011) and Walsh (2016) show promise. Doviak suggests that by thinking of right action in terms of the *optimal expression* of virtue over vice we can solve the problem of right action and two other related problems for agent-basing: the flouting of the deontic principle that 'ought implies can' and the problem of motivational overdetermination. Walsh proposes that if we understand right action as the *realization* of virtuous motives, then the problem of right action dissolves. However, while these proposals have much to recommend them, they are ultimately inadequate. My goal in this paper is to show that the accounts of Doviak and Walsh fail as general solutions to the problem of right action, and that in the end they solve only the problem of motivational overdetermination. The upshot is that we should take pause in our consideration of adopting agent-based theories.

With this goal in mind, the discussion below is organized thus. First, in the proceeding section, I consider how Doviak's proposal purports to address the three problems, and then in section three, I show that it is subject to counterexample. There I argue that, should the range of a subject's alternatives be plausibly expanded or constrained, Doviak's proposal to solve the problem of right action and respect the principle that 'ought implies can' leads to counterintuitive results. In section four, I explicate Walsh's proposal and show that it

² For more on the problem of moral luck, see Slote (1992: ch. 7); cf. Athanassoulis (2005). For an agent-based approach to metaethics, see Slote (2010); for discussion, see Kauppinen (2014).

³ See Brady (2004), Garrard (2000), Jacobson (2002), Johnson (2003), Russell (2008), Svensson (2010), and Swanton (2001).

solves the problem of right action only when a subject's motivational set is stipulated to contain motives a critic has little reason to countenance. In the fifth and final section, I assess whether Walsh's proposal may be extended so as to address the flouting of 'ought implies can' and the problem of motivational overdetermination. I conclude it solves only the latter. Overall, my strategy is to show that Doviak and Walsh can solve only certain versions of the adumbrated problems, and so, that their proposals are not general solutions with which agent-based approaches may advance.

2 Right Action and the Optimal Expression of Virtue

Doviak (2011) suggests his view solves the problem of right action and two related problems for agent-based theories. Before examining his view, allow me to lay out these two other problems.⁴

The first problem is that agent-based views appear to flout the deontic axiom that 'ought implies can' (Hurka 2001; cf. Jacobson 2002). Suppose there is a masochist who loves inflicting pain on innocent park-goers. They wait in the park until some joggers come by, and are faced with two alternatives. The masochist can either act from malice and inflict pain on the joggers, or they can let the joggers carry on, but only because of some deplorable sort of self-concern. In short, the masochist can only act from deplorable motives. Now, it is obvious that the masochist ought not to inflict pain on the joggers, nor ought they to abstain from doing so *only* because of a deplorable sort of self-concern. But since agent-based views take the rightness or wrongness of an action to be determined by the motives it expresses, it seems the masochist cannot do what they ought to; they can express only vicious motives. So, it seems agent-based views must deny that 'ought implies can.'

The second problem is that agents can have more than a single motive for action, and these motives can have conflicting moral worth (Driver 1995). Imagine a narcissistic philanthropist who has decided to donate money for the construction of a hospital in their hometown. The philanthropist is motivated by benevolence towards their fellow denizens and a vicious sort of pride, an unhealthy concern for their reputation. Suppose that either motive alone would be enough to engender the donation. If the philanthropist acts from both motives it seems their donation is both right and wrong—it seems to be problematically overdetermined.⁵ If we claim they act just from benevolence, the agent-based view seems either ad hoc or to flout its core thesis. If we claim they act just from an unhealthy concern for their pride, we must claim they act wrongly—that an intuitively right action is made wrong somehow by its flowing from vicious motives. And this just is the problem of right action all over again.

To solve these problems, Doviak suggests we understand right action as *the optimal expression of virtue over vice*. Let us unpack his account and see how it purportedly addresses the three problems.

Doviak understands virtues and vices as psychological dispositions whose expression make an agent admirable or deplorable, respectively. He takes it as uncontroversial that agents often act from numerous motives, sometimes all admirable, sometimes all deplorable, but, most often, a mix thereof. So, for any given action, an agent will likely express some degree of

⁴ For the sake of argument, I assume that these two other problems are genuine; some readers may not find them so. My concern is only whether or not Doviak's account can do what he claims it can.

⁵ For further discussion, see Sider (2003).

admirability and deplorability. (Some may find this contentious, but again, for the sake of argument, I follow Doviak.) The novelty of Doviak's proposal lies in his formalization of the above. He suggests that for each and every virtue and vice an agent expresses when they act we should assign some numerical value. For virtues, he suggests we assign a value within the range $[0, 1]$, and for vices, a value within the range $[0, -1]$. So, for example, if I donate to charity from both benevolence and an unhealthy concern for my reputation, I express, say, 0.5 benevolence and -0.4 unhealthy pride. After tallying all the admirability and deplorability values for some agent and their action, we are to sum them into a single value, what Doviak (2011: 267) calls *net intrinsic virtue value* (net-IVV). Doviak (2011: 267) says that "an action's net-IVV represents how virtuous (or admirable) the agent is *on balance* in performing [some] action." In the above example, supposing benevolence and unhealthy pride exhaust the relevant psychological factors, my net-IVV would be 0.1.

With the notion of net-IVV in hand, Doviak claims that an action is right just in case it *maximizes* net-IVV (for some subject S at time T). So, my donating was right just in case it maximized net-IVV compared to my alternatives. But, if the philanthropist genuinely had the alternative of donating *without* a partly vicious motive, and so, could have expressed just as much virtue while expressing less vice, they have failed to maximize net-IVV, and so, acted wrongly. To quote Doviak (2011: 268-9) at length:

[My view] implies that one cannot determine the moral status of a given action without examining the areatic values of its alternatives. It is not enough to know that the agent expresses virtue (or that he is admirable in some respect) in order to know that the action is right. If there is some other alternative available to the agent that has a higher net-IVV, then any virtuous but non-maximizing action is wrong. In addition, [my view] implies that an agent's being on balance vicious (or deplorable) in the performance of an action is not sufficient for that action's being wrong. If the agent is overall vicious in the performance of an action, but the performance of any of the other options would make the agent even more vicious, then the option that makes him least vicious (or least deplorable) will be obligatory.

So, according to Doviak, an action is right just in case it expresses the greatest net-IVV compared to the alternatives (for a subject S at a time T).

Return to our three problems: the problem of right action, the flouting of 'ought implies can,' and motivational overdetermination. Consider again Das's poolside savior, the man who saves a drowning child just so that he can impress and so sleep with the child's mother. How does Doviak maintain the poolside savior acts rightly despite his deplorable motives? It seems the savior has two alternatives: save the child just to impress their mother so as to sleep with her or not save the child and still try to sleep with their mother. The latter alternative expresses not just one deplorable motive, but two: the desire to manipulate the mother *and* callousness towards her drowning child. While the former alternative is not purely virtuous, it expresses at least minimally admirable motives, and so, scores positively on one measure of virtue. Given the savior's psychological makeup (let us suppose that he *really* wants to sleep with the child's mother), these may be his only genuine alternatives. So, of the alternatives before him, saving the drowning child, even taking into account his deplorable motives, is the action which maximizes net-IVV, and so, is right.

What of the masochist and the worry that agent-based views must reject the principle that 'ought implies can'? Since the right action is the one that, of the alternatives, maximizes net-

IVV, there will always be a right action, however deplorable an agent's motives. In Doviak's (2011: 269) words:

Intuitively, we know that [the masochist] has an obligation to *not* harm the [joggers], even if the only way [the masochist] can bring himself to do this is out of some pathetic form of self-concern. [My view] explains this obligation as follows: [the masochist] is more deplorable on balance under the circumstance in which he harms [the joggers] than he is under the circumstances in which he does not.

Since, on Doviak's view, there is always some action that maximizes net-IVV, and so, is right, there will never be a circumstance in which an agent cannot do what they ought to. So, agent-based views do not violate the principle that 'ought implies can.'

Lastly, consider again the narcissistic philanthropist and the problem of motivational overdetermination. The worry is that one and the same action can have virtuous and vicious motives, and so, be both right and wrong. Doviak's view seems to easily handle this worry. For, since it *balances* all the relevant psychological features of an agent to produce a single net-IVV, it does not allow for an action to be genuinely overdetermined. Although the philanthropist's donation might express both virtuous and vicious motives, on balance, it will yield a single net-IVV, and so, not be *both* right and wrong. Various motivated acts like this will only be more or less virtuous or vicious, but not both at once.

3 Problem(S) Solved?

While Doviak's proposal to collapse all motivational evaluations into a single value is ingenious, his view is inadequate. I will show this much via counterexamples.

Consider again Das's poolside savior. Intuitively, he does what he ought to—he performs the right action. For, it is uncontroversial that saving the drowning child is right. However, he has deplorable motives. Now, if the poolside savior's only alternatives are saving the child or not saving the child, Doviak can claim he acts rightly, that he does the right thing but for the wrong reason. But why suppose these are the only alternatives? Why not suppose there could be some third alternative: saving the child from a *less deplorable* motive, say, a sort of obsequious moral righteousness? If the poolside savior has this third alternative, then Doviak can no longer claim the creep does the right thing for the wrong reason; for, there is some alternative open to him that would increase his net-IVV (being obnoxiously righteous is closer to virtue than being a manipulative creep), and so, be uniquely right.

In short, it seems an agent can act rightly but for the wrong reason even though they could have acted *more rightly*. The poolside savior can do what is right even if he is not maximizing his expression of virtue, and this is something Doviak's account cannot handle. For, again, if an available alternative would yield a higher net-IVV, it is the uniquely right action according to Doviak.⁶

It seems Doviak's account only solves the problem of right action in a select number of examples, examples where alternatives are kept to a minimum. But why think these are the only examples facing agent-based views? Das's example is just one of many, and for an

⁶ I should note that my aim here is not to criticize the notions of maximization or optimality per se, but only in Doviak's identification of it with rightness.

account of right action to be adequate, it should be able to handle all plausible cases. An account which addresses just a subset of cases is no genuine account at all.

What about the other two problems? The results are mixed. Consider, instead of the masochist, Larry, who simply loves larceny. Larry gets a rush from stealing, and also sees it as a way to level the economic playing field between the haves and the have-nots. When thinking about how to spend his day, Larry considers the following two options: steal out of envy or steal out of boredom. The latter alternative maximizes Larry's net-IVV, for stealing because one wants a rush is closer to virtue than stealing out of a vicious inner-state. Given this, Doviak is committed to claiming that stealing out of boredom is the right action, and, moreover, that Larry ought to so act. Doviak will note that, since the latter alternative is right, that is, what Larry *ought* to do, and that Larry *can in fact* act on it, that his view does not violate the principle that 'ought implies can.' However, if a view implies that what an agent ought to do is steal for the rush of stealing—or anything else that seems clearly wrong—we have good reason to reject it.

It seems the above is an example where an agent ought to do, on Doviak's view, something we take to be obviously wrong, unlike the masochist example where one alternative is to do the intuitively right thing. But why should we think all pertinent (or broadly realistic, however extreme) examples include as a genuine alternative what we take to be the *intuitively right action*? Why suppose that the action available to an agent that maximizes their net-IVV—whatever the action is—is the intuitively right action? The two seem to come apart.

At this point, allow me to consider a few objections. First, regarding the flouting of 'ought implies can,' you might reply that agents simply always (barring various constraints, such as coercive forces) have as an available alternative doing the right thing. That is, Larry could simply not steal. If this is right, then Doviak needn't be committed to claiming that Larry ought to steal out of boredom.

In response, I do not wish to deny that in various important senses—senses of psychological, physical, and metaphysical possibility—Larry can refrain from stealing altogether. However, it certainly seems possible that some agents are such that what they take to be the alternative routes of action before them include only intuitively wrong actions. To truly settle the matter, we would need an account of what an *available alternative* is, and I have no intentions of providing one here. But a good first take on an available alternative is this: a psychologically-live option. Larry may be fully aware of the fact that he could walk in the park or get an ice cream cone instead of stealing, but that doesn't mean he's genuinely considering either of those alternatives. He might just be set on stealing. So, perhaps in something like the way in which spontaneously traveling to Antarctica tomorrow is not an alternative available to me right now—even though I might have the necessary funds, et cetera—refraining from stealing just isn't an option for Larry.

Second, given this response, you might object that in critiquing Doviak I'm trying to have it both ways. At first I critique him for not allowing more alternatives to be available to the poolside savior. Then, in the case of Larry, I limit the range of available alternatives; I deny Larry could simply refrain from stealing. But how can I insist in one case that an additional alternative be available and then deny such an alternative be available in another case?⁷

In response, my strategy is not to show that, for any relevant case of action, certain alternatives are or are not available. Such is a task far too great for a single paper, and one I have no idea how to accomplish. My strategy is just to show that there are plausible cases

⁷ Thanks to anonymous reviewers for bringing to my attention the need to consider this objection.

which Doviak's account cannot handle. Some of these cases include more alternatives than when presented by Doviak and others; some of the cases include fewer alternatives (or include only certain alternatives) than when presented by Doviak and others. So I'm not arguing that Doviak has erred in using examples that contain (or don't contain) certain available alternatives. I'm simply suggesting that the range of cases that need to be accounted for is larger than Doviak presents it.

Third, and in regard to the flouting of 'ought implies can,' you might object that, despite the oddity of calling boredom-inspired theft right, it is exactly what agent-based views need to say to avoid the flouting of 'ought implies can.' Although this consequence is deeply counterintuitive, it does not prove Doviak's view is wrong.

Now, while I do not doubt this move—call it the *affirmation of bottomless deplorability*—addresses the problem at hand, it does so only in a way that agent-based critics will find terribly unsatisfactory.⁸ For recall the main issue to be addressed by agent-based views is that of squaring their moral assessments with *commonsense* morality. Clearly, calling Larry's boredom-inspired theft right does not accomplish this.

At this point, Doviak could evoke a contrastivist account of right action. That is, he could claim that given Larry will steal one way or another, stealing out of boredom is right *compared* to stealing from a sense of envy.⁹ If contrastivism is right, then there may be no problem for Doviak.¹⁰

In response, it is important to note that Doviak's account is based on available *alternatives*, not mere contrast cases. What's available to Doviak in Larry's case are *Larry's alternatives*, not contrast cases wherein Larry, or someone very much like Larry, has other alternatives, say, refraining from stealing altogether. Plus, given that Doviak's aim is to secure for an agent-based theory the ability to claim an agent can do what we intuitively think they ought to, full-stop, a merely contrastive sense of 'rightness' doesn't do us much good. Commonsense suggests we ought not to steal, full-stop (barring extraordinary circumstances), and claiming that Larry ought to steal out of boredom, *compared* to stealing out of a sense of envy, doesn't secure for us that commonsense notion of full-stop rightness. I submit that, in some weak sense, perhaps we might *excuse* Larry for stealing, but that doesn't change the fact we believe he ought to not steal in general and in the first place.¹¹

Now, while I think Doviak's account fails to solve the above two problems, I do think it successfully obviates the objection from motivational overdetermination. For, again, if all relevant motives are collapsed, that is, *balanced*, into a single net-IVV, no action will ever be *both* right and wrong. Any action will just be *more or less* right or wrong, located at some point on a continuum between 1 and -1, and not pulled both 'this way' by virtue *and* 'that way' by vice.

Let's take stock. Doviak proposed we understand right action as the optimal expression of virtue over vice. His view purported to solve the problem of right action for agent-based virtue theories, and also two additional problems, the worry that such views flout the principle that 'ought implies can' and a worry that actions could be both right and wrong. I have shown

⁸ Note that Doviak must also endorse a converse position—call it the *affirmation of limitless admirability*—which claims that, no matter how virtuously an agent acts, so long as there is an alternative to act more virtuously, the agent does not act rightly.

⁹ For discussion of a contrastivist account, see Snedegar (2013).

¹⁰ Thanks to Julia Driver for brining this objection to my attention.

¹¹ Thanks to an anonymous reviewer for prompting me to make my discussion here clearer. For a discussion of excusing conditions (in some circumstances), see McMahan (2009).

Doviak's view is effective only in addressing the third problem, and so, is not an adequate account of right action (and more) for agent-based virtue theories.^{12,13}

4 Right Action and Motive Realization

Walsh (2016: 654) suggests that *an action is right just in case (and because) it realizes an agent's morally praiseworthy motive*. In brief, the idea is that when an agent acts from an admirable motive, and the goal or end of that motive obtains, they act rightly. Let's unpack.

First, how does Walsh understand a motive, and how does it differ from, say, an intention? Allow me to quote Walsh (2016: 654) at length.

A motive may be realized in a variety of ways, and no one of these ways constitutes the uniquely right or best action. For example, if the motive which underlies my intention to go to the gym three times a week is to lose a certain amount of weight, then that motive will have been realized if I do actually lose that amount of weight. Notice that this will still be the case even if my intention to go to the gym three times a week is *not* realized. Perhaps I only go twice a week, but still manage to lose the weight. This should alert us to the possibility that motives may still be realized *even if the intentions to which they give rise are not realized*. Briefly put then, intentions are such that their realization requires a specific sort of performance on the part of the agent (my intention to go to the gym three times a week can *only* be realized by my actually going three times a week). Motives, on the other hand, do not require a specific sort of performance for their realization; agents may realize them in a number of ways, each of which might involve the selection of different sets of specific intentions. [emphases original].

And what of realization? If what makes an action right is that it *realizes* a good motive, that it makes the world a certain way, is this just to place the moral worth of an action in its consequences? Is this to abandon the entire agent-based project? Not at all says Walsh. The state of affairs aimed at by some agent is not itself good or bad, is not itself an independent moral standard, but, rather, is 'good' only inasmuch as the virtuous agent is motivated to bring it about. The motives of the agent are still moral bedrock. Walsh's point is just that there are success criteria for determining whether an agent has in fact brought about the state of affairs their virtuous dispositions motivated them to bring about. The notion of realization just picks up on the fact that sometimes virtuous agents fail to make the world the way their virtues incline them to.

So, with all this in hand, to repeat, on Walsh's account, an action is right just in case it realizes an agent's virtuous motives. Conversely, an action is wrong just in case it fails to realize an agent's virtuous motives or realizes an agent's vicious motives.

¹² For further discussion and development of Doviak's view, see Ciurria (2012).

¹³ Another objection to Doviak's view worth registering is that it seems committed to counterintuitive comparative evaluations. Suppose that in making a modest donation, Jim expresses numerous virtues—generosity, kindness, and justice—to a small extent, resulting in a net-IVV of 0.6. And suppose that while fighting in a just war, Jane expresses the single virtue of courage to a great extent, resulting in a net-IVV of 0.6. Given that Jim and Jane have identical net-IVVs, Doviak must claim they are equally praiseworthy.

But who is more praiseworthy, someone who acts from minimal generosity, kindness, and justice, or someone who acts from great courage? Intuitively, the courageous hero is more praiseworthy than the modest but variously motivated alms giver. Thanks to Ge Fang for bringing this objection to my attention.

How does this account help us deal with the problem of right action? Consider again Das's poolside savior; he does the right thing but from deplorable motives. On Walsh's view, we needn't say the poolside savior acts wrongly because his saving the drowning child does not *realize* his deplorable motive, his motive to manipulate the child's mother into sleeping with him. As Walsh (2016: 658) puts it:

That motive is realized by the action of sleeping with the child's mother, not the action of saving the child's life. The latter action *helps* to realize the agent's motive, but does not in itself realize it. This occurs only when the agent actually sleeps the with child's mother. Thus it is the action of sleeping with the child's mother which is morally wrong, because it is manipulative and callous.

Walsh (2016: 658) goes on:

Importantly none of this prevents us from saying that, in saving the child's life, Das'[s] savior does in some sense act manipulatively. This is because his actions *express* his motive of manipulating the child's mother, even if they do not realize that motive. Of course, given that the agent's actions express a morally blameworthy motive, it is appropriate for us to find moral fault with his character.

So far, Walsh's account accommodates commonsense morality. But we are not yet in a position to say the poolside savior acts *rightly*; all we have done is show that he does not realize a deplorable motive by saving the child, and so, act *wrongly*.

Walsh (2016: 659ff) suggests that we posit an additional, subordinate motive: the poolside savior's motive to spare the mother from pain. If we posit this motive we can say the poolside savior acts rightly. By saving the child, the poolside savior *realizes a virtuous motive*, and so, does the right thing. Of course, this subordinate motive is accompanied by the *expression* (and is ultimately in the service) of a vicious motive, the motive to manipulate the mother. But, given that on Walsh's account right action is defined in terms of motive *realization*, the mere expression of vicious motives does not prevent us from saying the poolside savior *acts* rightly.¹⁴ So, Walsh claims, agent-based theories can accommodate the commonsense view that agents can do the right thing but for the wrong reason, or from deplorable motives. By 'stacking' motives and making a distinction between realization and expression, agent-based views can accommodate commonsense moral evaluations and so the problem of right action. So whereas Doviak urged us to *collapse* all motives and motivational evaluations into a single value, Walsh suggests we instead *stratify* them.

Walsh's account seems to have an edge on Doviak's. For instance, Walsh can say that the poolside savior acts rightly when he saves the drowning child just so that he can sleep with the mother because, *whatever his alternatives*, he in fact realizes a good motive when he saves the child. Of course, the poolside savior expresses a bad motive, but as noted, motive expression only factors into *agent* evaluation, not *act* evaluation. Also, looking ahead to the flouting of 'ought implies can,' Walsh needn't affirm *bottomless deplorability* (or *limitless admirability*): it does not matter that Larry's least vicious alternative is to steal out of boredom because doing so both expresses and *realizes* deplorable motives. By grounding right action in virtuous

¹⁴ An interesting objection that Walsh (2016: 658ff) addresses, but one I will not consider here, is that on his account it seems as though, so long as an agent does not realize some deplorable motive, they cannot act wrongly. For instance, if Gustav is motivated to influence an election but must first blackmail a government official, so long as the election is not in fact influenced, he does not act wrongly. Walsh responds in exactly the same way he suggests we account for the poolside savior's acting rightly: posit a subordinate motive that *is* realized.

motive realization, rather than maximal net-IVV, Walsh can avoid a commitment to intuitively wrong actions being right (and virtuous, but not maximally virtuous actions being wrong).

But it is not clear that Walsh's account is satisfactory. It seems plausible that the poolside savior could save the drowning child not because he had some virtuous, subordinate motive, but because he thought doing so was simply the best way to manipulate and so sleep with the child's mother. It seems possible to do the right thing without expressing or realizing a virtuous motive at all, whether primary or subordinate. So in positing a subordinate, virtuous motive Walsh has precisified the case in a way that his account can handle, but that critics can reject. Why might a critic simply allow a virtuous motive to be posited?¹⁵

Walsh might reply that the best way to explain why the poolside savior acts as he does is because he has this virtuous, subordinate motive. For, if he truly had no virtuous motives, why would he save the drowning child at all? How could a truly and wholly virtuous agent do what's right?¹⁶

In response, I simply do not find it plausible that the only or best explanation for the poolside savior's action is the existence of a subordinate, virtuous motive. It seems perfectly reasonable that he might just have vicious motives, and know that the best way to get what he wants is to save the child. Some agents, especially vicious, manipulative ones, are good at putting on a show; their doing the right thing shouldn't move us to imbue them with virtue. Indeed, to think the only or best explanation of an otherwise vicious agent's right action is the presence of virtue seems to beg the question against the critic of agent-basing. It seems to suppose that right action is only explicable in terms of virtuous motives, and this is exactly what is in question. So, I conclude that, at best, Walsh can only solve the problem of right action in cases where we suppose an otherwise vicious agent has, deep down, a modicum of virtue. But this means his account is not general, and so, not fully satisfactory.

5 Can we Extend Motive Realization?

But perhaps Walsh's account handles the alleged flouting of 'ought implies can' and the problem of motivational overdetermination? If Walsh can adequately address these two problems, his view is ahead of Doviak's at least, and a better contender for agent-basing in general.

How might the infrastructure of motive realization help retain the principle that 'ought implies can'? Recall the problem case: Larry can either steal from a sense of vicious envy or steal from a sense of boredom. Intuitively, Larry ought not to steal at all, but as stipulated, these are his only alternatives. While Doviak tried to solve this problem by claiming one of the alternatives was right, Walsh is committed to both alternatives—actions based on either of the two motives—being wrong. For both seem to express and, should Larry's larceny go as planned, realize vicious motives. So it seems Walsh really does have a problem here: he seems unable to say Larry *can* do what he *ought to*. One might try to posit an additional motive in Larry, one that is virtuous and genuinely available to be acted on, but this would just be to change the example, and so, not address the problem at hand.

Walsh must accomplish two tasks: (i) show that what Larry *ought to do* is refrain from stealing, and (ii) show that Larry *can in fact* refrain from stealing. Walsh can accomplish the first task, but

¹⁵ Thanks to Eric Brown for helping me to see this.

¹⁶ Thanks to an anonymous reviewer for bringing this response to my attention.

only by appealing to a separate account of agent-based action guidance.¹⁷ For, as the example stands, there is no virtuous motive that Larry has that could be realized (or even expressed) to get Walsh the right action—he can only say that acting on either motive is wrong. What sort of account of action guidance might Walsh rely on though? He can say, for example, that what an agent ought to do is what an ideally virtuous observer would *deem right* in the circumstances (see Kawall 2002, Van Van Zyl 2009). While I think this move is legitimate, it invites a worry. One might worry separating our account of right action from our account of action guidance is theoretically problematic. For, why have two accounts when some virtue theories simply collapse the two and have one? Why appeal to a theoretical, idealized agent when the motives of actual, non-ideal agents are supposed to be the basis of agent-based views like Slote’s and others? If an account of action guidance can do the work Walsh’s theory of motive realization is supposed to do, namely, solve the problem of right action, I see little reason for adopting Walsh’s account in the first place.

The bigger problem, I think, is that Walsh simply cannot accomplish the second task. Even if we establish, in agent-based terms, that Larry ought not to steal, the example is such that he will steal. Since Larry has no virtuous motives, he cannot do what he ought to, and so, it seems that Walsh’s account does violate the principle that ‘ought implies can.’ Walsh might say the example is impossible, or that, so described, Larry is psychologically deviant or deficient, and so, no counterexample to his view. But as I noted in my discussion of Doviak above, it seems the case of Larry is plausible and my goal isn’t to show that Walsh (or Doviak) cannot solve *some* versions of the three problems, but only that he (and Doviak) cannot solve all plausible versions. While this puts Walsh at no disadvantage to Doviak, it does cast doubt on his account is general.

What of the third problem, the problem of motivational overdetermination? Here I think Walsh’s account can be successfully extended. The narcissistic philanthropist has both virtuous and vicious motives, but, for one, they needn’t both be realized by donating funds for a hospital. Plausibly, one motive might just be expressed. Donating the funds is the right action, and so, we can suppose the vicious motive of enhancing the philanthropist’s reputation is merely expressed, but their virtuous motive of benefitting her fellow denizens is realized. (Or, perhaps both are realized, but at different times, or in different respects.) Second, we might structure the philanthropist’s motives such that one is in service of the other, that is, so they are not the basis of the self-same action. It is reasonable to understand the vicious motive as, say, one whose realization just is the further motivation to benefit one’s fellow denizens. So understood, the donation itself would not be overdetermined, both right and wrong—the formation of the virtuous motive would be somehow deviant, and the donation right. And while it is odd to say a virtuous motive could be deviant, it seems psychologically and phenomenologically plausible, and, importantly, gets around the problem at hand.

6 Conclusion

Agent-based virtue theories must deal with a serious problem: how an agent can do the right thing but for the wrong reason and, conversely, how an agent can do the wrong thing but for the right

¹⁷ Walsh (2016: 660) notes such an appeal is open to him, and mentions one account proposed by Van Van Zyl (2009), but does not pursue or adopt it (although he mentions it is compatible with his view). In fact, he seems unconvinced that he or other proponents of agent-basing need to adopt an account of action guidance at all. He writes (2016: 660), “I am not convinced that agent-based theories (or indeed any other approach to virtue ethics) ought to be in the business of providing principles of action guidance. A rejection of this undertaking has been persuasively argued for by Julia Annas...”

reason. Doviak and Walsh have proposed interesting accounts of right action purporting to address this and related problems, but, ultimately, I think neither account is totally successful. Both accounts address only certain versions of problem cases, and neither account addresses all three problems—the problem of right action, the flouting of ‘ought implies can,’ and the problem of motivational overdetermination—in a satisfactory way. Should the above arguments be on track, neither account addresses the problem of right action; they solve only the problem of motivational overdetermination. If nothing else, I have shown that recent and promising agent-based views need work, and, perhaps more significantly, that agent-basing in general continues to conflict with commonsense morality. This should give us pause in considering it a plausible normative theory.¹⁸

References

- Athanassoulis N (2005) Common-sense virtue ethics and moral luck. *Ethical Theory Moral Pract* 8:265–276
- Brady M (2004) Against agent-based virtue ethics. *Philos Pap* 33:1–10
- Ciurria M (2012) A new mixed view of virtue ethics, based on Daniel Doviak’s new virtue calculus. *Ethical Theory Moral Pract* 15:259–269
- Das R (2003) Virtue ethics and right action. *Australas J Philos* 81:324–339
- Doviak D (2011) A new form of agent-based virtue ethics. *Ethical Theory Moral Pract* 14:259–272
- Driver J (1995) Monkeying with motives: agent-basing virtue ethics. *Utilitas* 7:281–288
- Garrard E (2000) Slote on virtue. *Analysis* 60:280–284
- Hurka T (2001) *Virtue, vice and value*. Oxford University Press, New York
- Hursthouse R (1999) *On virtue ethics*. Oxford University Press, New York
- Hursthouse R and Pettigrove G (2016) Virtue ethics. In: Edward N. Zalta, ed., *Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/ethics-virtue/>
- Jacobson D (2002) An unsolved problem for Slote’s agent-based virtue ethics. *Philos Stud* 111:53–67
- Johnson R (2003) Virtue and right. *Ethics* 113:810–834
- Kauppinen A (2014) Moral sentimentalism. In: Edward N. Zalta, ed., *Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/moral-sentimentalism/>
- Kawall J (2002) Virtue theory and ideal observers. *Philos Stud* 109:197–222
- McMahan J (2009) *Killing in war*. Oxford University Press, New York
- Russell D (2008) Agent-based virtue ethics and the fundamentality of virtue. *Am Philos Q* 45:329–347
- Sider T (2003) What’s so bad about overdetermination? *Philos Phenomenol Res* 67:719–726
- Slote M (1992) *From morality to virtue*. Oxford University Press, New York
- Slote M (2001) *Morals from motives*. Oxford University Press, New York
- Slote M (2010) *Moral sentimentalism*. Oxford University Press, New York
- Snedegar J (2013) Reason claims and contrastivism about reasons. *Philos Stud* 166:231–242
- Svensson F (2010) Virtue ethics and the search for an account of right action. *Ethical Theory and Moral Practice* 13:255–271
- Swanton C (2001) A virtue ethical account of right action. *Ethics* 112:32–52
- Van Zyl L (2009) Agent-based virtue ethics and the problem of action guidance. *J Moral Philos* 6:50–69
- Walsh JP (2016) Agent-basing, consequences, and realized motives. *Ethical Theory Moral Pract* 19:649–661
- Zagzebski L (2004) *Divine motivational theory*. Cambridge University Press, New York

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

¹⁸ I would like to thank Niklas Andersson, Anne Margaret Baxley, William Bell, Eric Brown, Chris Colacchia, and Ge Fang for helpful comments on this paper. I would also like to thank two anonymous reviewers for the journal for providing very detailed comments. Julia Driver deserves special thanks for reading multiple (and likely, not so great) drafts and brining important material to my attention.