# Externalism and Self-governance

JOHN SKORUPSKI
*University of St Andrews*

What outcomes are good, and what there is reason for one to do, is not generally determined by what one thinks or even what one has reason to think. But is a similarly 'externalist' account of the distinctively moral concepts, the concepts of moral duty or obligation, of moral wrongness, blameworthiness and guilt, appropriate? I argue not; and on that basis I suggest that an externalist account is not appropriate for the concept of a virtue either.

Julia Driver and I would both, I believe, accept a theory of the good, and a theory of reasons for action, which is 'externalist'. That is, we both think that what outcomes are good, and what there is reason for one to do, is not determined by what one thinks or even what one has reason to think, except in occasional or indirect ways. Also, we share a broadly teleological conception of ethics. Where we disagree, however, is on the question of whether externalism is appropriate as an account of the distinctively moral concepts, the concepts of moral duty or obligation, of moral wrongness, blameworthiness and guilt.

Driver's book is not in fact mainly about any of these concepts; it is about the notion of virtue. She argues for an externalist account of virtue. I am not completely sure what I think about that, because I am not sure what I think about the concept of virtue. I do not really know whether it plays any stable role in our present-day moral discourse. In particular, do we have a clear distinction between moral virtue and other sorts of admirable character traits, and in so far as we do, is that not driven by a prior grasp of the notion of moral wrongness?[1] Personally I find it easier to know what duty is, and how I know my duty, than to know what the virtues are, and to know what is good about them.

For this reason my approach here will be indirect. I shall start by laying out some obvious implausibilities about externalism, distinguishing three kinds of concept you could be an externalist about: virtue, duty and blameworthiness. The case in which externalism seems to me to be most obviously wrong is that of blameworthiness.

In the second section I shall argue that duty, moral wrongness and blameworthiness are conceptually linked, and that what lies at the heart of this linkage is our conception of the moral agent as

[1] Consider: (i) generosity, fair-mindedness, prudence; (ii) courtesy, reliability, calmness, open-mindedness, good humour; (iii) authenticity, sincerity, spirituality, nobility, aesthetic sensibility; (iv) self-respect, chutzpah, boldness, hipness, attitude and cool. They all have something to be said for them. Are they all virtues? If not, why not?

© 2004 Cambridge University Press
DOI: 10.1017/S095382080300102X
*Utilitas* Vol. 16, No. 1, March 2004
Printed in the United Kingdom

https://doi.org/10.1017/S095382080300102X Published online by Cambridge University Press

*self-governing*. Central to this conception is a requirement that has been particularly important in modern moral (and political) thought, and that is linked to the idea of conscience. It is that moral agents must be able to tell, for themselves, in concrete settings of choice, what is morally right and wrong. They do not need to be told. Now this central requirement of morality as self-governance rules out externalism as an account of the moral concepts – because externalism makes it impossible for me to know my duty in the way that the self-governing conception of the moral life requires that I should be able to know it.

Finally, in the third section, I shall say something about why this seems to me to make externalism about virtue implausible too.

## I. EXTERNALISM

Here are some passages in which Driver states her externalism about virtue:

> [A] virtue is a character trait that systematically produces a preponderance of good. (p. xvii)
> [A] virtue is a character trait that leads to good consequences systematically. (p. xviii)
> [A] moral virtue is a character trait that produces good consequences for others. (p. 46)
> [W]hat is definitive is this: a virtue is a character trait (a disposition or cluster of dispositions) that, generally speaking, produces good consequences for others. (p. 60)
> As long as one's psychological states systematically lead to the good, that is sufficient for virtue. (p. 61)
> [A] virtue is a character trait that produces more good (in the actual world) than not *systematically*. (p. 82)
> [A] virtue is a disposition that systematically produces good. (p. 89)
> *x* is a virtue iff it is a character trait that produces actual good consequences overall or systematically. (p. 95)

Why does she take this view, given that she is well aware of its counterintuitive aspects? Well, what she says in its favour is that: 'The advantage of this account is that it preserves the connection between the agent and the world' (p. 70). As I understand her she simply means that on this account, virtuous actions have systematically good effects – whereas on the internalist account, they may or may not have. That does not seem much of an advantage to me, because it seems *true* that a virtuous action may have a bad effect. At least it is true if 'virtuous action' means 'action which proceeds from a virtuous disposition of the

https://doi.org/10.1017/S095382080300102X Published online by Cambridge University Press

agent'. It happens all the time. Kind acts misfire, honest dealing leads to disaster, and so forth.

But can it happen systematically? Here we have to reach for a more philosophically contrived example. So consider Satan and Planet X. This is a planet we know nothing of, on which there are billions of people under Satan's power. Long ago Satan set up an extremely stable mechanism which systematically links Earth and Planet X, in such a way that acts of honesty or kindness on Earth produce, *via* this systematic mechanism – and unknown to us – immense suffering on Planet X. On the other hand, even the mildest acts of dishonesty or cruelty on this earth have hugely beneficial pay-offs for the inhabitants of Planet X.

This is a bit like an example Driver herself discusses, about Joe and an evil demon (pp. 78 ff.). But I want to emphasize that we are considering this as an epistemic possibility about the world we *actually* live in, not as a counterfactual.

Suppose then that Satan comes along tomorrow and tells us the truth about Planet X. That will be a very dismaying piece of news, throwing us into confusion about what to do. From that point, certainly, we are going to have to think very carefully about whether *it is morally right* to do the honest or kind thing. But I do not think we would and for my part I do not think we should come to *any* of the following three conclusions:

1. Kindness and honesty have never been virtues; dishonesty and cruelty have never been vices. (externalism about virtue and vice)
2. It has never been morally obligatory to act honestly or to refrain from cruelty; equally, acts of dishonesty and cruelty have never been morally wrong. (externalism about moral right and wrong)
3. Acts of dishonesty and cruelty have never been proper objects of blame, honesty and kindness have never been admirable traits of character. (externalism about moral assessment of the agent)

As I said at the beginning, it seems easiest to see that externalism about the moral assessment of the agent is false. But from that a pretty clear path leads straight to internalism, at least about (2). This is my topic in the next section.

## II. MORALITY AS SELF-GOVERNANCE

We think of moral agents as *self*-governing and *self*-disciplining. Moral duty is an *internal* discipline on action, in that one determines its content for oneself and accepts guilt (self-blame) for non-compliance.

https://doi.org/10.1017/S095382080300102X Published online by Cambridge University Press

On this conception of morality, moral agents must be able to tell for themselves what is right and wrong: they must be able to ask and answer the question, 'What is my duty here and now?' They must be morally self-auditing.

This is the conception of morality as self-governance, a phrase I take from J. B. Schneewind's study of the evolution of modern moral thought, *The Invention of Autonomy*.[2] The story, as he tells it, involves a shift in the modern period from what he calls the morality of obedience to the morality of self-governance:

> The new outlook that emerged by the end of the eighteenth century centred on the belief that all normal individuals are equally able to live together in a morality of self-governance. All of us, on this view, have an equal ability to see for ourselves what morality calls for and are in principle equally able to move ourselves to act accordingly . . . These two points have come to be widely accepted – so widely that most moral philosophy now starts by assuming them.[3]

One of the most acute commentators on this conception of morality as self-governance is Hegel. He calls self-governance the 'right of the subjective will', and he analyses the 'moral point of view' as a one-sided or obsessive concentration on this conception, which he thinks can be dangerously empty or even corrupting. However he also thinks that the right of the subjective will is preserved in the form of ethical living-together ('ethical life', '*Sittlichkeit*'), which he favours. Self-governance is the 'subjective or "moral" freedom' to which Hegel (like Schneewind) thinks modern European ideals aspire:

> This subjective or 'moral' freedom is what a European especially calls freedom. In virtue of the right thereto a man must possess a personal knowledge of the distinction between good and evil in general: ethical and moral principles shall not merely lay their claim on him as external laws and precepts of authority to be obeyed, but have their assent, recognition, or even justification in his heart, sentiment, conscience, intelligence, etc.[4]

Notice that there are two elements in the passage from Schneewind, cutting across, I mean, the separation into two points he makes himself: (i) there is the self-governance conception of what it is to be a moral agent, and (ii) there is the moral egalitarianism, which says that we are

[2] J. B. Schneewind, *The Invention of Autonomy* (Cambridge, 1998).

[3] Ibid., p. 5.

[4] *Encyclopaedia*, §503, trans. William Wallace, in *Hegel's Philosophy of Mind*, being part three of the *Encyclopaedia of the Social Sciences* (1830), trans. William Wallace, together with the *Zusätze* in Boumann's text (1845), trans. A. V. Miller (London, 1971). 'Assent, recognition, or even justification . . . etc.' neatly signals his reservations about how far one can expect completely individual moral insight in everyone: about the origin of moral knowledge in shared ethical life and the different epistemic relations different individuals will have to it.

https://doi.org/10.1017/S095382080300102X Published online by Cambridge University Press

all equally, indeed on some Kantian views absolutely, moral agents.[5] Both are interesting, and they arguably produce a characteristic tension between the self-determinist and the egalitarian strands of modern moral thought. As Driver interestingly notes at various points in her book, classical notions of the virtuous person are decidedly elitist in their implications.[6] However here our focus is on (i), the conception of self-governance.

According to this, *moral* agents recognize for themselves what is right and wrong – they do not just fall in with what has been commanded by God, or come to that, with what has been commanded by an expert panel of scientific consequentialists. They must, in some sense, follow their consciences, and no doubt it is because modern moral thought so strongly highlights the self-governance conception that it becomes so preoccupied with conscience – with questions about its educability and fallibility, and about the sense in which it commands. (I am not saying, by the way, that self-regulation driven by blame and guilt only exists in the modern period – personally I think it is quite arguable that it is a fundamental human pattern that exists in any society. What is characteristic of modernity is rather the way it distinguishes this from other internalized norms of behaviour and foregrounds it in an ethical vision of equal morally self-governing agents.)

If morality is self-governance, then moral self-audit is basic to morality. And now the fundamental point is that the epistemology of moral self-audit is incompatible with externalism. I can and should accept guilt only if *I could have seen that* what I did was wrong. That can be developed by way of a connection between morality, blame and reasons for action as follows:

(A) X is blameworthy (absent extenuation) for doing Y only if X's epistemic state gave X sufficient reason to believe that there was sufficient reason for X not to do Y.
(B) It is morally wrong for X to do Y if and only if X is blameworthy (absent extenuation) for doing Y.

---

It is morally wrong for X to do Y only if X's epistemic state gave X sufficient reason to believe that there was sufficient reason for X not to do Y.

[5] Cf. Kant: 'What *duty* is, is plain of itself to everyone, but what brings true lasting advantage, if this is to extend to the whole of one's existence, is always veiled in impenetrable obscurity . . . But the moral law demands compliance from everyone, and indeed the most exact compliance. Appraising what is to be done in accordance with it must, therefore, not be so difficult that the most common and unpractised understanding should not know how to go about it, even without worldly prudence' (*Critique of Practical Reason*, trans. and ed. Mary Gregor (Cambridge, 1997), p. 33).

[6] e.g. pp. 36, 53. Compare n. 4 above.

https://doi.org/10.1017/S095382080300102X Published online by Cambridge University Press

This is the necessary condition that is violated by an externalist account of moral wrongness. But self-governance also implies something stronger – viz. that you can tell when an action would be blameworthy, and thus, given (B), when an action would be morally wrong. This must mean that your action's blameworthiness, and hence wrongness, is derivable from your epistemic state. Which leads us to the following principle:

*If you have a moral obligation you can tell that you have.*

Following Hegel, this is a precondition of 'subjective' or 'moral' freedom. One might call it the principle of personal moral insight. But what restrictions should be placed on it? Here are three possible restrictions:

1. unless you lack the relevant information
2. unless you lack the moral insight to tell that you have
3. unless you lack the time, in context, to work it out.

Now internalism about moral obligation, as it follows from the self-governance conception, discards (1). It cannot be that you had a moral obligation which you could not know that you had because you lacked the relevant information. That would be incompatible with the possibility of moral self-audit. For in that case, even if you could stop the clock and think as well and as hard as you could, you might not be able to work out what your moral obligations in this context were because they might be shrouded from you behind a veil of ignorance. So moral self-audit requires that your moral obligations in a given context are relative to your epistemic state in that context. (You may of course have had an anterior duty, which you failed to fulfil, to improve your epistemic state, but that is another matter.)

Now the points just made do not rule out restrictions 2 and 3. I do think there is a pressure in modern moral thought towards rejecting them as well – rejecting all these possible restrictions and adopting an unconditional version of the principle of personal moral insight for all human beings. Does not what *you* had a moral obligation to do depend on what *you* could have worked out for *yourself* in the time available? On the other hand, lack of moral insight and lack of time may be regarded as circumstances that extenuate you from blame even where you did something morally wrong (see (B) above). However, we do not need to decide this for present purposes. For present purposes the important point is that the principle of moral insight is not restricted by condition (1).

Let me also note an analogue in epistemology: a certain notion of *warranted belief*, which becomes indispensable if we think of epistemic

https://doi.org/10.1017/S095382080300102X Published online by Cambridge University Press

agents as self-governing. Again it goes with the notion of self-audit: self-governing epistemic agents must be able to ask: 'What am I warranted in believing here?' In particular, they must be able to decide whether they need to acquire further information to get a warrant for believing or disbelieving. In this case too, then

*If you have a warrant to believe that p you can tell that you have.*

And once again of the three possible restrictions –

1. unless you lack the relevant information
2. unless you lack the insight to tell that you have
3. unless you lack the time, in context, to work it out –

this notion of warrant discards restriction (1).

Duty, I conclude, supervenes not on the reasons there *are* for you to act, nor on the reasons you *believe* there are for you to act. What it supervenes on is rather tricky to get right, but it is nonetheless normative and internalist: something like, what reasons you have warrant to believe there are, or even: what reasons you would have been justified in context in taking yourself to have a warrant to believe there are. Something like this follows from the conception of morality as self-governance – or of moral agents as morally self-auditing agents.

It can still be asked what grounds can be given for thinking that there is or can be a normative system that corresponds to the self-governance conception of morality. The question goes beyond what can be discussed in this article, but we can notice three possible lines of reply:

1. It is a metaphysical truth about us as agents that we are bound by a system of self-governance. (Kant)
2. It is both natural to us and useful that we do acknowledge a system of self-governance.
3. The ideal of ourselves as bound by a system of self-governance (as 'morally free') is one that we wish to live up to.

It seems to me that even if we reject a Kantian metaphysics of agency, (2) and (3) give us plenty of scope for adhering to the self-governance conception, and thus to internalism about duty. So I see no reason why an ethics which is teleological in a broad sense should have to reject the self-governance conception of morality.[7] It is not a conception

[7] See my *Ethical Explorations*, ch. 3, for a discussion of teleological ethics in this broad sense.

https://doi.org/10.1017/S095382080300102X Published online by Cambridge University Press

which is inherently 'deontological', though no doubt it is one of the main contributors to the plausibility of deontological ethics.

## III. HOW DOES VIRTUE RELATE TO DUTY?

But how does virtue relate to duty? Let us distinguish, to begin with, three domains of the practical and some of their concepts:

1. The good: 'teleology', what ends or objectives should be pursued.
2. Excellence and ideals – what should be admired, despised; pride, shame, etc.
3. Morality: obligation, duty, conscience, moral wrongness, blameworthiness, respect.

Consider the first domain, the domain of the good. The good in general is that which there is reason to pursue. And as I have already noted, reasons for action relate to the facts, not to what we are warranted in thinking to be the facts. Here we are talking about what there is reason for me to do, not about what there is reason for me to believe that there is reason for me to do. Similarly with the domain of ideals. What is admirable depends on the facts, not on what we are warranted in believing to be the facts. In general, what there is reason to admire, despise, etc. depends on the facts, not on what we are warranted in believing to be the facts.

But as we have just seen, the same is not true of morality: one's duty is determined by the reasons one is warranted in believing there are, not by the reasons there are. The word 'right', so often used, can be very confusing here. 'Right action' can mean morally obligatory action. Or it can mean, action which there is most reason to do. But those are different senses, and it is a very implausible ethical position that holds that the morally obligatory action is always identical with the action there is most reason to do.

So what about the notion of virtue? Characterizing the notion of a moral virtue seems to me to be an exceptionally difficult task. For a start it seems evident that it is far from clear-cut: there are various excellences of character which at best lie on the border line between non-moral excellence and moral virtue.

'Virtue', in so far as we have a clear modern way of talking about it at all, sits across (2) and (3). On the one hand, a virtue, it seems to me, must be something naturally or spontaneously admired. But on the other, if we are to treat it as a moral notion, it must be narrower than that: it must somehow pull its weight in explaining how the person who has it does what is *morally* required. So here is a partial attempt

https://doi.org/10.1017/S095382080300102X Published online by Cambridge University Press

to capture that latter point:

> A character trait is a virtue only if it is a disposition to *respond to* some class of warrantable reasons[8] when they become, in certain contexts, *morally salient* (or: to act from these morally salient reasons non-accidentally).
>
> A warrantable reason is *morally salient* in a context when failure by a moral agent in that context to act in accordance with it is (absent extenuation) blameworthy.

I present this only as a necessary condition. (Concern for the respect of morally good people may fit the condition – but is it a virtue? The reply may be that a person acting only from this concern is not acting from morally salient reasons alone, or as such – but the question then is whether we sufficiently understand 'alone' or 'as such'.) Consider in this light the case of Huckleberry Finn, discussed by Driver. Does he act virtuously in helping Jim, even though when he does so he believes he is doing the morally wrong thing? Driver answers in the affirmative and I agree with that. Later she comes to the following general conclusions, which I also agree with:

> First, it is not clear at all that any unified account of virtue can be given when that account requires some particular type of mental state to be common to all virtues.
>
> Secondly, and possibly more radically, when a person possesses a virtue, this does not mean that in acting in accordance with that virtue, he is doing something he thinks to be morally good. (p. 58)

The case of Huck seems to me to show that the second conclusion is right. He is responding to reasons which are morally salient, even though he does not recognize that that is what they are. Of course this is delicate – his kindness in helping his friend must somehow connect to the moral salience of reasons of kindness in that particular situation. Had he been equally disposed by his kindness to friends to help a friend who's a serial rapist or a murderer to escape justice, I doubt that we would call his kindness to friends a virtue. Still, as I say, I agree with these points Julia Driver makes about virtue, and in particular with the point that you can act virtuously without recognizing that you are acting in accordance with duty.

But I do not agree that this sustains her externalist view of virtue. On the contrary, it is consistent with the characterization of virtue that I have just given, and that is an internalist characterization. For my characterization connects virtue to a disposition to respond to reasons

[8] Reasons which one is warranted in thinking one has.

https://doi.org/10.1017/S095382080300102X Published online by Cambridge University Press

which are, in certain contexts, morally salient. And given my defence of internalism about blame, these 'morally salient' reasons must be warrantable reasons. That does not mean that the virtuous agent must recognize those reasons, in those contexts, *as* morally salient. He may even deny that they are, as Huck does. But still, we assess the virtuousness of his disposition to act from reasons which he is warranted in thinking exist, or perhaps just reasons he thinks he has. Either way, these can diverge from the reasons for action that there really are – as the case of Satan and Planet X shows.

jms2@st-andrews.ac.uk

https://doi.org/10.1017/S095382080300102X Published online by Cambridge University Press