10

Manifestations of Virtue

Arden Ali

Not all permissible acts are praiseworthy. Even an uncaring person might step in front of a drunk driver's vehicle to stop them from driving off and endangering the lives of others. But if we find out that this person stopped the drunk driver with hopes of being injured and securing an insurance payout, I doubt many of us would be inclined to praise him for his behavior. Similarly, a mean-spirited person might help her friend beat an alcohol addiction, but only because she needs leverage in their friendship. Such an act is permissible, perhaps even required, but the mean-spirited person doesn't appear to deserve praise.

The difficulties facing any virtue theory of permissibility are familiar. Even those who are sympathetic to virtue theory in general are skeptical about attempts to analyze deontic concepts in terms of virtuous traits of character.[1] Recently, an even deeper skepticism about virtue theory has emerged. These philosophers resist any attempt to analyze *praiseworthy* action in terms of the virtues.[2] Their objection is that the praiseworthiness of an action does not appear to depend on the virtuousness of the person who performs it. In this chapter, I examine this objection closely. I reject an ambitious reply from some virtue theorists who insist that some fully praiseworthy acts *do* in fact issue only from virtue. In its place, I argue that virtue plays an indirect but essential role in explaining moral worth. Though praiseworthy acts are not manifestations of virtue, these acts are manifestations of motivational states that are themselves inexplicable without an allusion to the virtues.

[1] For example, see Williams 1995 and Thomson 1997.
[2] When I say 'virtue' in this chapter, I mean 'virtuous disposition' unless otherwise specified.

Arden Ali, *Manifestations of Virtue* In: *Oxford Studies in Normative Ethics, Volume 10*. Edited by: Mark Timmons, Oxford University Press (2020).
DOI: 10.1093/oso/9780198867944.003.0011

In the first section, I describe the skeptical argument against the virtue theory of praiseworthiness and dismiss an ambitious reply to it. In the second and third sections, I argue that the core premise in the skeptical argument must be revised if the argument has any chance of success. In the fourth section, I bring out an implicit premise in the skeptical argument, ultimately showing that the newly uncovered assumption is questionable in the light of the revised premise from earlier sections. I conclude that the skeptical argument is far from insurmountable for the virtue theory of praiseworthiness.

1. The Skeptical Argument and the Ambitious Reply

Though they grant that acting from virtue is sufficient for full praiseworthiness, my opponents are doubtful that acting from virtue is necessary in order to warrant full praise for an act. Roger Crisp (2015) imagines a malicious gangster, Ronnie, who once saves someone from danger because he genuinely wishes to help. Although Crisp accepts the value of Ronnie's good motives, he wonders: "should we accept that Ronnie would have been *more* praiseworthy for acting in the way that he did had his action been based on a disposition? I think not" (2015: 14 his emphasis). Since Ronnie's act is fully praiseworthy, Crisp concludes that the virtues have nothing to add to an account of praiseworthiness.

The argument implicit in Crisp's example is this:

(P) A right act is fully praiseworthy if the act is motivated by the appropriate occurrent attitudes.

(C) So, the concept of virtue per se plays no role in the analysis of praiseworthy action.

Obviously, there are a number of steps elided in the argument above. But the core premise is stated explicitly: right actions may be fully praiseworthy if they are motivated by the appropriate occurrent attitudes. This is precisely the situation in Ronnie's case. We are meant to assume that his behavior is motivated by the appropriate occurrent

attitudes—namely, the desire to help someone in need—and he appears fully praiseworthy for his benevolent act. The fact that Ronnie is fully praiseworthy implies that possession of the relevant virtue of character is not necessary for full praiseworthiness. If virtue is not necessary for full praise, and a philosophical analysis of praiseworthiness should identify necessary (and sufficient) conditions for full praiseworthiness, then the concept of virtue cannot provide a philosophical analysis of praiseworthiness. So, Crisp concludes, the concept of virtue per se has no role to play in the analysis of praiseworthiness in general.

There is no shortage of examples to support the core premise in the skeptical argument. Thomas Hurka describes a cowardly soldier who jumps on a grenade in order to save the lives of his comrades on one occasion.[3] Hurka claims that such an act is fully praiseworthy even though the soldier lacks the virtue of bravery by stipulation.[4] Similar examples appear in Julia Markovits' account of moral worth, which differs from Nomy Arpaly's account precisely because Markovits denies the role of an agent's dispositions in the assessment of her actions.[5] Markovits imagines a fanatical dog-lover who saves a drowning stranger, but who would ignore the stranger if his dog were present.[6] The dog-lover clearly lacks virtue, but Markovits insists that his act is morally worthy as long as he is motivated by the right-making reasons on the occasion that he acts.

If this skeptical argument succeeds, virtue theory will have been almost fully ousted from ethical theory. The virtue theorist not only struggles to offer an account of acting *rightly*; they also cannot say much about acting *well*. In fact, the skeptical argument above is associated with doubts about the general moral significance of the virtues. Following G.E. Moore, many philosophers deny that the virtues bear

[3] This example appears in Hurka 2006: 72 and Hurka 2013: 12–13.

[4] In the same paper, Hurka (2006) describes walking down the street with a companion, who stops to give $20 to someone in need "apparently from concern for that person for his own sake" (71). We can imagine that Hurka's companion is generally unkind and mean to strangers and so this action may be thoroughly out of character. Hurka insists that his companion's act is fully praiseworthy. If his judgment is correct, praiseworthiness simply requires the right kind of motive; an agent's virtue is largely beside the point.

[5] See Arpaly 2002 and Markovits 2010 for this debate.

[6] See Markovits 2010: 210.

moral value intrinsically.[7] For Moore, when someone acts from a good motive, "it cannot be denied that the state of the man's mind, in performing [the act], contains something intrinsically good" (1903: 177). But, in the same breath, Moore denounced Aristotle's admiration for the virtues as a "gross absurdity" (1903: 176). He claimed that the exercise of a virtuous disposition "has, in general, no intrinsic value whatsoever" (1903: 176). Doubts about the role of virtue in moral philosophy are likely to follow: if the virtues are not themselves bearers of moral value, why should moral philosophers place any emphasis on them?

The skeptical argument therefore demands a response from virtue theorists. The ambitious reply involves simply rejecting the premise. In a discussion of related issues, Rosalind Hursthouse suggests her inclination to such an approach.[8] She imagines very wicked people who "pursue bad ends and characteristically do terrible things without a qualm" but who "may nevertheless be capable of actions that (at least apparently) are quite splendid" (1999: 146). Suppose one of these people does something "spectacularly self-sacrificial or unusually demanding" for apparently good reasons on an occasion. Would their act warrant full praise? Hursthouse is doubtful. To say that such people are morally motivated even on these special occasions conflicts with "our recognition of the fact that this man is wicked—how can we ascribe something as noble or high-minded as 'moral motivation' or a sense of duty to someone so vile?" (1999: 146). Hursthouse's remarks imply that someone who is not virtuous (at all) in some way cannot act well in the relevant regard, so having good occurrent motives is not sufficient for full praiseworthiness at least in some cases.

There is a reasonable challenge facing anyone who makes this ambitious reply. The ambitious reply amounts to the claim that an act merely performed from the appropriate occurrent motives is morally deficient. It says that the act would have been *more* praiseworthy had it issued

[7] Explicit doubts about the intrinsic moral value of the virtues appear in Moore 1903: 177, Sidgwick 1907: 393, Hooker 2002, Hurka 2006, Markovits 2010, and Crisp 2015. Crisp (2010: 23–4) says that a commitment to the intrinsic moral value of the virtues is the controversial claim of "explanatory virtue ethics."

[8] To be clear, Hursthouse isn't committed to this ambitious reply, but it is suggested by her remarks. At other points in the text, she appears to take the more modest approach discussed later in this chapter.

from virtue. Such a reply requires that the virtue theorist explain *why* occurrent motives are insufficient for full moral worth. Unsurprisingly, this challenge has been a tall order.

Consider Bradford Cokelet's attempt to meet this demand. He describes an agent named Randy, who does not value his son or their relationship very much but attends his son's soccer game on one occasion. Cokelet says that "the value and worth of Randy's showing up (for good reasons) will wax and wane with his overall values; his act of showing up will 'mean less' and be less good and praiseworthy, if he is a bad dad than if he is a good one" (2015: 240). Since his overall values are an embodiment of his character, Randy is more praiseworthy if he acts from virtue and not merely from good occurrent motives.

In my view, Cokelet's account fails to discharge the challenge facing the ambitious reply. He sets out to defend the limited claim that one's character can affect the ethical value of one's acts, and though he may secure his intended conclusion, his account does not respond adequately to the skeptic about virtue theory. To see why, attend to Cokelet's claim that the worth of Randy's act "is a function of its interpersonal meaning" (2015: 239). The value of attending his son's game is that it communicates Randy's support to his son. This message would not be communicated if Randy were generally unsupportive or only superficially supportive on this occasion. Similarly, the act of reaffirming or renewing one's wedding vows is valuable in virtue of its interpersonal meaning. The ethical value of the act may rest in the message it communicates to one's partner, who might otherwise feel anxiety and insecurity in the absence of the commitment. Cokelet's most persuasive examples follow this model.

But, if the relevance of virtue to moral worth is limited to cases of interpersonal meaning, the virtue theory remains on shaky ground. It is plausible that one's character is relevant to the ethical value of acts with interpersonal meaning because these acts would be *valueless* unless they issued from the agent's deep convictions and commitments. So, it is no surprise that Randy's act is more valuable if he is a good dad; being a good dad is required for the act to have value in the first place. In other words, Randy's act only has greater value if it issues from good character because it is the kind of act that wouldn't have *any* value if it didn't issue

from character in some sense. Similarly, it is not valuable to repeat one's wedding vows if the speech is not accompanied by broad dispositions to be devoted, committed, and loving to one's partner. In the absence of these dispositions, the reaffirmation of wedding vows appears ethically empty.

On Cokelet's view, then, virtue bears no *general* relationship to praiseworthiness. The virtues are relevant to praiseworthiness only when the value of the act in question depends entirely on the dispositions of the agent. If the virtues are relevant to moral worth because they are related to special communicative acts, the skeptic may still make a strong case against the virtue theory of praiseworthiness. They may argue that the concept of virtue needn't appear in the analysis of praiseworthy action, even though it must appear in any view about the moral significance of special communicative acts. In other words, it may be constitutive of some ethically-valuable acts that they express someone's virtuousness and so it may be that someone couldn't perform those special acts without being virtuously disposed to some extent. But this doesn't show that virtue is related to praiseworthiness; it merely shows that some ethically valuable acts are unavailable to those who lack the relevant disposition entirely.

Cokelet does attempt to extend his account beyond cases of interpersonal meaning. He claims that the value of giving $20 to someone in need is diminished if the agent undervalues the loss of $20. He says: "if the person has been inducted into a strange religious cult, disvalues having money in her wallet, and is therefore disposed to give her money away to anyone who asks, then her act is less generous than it would be if she values what she gives away" (2015: 242–3). In this case, the agent who gives up $20 does not properly value the money; her valuation of the cash is uncalibrated to its actual value. But one wonders whether her act does not seem praiseworthy because it is not performed from the right kind of occurrent motive, e.g. the desire to help a stranger by giving up something valuable that (she believes) the stranger can effectively use. It is unclear that any member of the strange cult has this kind of motive for giving up cash. She does not value the item highly, nor does she apparently think it is very useful to have money. So, although we may want to withhold praise from such a person, this intuition can be

explained simply in terms of her occurrent motives. We needn't appeal to virtue to understand the case.

2. Deviance Problems for the Core Premise

I am going to pursue a different line of response. In this section (and the next section), I will show that the core premise in the skeptical argument requires significant revision. Later, I will show that an implicit assumption in the argument looks suspicious in the light of the revised core premise. To begin, recall the reasoning in the skeptical argument: a right act is fully praiseworthy if the act is motivated by the appropriate occurrent attitudes, so virtue per se has no role in the analysis of praiseworthy action.

The single premise in the argument above doesn't depend on any theory about the precise propositional content of the appropriate occurrent attitudes. If one is inclined to Kant's view of moral worth, then the appropriate propositional contents may involve direct reference to moral rightness or duty. If one is inclined to recent alternatives to Kant, then the appropriate propositional contents may specify the facts in virtue of which the act is morally right.[9] Nothing should hang on this choice about the content of the appropriate occurrent attitude.

Although the core premise doesn't take a stand on the content of the appropriate occurrent attitudes, it does obviously require that there are such things as occurrent attitudes. There is significant debate about the nature and intelligibility of these mental states. Some philosophers identify occurrent attitudes with mental states that an agent is presently-entertaining, i.e one's belief that Aristotle tutored Alexander became occurrent when reading this very sentence.[10] Others treat occurrent attitudes as active mental states, i.e. one's desire to buy cheese may be activated when in the grocery store but may otherwise remain inactive.[11] Neither of these uses of 'occurrent' seem to match the use of the term in the skeptical argument. Opponents of the virtue theory seem to contrast

[9] See Arpaly 2002 and Markovits 2010.
[10] See Braddon-Mitchell & Jackson 2006: 145 for this description.
[11] See Goldman 1970: 87–8.

occurrent states with dispositional states. On such a view, an occurrent attitude is a token mental state that could be ascribed independently of any dispositional properties of the agent. The presence of an occurrent attitude in this sense doesn't require the presence of any dispositions in the agent. In short, as I understand it, an occurrent attitude is an entirely non-dispositional one. The most familiar version of an occurrent attitude in this sense is the representational conception of belief, according to which someone believes a proposition just in case there is a representation with the content of the proposition in her mind (typically, in the 'belief box').

One needn't think that all occurrent attitudes in this sense are beliefs. There may be occurrent desires, too. An occurrent desire is akin to what Michael Smith and Philip Pettit (1990) call a 'foregrounded' desire. For them, a desire "figures in the foreground if and only if the agent reaches [a] choice via the recognition that he has that desire and that [his chosen] option has the desirable property…of promising to satisfy the desire" (Smith & Pettit 1990: 567–8). Importantly, Smith and Pettit do not think the relevant kind of recognition is always conscious; they say: "a desire may be in the foreground, as in implicit deliberation, without being consciously considered" (Smith & Pettit 1990: 568). They hold that foregrounded desires appear "somewhere in the process leading to action" and thus have the property of "engaging deliberation" (1990: 566, 568). To say that someone acts from a foregrounded desire is to say that this desire appears in the best representation of the deliberative episode that led to their action.

This gives an initial characterization of agents like Crisp's Ronnie or Markovits' dog-lover. These agents have an attitude with the appropriate propositional content tokened in their mind or foregrounded in their deliberation (e.g. 'I think someone is in dire need of help!' or 'I really want to help that person!') and this attitude moves them to act on the occasion when they help someone in need. Crucially, their attitude is occurrent; their possession of the relevant attitudes does not depend on the truth of any dispositional claims about them.

Though I think this is one natural way to interpret talk of occurrent attitudes, it seems plainly clear that the core premise in the skeptical argument is false when 'occurrent attitude' is interpreted in this way.

If skeptics do use the term in this way, they cannot insist that acts performed from the appropriate occurrent attitudes are fully praiseworthy. This is because any occurrent attitude may trigger action deviantly or play the wrong role in deliberations. Consider the following simple case:

Suppose that people who are praiseworthy for telling jokes are motivated by an occurrent desire to make other people laugh. Alp may have the desire to make you laugh and his recognition of this desire may make him nervous and uptight. When he notices his nervousness, and sees how badly he wants to impress you, he might anxiously repeat a joke, even though he thinks that the joke is not worthwhile and has committed himself to never repeating it. The best representation of the deliberative episode that led to Alp's behavior includes the desire to make you laugh. In fact, if you inquired about why he told the joke, Alp might say that he told the joke because he wanted to make you laugh. It was precisely this thought that set him on the course to telling the joke. But, although the desire to make you laugh moved Alp, this desire moved him deviantly. That is, the appropriate desire moved Alp via his nervousness. Given his commitment to never repeating the joke, Alp seems to have made a mistake in telling the joke on this occasion. His behavior is apparently *accidental* in the sense that usually precludes praiseworthiness. Since I doubt that my opponents want to allow that accidents in this sense are fully praiseworthy, the core premise in the skeptical argument cannot be true as stated.

The problem raised in the example above is not simply the general problem of causal deviance that worried philosophers like Donald Davidson.[12] The kind of deviance at issue here needn't involve any interference from the agent's emotional or physiological states. The desire to make someone else laugh (or belief that telling a joke will make someone else laugh) might move Caligula to tell someone a joke. If Caligula tells a joke, his desire to make someone else laugh might figure in the best representation of the deliberative episode that led to his act. But suppose Caligula is moved in this way only because he also knows that laughing will cause this person tremendous pain. If the appropriate desire in this case is the desire to make someone laugh, Caligula is

[12] See Davidson 1973.

moved by the appropriate occurrent desire. But no one wants to allow that Caligula's act is fully praiseworthy. Thus, occurrent motives of this sort are not sufficient for full moral worth and the core premise in the skeptical argument is false (again, as stated).

Notice that it makes no difference to change the content of the appropriate attitude in Caligula's case. Even if the appropriate desire were different—for example, suppose it were the desire to lighten someone's mood—the problem would remain. Imagine that Caligula acts from the desire to lighten my mood, but only because he knows that I will make a terrible moral mistake if my mood is elevated. My opponents do not want to allow that such a person is fully praiseworthy, so they must revise the premise in their argument.

One might revise the model by adding the further stipulation that all praiseworthy actions are motivated by *foregrounded desires that are also noninstrumental*. Noninstrumental desires are held for their own sake; they do not owe their existence to the presence of any other desire. Caligula's desire to make someone else laugh is paradigmatically instrumental. He wants to make someone else laugh because he desires to cause someone else pain and believes that telling the joke is a means to achieving this end.

It is hard to make sense of the idea that there could be noninstrumental desires that are occurrent in the sense under consideration here. But, putting this worry aside, it seems that the current proposal is too conservative in any case. Some praiseworthy acts are performed by people for whom the appropriate noninstrumental desire is missing from the deliberative episode preceding their act. One such case has figured prominently in the literature on moral worth. Mark Twain's Huckleberry Finn struggles to decide whether he ought to help Jim escape slavery or write a letter to Jim's captors to tell them where to find him.[13] Huck supposedly believes that it is wrong to help Jim escape slavery, but he does so anyway out of respect for Jim's personhood. Those who discuss the case tend to claim that Huck's behavior is praiseworthy even though

[13] See especially Chapter XVI of Mark Twain's *Huckleberry Finn*. The example is discussed in detail in Bennett 1974, Arpaly 2002, Markovits 2010, and Sliwa 2015.

Huck himself believes he is doing something terribly wrong.[14] Even further, Huck supposedly acts out of respect for Jim's personhood even though Huck would not assert that Jim is a person and couldn't be said to believe this proposition. This means that the desire to respect Jim's personhood cannot figure in the foreground of Huck Finn's decision to help Jim escape slavery. Recall that a desire "is present in the foreground of the decision if and only if the agent believed he had that desire and was moved by the belief that a justifying reason for the decision was that the option chosen promised to satisfy that desire" (Smith & Pettit 1990: 568). Huck seems unaware that he wants to respect Jim's personhood. It is unlikely that Huck's moral knowledge is sophisticated enough to attribute such an attitude to him. Further, he certainly does not think that this desire justifies helping Jim escape. In fact, Huck believes that helping Jim escape slavery is morally unjustified.

If Huck Finn is praiseworthy for helping Jim, then the revised model fails to vindicate the core premise. On this revised model, praiseworthy actions are motivated by the appropriate foregrounded noninstrumental desire. But Huck Finn is supposed to show that someone can be praiseworthy for what they do even if the appropriate noninstrumental desire is missing from the deliberative episode preceding the act. This means that the core premise is false on the revised model under consideration and so opponents of the virtue theory cannot help themselves to it.

3. Delusion Problems for the Core Premise

So far, I have argued that being motivated by the appropriate occurrent attitude cannot secure that an agent deserves praise for their right act. This is because the appropriate occurrent attitude (i.e. the occurrent attitude with the appropriate content) may move the agent to act in deviant ways or it may play the wrong role in her deliberations. To salvage the skeptical argument, my opponent might now introduce the idea of

[14] There are some doubts about Huckleberry Finn's praiseworthiness. See, for example, Sliwa 2015. Still, I take it that it would be uncomfortable for my opponents to rest too much of their case against the virtue theory on the rejection of the standard verdict in the Huck Finn case.

motivating reasons for acting. The appropriate occurrent attitude appears in Caligula's deliberation, i.e. he desires to make someone laugh, but it is far from clear that Caligula's *motivating reason* for telling the joke is that it will make someone else laugh. Plausibly, his motivating reason for telling the joke is more cruel: telling the joke will cause his intended audience extreme pain. So, as long as they may help themselves to the concept of a motivating reason, the skeptic can maintain her argument against the virtue theory.

What does it mean to take some consideration as one's reason for acting? There are two uses of the term 'motivating reason' in the literature. First, some philosophers use this term to refer to the consideration that rationalizes an agent's behavior from an explanatory perspective. Call this the *thick* conception of motivating reasons. I will set this view aside until the second half of this section because nearly everyone agrees that thick motivating reasons are not mere occurrent attitudes.

Second, some philosophers use this term to refer to the consideration "in the light of which the agent did that action" (Dancy 2000: 1). Call this the *thin* conception of motivating reasons. In this sense, thin motivating reasons are the "considerations that someone took to count in favor of an action, whether or not they actually count in favor of it—those considerations someone *treated as* 'normative' reasons" (Hieronymi 2011: 411). T. M. Scanlon (1998: 19) uses the term 'operative reasons' to refer to this concept; Jonas Olson and Frans Svensson (2005: 205-6) chose the term 'deliberative reason' for it; and John Hyman says these are facts that "someone is said to have been guided by" and which the agent "took [] into consideration, when he modified his thought or behavior in some way, or decided what to think or what to do" (2011: 361). With thin motivating reasons in mind, Kieran Setiya highlights two key properties:

> The first is that we choose the reasons on which we act. There are many reasons for which I might decide to write a book: personal satisfaction, a fragment of immortality, professional ambition. I am not passive in the face of this: even if I believe that books give their authors a kind of immortality, and even if I think that this is a reason—a good

> reason—to write a book, it may not be *my* reason for doing so. That is up to me… (Setiya 2007:39).
>
> The second insight is that we know without observation not only *what* we are doing, but *why*. Just as I can only do intentionally what I think I am doing…I know what my *reasons* are without having to find out. I don't have to ask myself why I am walking to the shops, if my reason for doing so is to buy a hammer. (If I do have to ask myself, and no answer can be found, my action has become detached from reasons; I am doing it aimlessly, if I am doing it intentionally at all.)
>
> (Setiya 2007: 40).

The thin sense of motivating reasons is the only sense of the term for which 'acting for a motivating reason' involves only an occurrent attitude. On the simplest view, I do φ for the thin motivating reason that p just in case I believe that I am doing φ for the reason that p. Setiya's own view is not unlike such an account: "[to] take p as one's reason for doing φ is to have the desire-like belief that one is doing φ for the reason that p" (2007: 46).

Could such an account vindicate the argument against virtue theory? I don't think so. The trouble for such an account is not deviance, but *delusion*. The content of someone's thin motivating reason is fixed by that person's belief about what she is doing and why. If these beliefs are occurrent, then they are not grounded in any dispositional properties of the agent. So, it is possible that someone could believe she were acting because p, and so the fact that p may be her thin motivating reason for acting, but her broader dispositions may show that she is totally deluded about her motivations. Markovits acknowledges the possibility of this kind of delusion: "If people were always perfectly self-aware and sincere, their account of what prompted them to choose to act as they did would always provide us with their [actual motives]. But people can, of course, be self-deceptive about their motivations" (2010: 222). A selfish person might perform an act of charity, and be convinced that she acts out of moral concern, but it might be apparent that she is moved only because her interests happen to be served from altruism in the present case. Perhaps the agent would not have acted charitably if it were not in her

selfish interests to do so. If she deeply and sincerely wants to believe that she is a charitable person, she might believe she is acting for the morally appropriate reasons, and thus she may have thin motivating reasons with the appropriate propositional content. I may want to believe that I am a good person, and so convince myself that I am motivated to recycle out of concern for future generations or those most affected by resource waste. But it takes no stretch of the imagination to suppose that I am entirely deluded about the goodness of my character. In other words, people can form beliefs about why they are acting—and so technically act for those thin motivating reasons—even though they appear moved by other considerations entirely. The upshot is that deluded people may have thin motivating reasons with the appropriate content, but they are certainly not praiseworthy for their act. So, if the skeptic's premise is understood as a claim about thin motivating reasons, the core premise is plainly false.

There is room to cast doubt on the possibility of the kind of delusion at work in the example. Is it possible to be delusional about one's own motivations? It may be impossible to delude oneself about one's motivations intentionally. If someone intentionally deludes himself about the facts that motivated him to act, then he must have knowledge of the facts that actually motivated him to act, since otherwise it wouldn't be possible to intentionally conceal them.[15] But the example doesn't require that the agent intend to delude himself. The kind of delusion at work in the example may be a simple case of motivationally biased belief. An agent's beliefs can be motivationally biased by their desires, wishes, anxieties, and emotions without involving their capacity for intentional action.[16] This is enough to establish the delusional case described above.

What is left for the core premise in the skeptic's argument? If the premise used in the skeptical argument against virtue theory is true, then my opponents do not appear to have occurrent mental attitudes in mind at all. Markovits notes that someone acts from good motives when the appropriate consideration appears in the rationalizing explanation

[15] For this point, see Paluch 1967. For a discussion of ways around this problem, see Bermúdez 2000.

[16] See Johnston 1988 and Mele 2001.

of her action.[17] In making this claim, she adopts the *thick* conception of motivating reasons. As commonly understood, thick motivating reasons are discerned from what R. Jay Wallace calls the *third-personal explanatory perspective*:

> Here we are contemplating an action that has already been performed, and asking why it was done; in posing this question, we abstract from the immediate deliberative horizon of the agent, and adopt a standpoint that brings the agent herself into view, as an object of reflection.
>
> (Wallace 2003: 432)

One takes the third-personal explanatory perspective in order to identify the considerations that moved someone to act, granting that those considerations may not appear in her first-personal reflective or deliberative thought. From this perspective, one can determine why someone acted in the light of the agent's broader self, which includes the agent's cognitive, conative, and affective attitudes at the first-personal, sub-personal, or non-personal levels. But this way of explaining someone's behavior—that is, the attribution of thick motivating reasons—cannot be done merely by appeal to occurrent states. When attributing thick motivating reasons, one attributes clusters of attitudes that are intricately related to one another in functional relationships, which typically grounds familiar dispositional ascriptions. As a result, an explanation from the third-personal explanatory perspective involves the attribution of dispositions that arise from the relationship between the agent's broader attitudes.

Skeptics about the virtue theory have something like the third-personal explanatory perspective in mind when they talk about good motivation. This is why the third-personal explanatory perspective solves the problems of delusion and deviance that plague the earlier proposals. A selfish person can delude themselves into thinking that they are altruistically

[17] She says: "I cannot be using the term 'motivating reason' as it is often used—to pick out (exclusively) belief-desire pairs. I propose that motivating reasons are the kinds of facts we are after when we ask about an agent, 'what were her reasons for acting as she did?'—those that appear in what have been called 'rationalizing explanations' " (Markovits 2010: 221).

motivated, but their broader disposition will bear their selfishness on its face. It will plainly show that the selfish person is disposed to act altruistically only when their interests happen to align with the altruistic thing to do. Moreover, Caligula's disposition precludes him from having good motives. He is disposed to tell the joke when it will cause his listener pain; if he were acting from good motives, he would be disposed to tell the joke when it will make his listener laugh.[18]

This admission has implications for the core premise in the skeptical argument. If the skeptic abandons the idea that praiseworthy action issues from occurrent mental states, the core premise in her argument must be revised to reflect this change. The revised skeptical argument looks something like the following:

(P) A right act is fully praiseworthy if the act is the manifestation of the appropriate broad dispositional state of the agent. Call this disposition *S*.
(C) So, the concept of virtue per se plays no role in the analysis of praiseworthy action.

The skeptic can no longer differentiate herself from the virtue theorist with the vague appeal to occurrent mental attitudes. Instead, she must be explicit about her reliance on the third-personal explanatory perspective, and the use of functional and dispositional forms of explanation, in giving her account of acting well. The acts that deserve praise are the manifestation of the agent's broader dispositions, even for those who are opponents of the virtue theory.

4. A Modest Reply to the Skeptical Argument

On my most charitable reading of my opponent's premise, acting well involves the manifestation of a special disposition. In the revised skeptical argument, I used disposition *S* to refer to this special

[18] See Hyman 2013 and Setiya 2016: 17 for the idea that dispositions may be useful in avoiding deviance generally.

disposition. This admission changes the landscape for the argument against the virtue theory. We are no longer comparing someone who acts from a disposition against someone who acts from an occurrent attitude. Rather, the relevant comparison is someone who manifests the virtuous disposition and someone else who manifests disposition *S*. With this in mind, the inference in the argument against virtue theory begins to look suspicious. Here's why:

If disposition *S* is supposed to explain praiseworthiness, it must be possible to give some general characterization of it. This disposition has at least two features. First, disposition *S* is broad enough that it rules out cases of delusion and deviance. Second, disposition *S* deviates from virtue in some significant regard; it is narrower, more ephemeral, or less robust than the virtuous disposition.[19] The task of balancing these features is far from insurmountable. But I want to insist that this task cannot be completed without an allusion to virtue as an indispensable guidepost.

It is not obvious that disposition *S* is inherently significant. Without more, any candidate specification of disposition *S* may appear arbitrary. What is inherently significant about acting from a disposition that falls short of virtue in *x* ways? Why does the disposition that falls short of virtue in *x* ways have significance while other dispositions (for example, those that fall short of virtue in *y* or *z* ways) do not? These remarks highlight an implicit premise in the skeptical argument as presented so far:

(P1) A right act is fully praiseworthy if the act is the manifestation of the appropriate broad dispositional state of the agent. Call this disposition *S*.

(P2) It is possible to explain the moral significance of disposition *S* without reference to the concept of virtue. That is, virtue does not have explanatory priority over disposition *S* in the analysis of praiseworthy action.

[19] Some believe the virtues may be quite fine-grained. Social psychologists have noted that people exhibit only highly contextualized dispositions or 'local' traits' (Doris 2002: 64). In light of the data, these psychologists and philosophers have suggested that we replace attributions like 'courageous' with attributions like 'sailing-in-rough-weather-with-one's-friend-courageous' (Doris 2002: 115). If virtues are narrow in this way, then the dispositions involved in acting from good motives must be *even narrower*. Nothing said here hinges on the situationist critique of virtue.

(C) So, the concept of virtue per se plays no role in the analysis of praiseworthy action.

The implicit premise says that it is possible to explain disposition *S* without appeal to the concept of virtue. This assumption is required to reach the final skeptical conclusion. But what can be said to defend it? This first option is to claim that disposition *S* has brute moral significance. But then my opponent's account is empty. It may be true that one can analyze praiseworthiness in terms of manifestations of disposition *S*, but this explanation is possible only because disposition *S* has been given a circular characterization. It is simply *stipulated* that disposition *S* is the disposition that one must manifest in order to warrant praise.

The second option is more appealing. My opponents may grant that disposition *S* is not inherently significant, but they may insist that it derives its significance from its relationship to other normative concepts. For example, one might claim that disposition *S* is responsive to the right-making normative reasons.[20] In other words, one might say that disposition *S* embodies sensitivity to the facts in virtue of which an action is morally right. Now the charge of arbitrariness reemerges. The relevant virtuous disposition presumably *also* embodies sensitivity to the right-making facts.[21] If disposition *S* falls short of virtue, then it somehow embodies *less* sensitivity to the right-making facts when compared to the virtuous disposition. This could mean that disposition *S* is simply narrower in breadth than virtue; while the virtuous person is disposed to act whenever the right-making features obtain, the person who manifests disposition *S* is disposed to act only in some of those cases. But, of course, there are *many* dispositions narrower or less sensitive in this regard than the virtuous disposition, so again it becomes arbitrary to insist that disposition *S* is *the* morally significant disposition for praiseworthiness.

[20] This is the route most appealing to Markovits (2010).

[21] After all, the difference between the virtuous agent and the merely praiseworthy agent is not that the former responds to *different* right-making features than the latter. If there were truly two sets of right-making features, it would be unclear why it is apparently true that "morally worthy actions are the building blocks of virtue" (Markovits 2010: 203). That is, if the virtuous person and the merely praiseworthy person respond to entirely different right-making features, then it would be inexplicable why people usually develop the virtues *through* patterns of morally worthy behavior.

The problem noted above will follow my opponents if they explain the derivative significance of disposition *S* in terms of morally good backgrounded desires or any other attitude. If the significance of disposition *S* is that it embodies a morally good backgrounded desire, and the virtuous person shares these morally good desires, then it must be true that disposition *S* embodies the desire to some *lesser* extent than the virtuous disposition (or, more likely, disposition *S* embodies a *weaker* desire than the desire embodied by the virtuous disposition). But presumably many dispositions embody morally good desires to a lesser extent than the virtuous disposition; why should disposition *S* have special significance?

Notice that one cannot rely on first-order moral theory to explain why disposition *S* is significant. As far as I can tell, moral theory tells us something about *virtue*, not disposition *S*. If utilitarianism is true, then virtue is the disposition to maximize utility.[22] If Rossian pluralism is true, then virtue consists in the seven distinct dispositions that embody respect for each of Ross' prima facie duties.[23] It is not clear how one is supposed to fix the parameters for disposition *S* in the light of these characterizations of virtue. So, it cannot be assumed that the moral significance of disposition *S* will be fixed by first-order moral theory.

Admittedly, there may be space for a moral theory that distinguishes *fundamental* right-making reasons and *non-fundamental* right-making reasons. Suppose that the fundamental reason to save another person's life is that doing so treats them as an end in themselves. Suppose there are also non-fundamental reasons why someone should save the person, e.g. 'she is my best friend' or 'she is in danger.' There would be no puzzle about the moral significance of disposition *S* if the correct moral theory provided two sets of reasons in this way. One could simply say that disposition *S* embodies respect for the non-fundamental reason to act while the virtuous disposition embodies respect for the fundamental reason to act. The significance of both dispositions is explained by this kind of moral theory.

This kind of moral theory is certainly possible. But it will be incumbent on any such theory to explain why the virtuous disposition and

[22] This point is made in Foot 1983. [23] Ross 1930.

disposition *S* diverge so drastically. It is hard to see why they should. Any rationale for treating disposition *S* as responsive to less fundamental right-making reasons will apply equally well to the virtuous disposition. Suppose the divergence is a response to concerns about an agent's having one-thought-too-many.[24] It would be oddly detached or impersonal to save another person in order to treat them as an end in themselves, so it would be a mistake to insist that praiseworthy agents act in response to such fundamental reasons. But why then should the virtuous person respond to those reasons? If the fundamental reasons are odd enough to qualify as unsuitable for the merely praiseworthy agent, why should they be suitable for the virtuous agent? There may be ways to answer these queries. But it is hard to see how my opponent can reply to them without assigning some special moral significance to acts performed from virtue.

The virtue theory of praiseworthiness can avoid this predicament. Earlier I said that Hursthouse may be prepared to accept that fully praiseworthy actions must issue from virtue. But elsewhere she appears prepared to accept a more modest claim:

> what is both necessary and sufficient for a virtuous act to be 'morally motivated' is that it is done from a state of character that adequately *resembles* the state of character from which the perfectly virtuous agent acts. The central idea, to repeat what I said at the outset, has been that ascribing 'moral motivation'…is ascribing something that goes far beyond the moment of action. It makes a claim about what sort of person the agent is—a claim that goes all the way down.
>
> (Hursthouse 1999: 123 my emphasis).

If one understands the moral significance of disposition *S* in terms of its relation to virtue, one can rely on virtue theory to explain disposition *S*. Suppose one takes the position that people who act well manifest a disposition that *approximates* or *resembles* virtue. The specific kind of approximation or resemblance is a substantive issue. It may be that disposition *S* is *ephemeral* virtue; someone who acts praiseworthily is

[24] See Markovits 2010: 226–30.

momentarily disposed like the virtuous person. It may be that disposition *S* is *prodromal* virtue; someone who acts praiseworthily is disposed like someone on the cusp of virtue. It may be that disposition *S* is *nascent* virtue; someone who acts praiseworthily is committed to developing virtue. Or, if the virtues are modular or decomposable into discrete areas of mastery, it may be that disposition *S* is simply *partial* virtue; someone who acts praiseworthily has a disposition that reflects her achievement of some important stage in the development of full virtue. There are many possible views, but we needn't survey all of them here. The important point is that none of these proposals can be charged with arbitrariness. That is, if disposition *S* is ephemeral, prodromal, nascent, or even partial virtue, its moral significance is explicable and non-arbitrary. One can argue that disposition *S* characterizes an important or noteworthy stage in the development of full virtue, and since the virtuous disposition has moral significance, disposition *S* also has (derivative) moral significance. There is no lingering puzzle about the special status of disposition *S*.

This gives us reason to take the virtue theory of praiseworthiness seriously. At the very least, it gives us reason to doubt that the second premise in the skeptical argument is true. The second premise in the argument proposes that disposition *S* is the fundamental explanatory notion in the analysis of praiseworthy action. But disposition *S* doesn't appear to have brute significance and the obvious proposals to explain its significance seem either arbitrary or unmotivated.

5. Conclusion

In the preceding section, I proposed that someone acts well when she approximates the relevant virtue in some respect. I have not taken a stand on the nature of this approximation or resemblance. This view becomes attractive when the weaknesses of the skeptical argument are made explicit. If the core premise in the argument requires the revision described in sections two and three, then the implicit premise in the argument—namely, the assumption that virtue is not explanatorily prior to disposition *S*—begins to look suspicious.

The upshot is that the virtues cannot drop out of the analysis of praiseworthy action, even for those that prefer to speak in terms of good motives. When I say that the virtues must appear in our analysis of praiseworthiness, I am making a claim about *explanatory priority*. This means that the final analysis of praiseworthiness should appeal to virtue rather than good occurrent attitudes. Views about explanatory priority are familiar. When consequentialists take the position that goodness is prior to rightness, they take a stand on the explanatory priority of goodness over rightness. Timothy Williamson's 'knowledge first' approach in epistemology is best understood as the claim that knowledge has explanatory priority over other epistemological notions.[25] When Michael Dummett says that language is prior to thought, he too makes a claim about explanatory priority.[26] The claim is this chapter is meant to have the same force as these more familiar positions in analytic philosophy.

It should be noted that the argument offered in this chapter is consistent with neutrality about the epistemological priority between virtuous dispositions and praiseworthy acts. One could hold that all praiseworthy acts must be identified by first identifying the relevant virtue, or, on the other hand, one could hold that all virtuous dispositions must be identified by first identifying relevant cases of praiseworthy conduct. I suspect an intermediate (albeit perhaps unsatisfying) position on this issue is correct. In some cases, our grasp of the relevant virtue may be clear, so it will make most sense to identify praiseworthy acts by first identifying the relevant virtue. In other cases, when our grasp of virtue is less complete, the opposite strategy may be more appropriate. This is the view that Gopal Sreenivasan calls the 'modest agent-centered view,' according to which "some [virtuous] acts can be identified as [virtuous] without reference to any [virtuous] person, while other [virtuous] acts *cannot* be so identified except by reference to a [virtuous] person" (2017: 254–5 his emphasis).

I should make a final remark about the cases that motivate the argument against virtue theory. Crisp insists that the malicious gangster, Ronnie, can be fully praiseworthy for an act of kindness. This possibility

[25] Williamson 2000: 2–5. [26] Dummett 1991: 315.

is left open by my argument here. I wanted to show only that Crisp has no grounds to rule out the basic virtue-theoretic explanation of Ronnie's praiseworthiness. According to this explanation, Ronnie is praiseworthy because he manifests a disposition resembling virtue. The same goes for Hurka's cowardly soldier, who approximates the virtue of bravery despite not possessing it, and Markovits' dog-lover, who also approximates bravery in some salient respect when he saves a drowning stranger. Is it plausible that any of these agents resemble or approximate virtue? I think so, though I suspect that the kind of resemblance may differ in each case. Hurka's cowardly soldier may be an example of ephemeral or short-lived virtue; Crisp's malicious gangster may show nascent or prodromal virtue; and Markovits' dog-lover may have nearly mastered virtue in the relevant respect. In principle, the virtue theory allows for such pluralism about moral praiseworthiness. That is, if the virtue theory is correct, there may be many different but equally credible grounds for attributing moral worth to someone's actions. A proper virtue theory of praiseworthiness must take a stand on the exact nature of all the ways that praiseworthy actions can resemble or approximate virtue. I cannot complete this work in the space that remains. But I hope to have shown that the common skeptical argument does not justify the wholesale rejection of this approach to moral worth.[27]

References

Arpaly, Nomy. 2002. "Moral Worth." *Journal of Philosophy* 99 (5): 223–45.

Bermúdez, José Luis. 2000, "Self-Deception, Intentions, and Contradictory Beliefs." *Analysis* 60 (4): 309–19.

[27] For help and encouragement, special thanks to Nomy Arpaly, Olivia Bailey, Nashid Chaudhury, Brad Cokelet, Nilanjan Das, Sanford Diehl, Ryan Doody, Lyndal Grant, James Harold, Sally Haslanger, Richard Holton, Brendan de Kenessey, Joshua Keton, Zoe Johnson King, Doug Kremm, Julia Markovits, Sam Mitchell, Kristi Olson, Sofia Ortiz-Hinojosa, Milo Phillips-Brown, Bernhard Salow, Scott Sehon, Kieran Setiya, Keshav Singh, Mark Timmons, Katia Vavova, Kenneth Walden, Stephen Yablo, and an anonymous reviewer. Thanks also to audiences at the Workshop in Normative Ethics at the University of Arizona, DePauw University, Mount Holyoke College, MIT, Australian Catholic University, and the City University of New York Graduate Center.

Bennett, Jonathan. 1974. "The Conscience of Huckleberry Finn." *Philosophy* 49 (188): 123–34.

Braddon-Mitchell, David & Frank Jackson. 2006. *Philosophy of Mind and Cognition*. 2nd ed. Oxford: Blackwell.

Cokelet, Bradford. 2015. "Dispositions, Character, and the Value of Acts." In *Character: New Directions from Philosophy, Psychology, and Theology*, edited by Christian Miller, R. Michael Furr, Angela Knobel, and William Fleeson. Oxford: Oxford University Press.

Crisp, Roger. 2010. "Virtue Ethics and Virtue Epistemology." *Metaphilosophy* 41 (1/2): 22–40.

Crisp, Roger. 2015. "A Third Method of Ethics?" *Philosophy and Phenomenological Research* 90 (2): 257–73.

Dancy, Jonathan. 2000. *Practical Reality*. Oxford: Oxford University Press.

Davidson, Donald. 1963. "Actions, Reasons, and Causes." *Journal of Philosophy* 60 (23): 685–700.

Davidson, Donald. 1973. "Freedom to Act." In *Essays on Freedom of Action*, edited by Ted Honderich. London: Routledge and Kegan Paul.

Doris, John M. 2002. *Lack of Character: Personality and Moral Behavior*. Cambridge: Cambridge University Press.

Dummett, Michael. 1991. "The Relative Priority of Thought and Language" in his *Frege and Other Philosophers*. Oxford: Oxford University Press.

Foot, Phillippa. 1983. "Utilitarianism and the Virtues." *Proceedings and Addresses of the American Philosophical Association* 57 (2): 273–83.

Goldman, Alvin. 1970. *A Theory of Human Action*. Englewood Cliffs, NJ: Prentice-Hall.

Hooker, Brad. 2002. "The Collapse of Virtue Ethics." *Utilitas* 14 (1): 22–40.

Hurka, Thomas. 2006. "Virtuous Act, Virtuous Dispositions." *Analysis* 66 (289): 69–76.

Hurka, Thomas. 2013. "Aristotle on Virtue." In *Aristotelian Ethics in Contemporary Perspective*, edited by Julia Peters. New York: Routledge.

Hyman, John. 2011. "Acting for Reasons: Reply to Dancy." *Frontiers of Philosophy in China* 6 (3): 358–68

Hyman, John. 2013. "Desires, Dispositions and Deviant Causal Chains." *Philosophy* 89: 83–112.

Hieronymi, Pamela. 2011. "Reasons for Action." *Proceedings of the Aristotelian Society* 111: 407–27

Hursthouse, Rosalind. 1999. *On Virtue Ethics*. Oxford: Oxford University Press.

Johnston, Mark. 1988. "Self-Deception and the Nature of Mind." In *Perspectives on Self-Deception*, edited by B. McLaughlin and A. O. Rorty. Berkeley: University of California Press.

Korsgaard, Christine M. 2008. *The Constitution of Agency: Essays on Practical Reason and Moral Psychology*. Oxford: Oxford University Press.

Lewis, David. 1979. "Counterfactual Dependence and Time's Arrow." *Noûs* 13: 455–76.

Markovits, Julia. 2010. "Acting for the Right Reasons." *Philosophical Review* 119 (2): 201–42.

Mele, Alfred. 2001. *Self-Deception Unmasked* Princeton: Princeton University Press.

Moore, George Edward. 1903. *Principia Ethica*. New York: Dover Publications.

Olson, Jonas & Frans Svensson. 2005. "Regimenting Reasons." *Theoria* 71 (3): 203–14.

Paluch, Stanley. 1967. "Self-Deception." *Inquiry* 10: 268–78.

Ross, William David. 1930. *The Right and the Good*, edited by Philip Stratton-Lake. New York: Oxford University Press.

Scanlon, Thomas. 1998. *What We Owe to Each Other*. Cambridge: The Belknap Press of Harvard University Press.

Setiya, Kieran. 2007. *Reasons without Rationalism*. Princeton, NJ: Princeton University Press.

Setiya, Kieran. 2016. *Practical Knowledge: Selected Essays*. Oxford: Oxford University Press.

Sidgwick, Henry. 1907. *The Methods of Ethics*. 7th ed. London: Macmillian.

Sliwa, Paulina. 2015. "Moral Worth and Moral Knowledge." *Philosophy and Phenomenological Research* 93 (2): 393–418.

Smith, Michael & Philip Pettit. 1990. "Backgrounding Desire." *Philosophical Review* 99(4): 565–92.

Sreenivasan, Gopal. 2017. "Acts, Agents, and the Definition of Virtue." *Oxford Studies in Normative Ethics Vol. 7*. Oxford: Oxford University Press.

Thomson, Judith Jarvis. 1997. "The Right and The Good." *Journal of Philosophy* 94 (6): 273–98.

Twain, Mark. 1884. *Huckleberry Finn*. Edited by Walter Blair and Victor Fisher (1988). University of California Press.

Wallace, R. Jay. 2003 "Explanation, Deliberation, and Reasons." *Philosophy and Phenomenological Research* 76 (2): 429–35.

Williams, Bernard. 1995. "Replies." In *World, Mind, and Ethics: Essays on the Ethical Philosophy of Bernard Williams*, edited by J.E.J. Altham and Ross Harrison. Cambridge: Cambridge University Press.

Willamson, Timothy. 2000. *Knowledge and Its Limits*. Oxford: Oxford University Press.