

STANDARD BEARERS

DAVID SOSA

david_sosa@austin.utexas.edu

ABSTRACT

In both ethics and epistemology an important question is whether justification is a fully internal or a partly external matter. In view of analogies between relevant considerations in each area, I recommend distinguishing, as basic and independent *subjects* of normative status, (i) agents and (ii) what they do. Evaluations of subjects, on one hand, and of their beliefs and actions, on the other, are less intimately related than is presupposed. This helps resolve internalism/externalism controversies in both domains. An important related advantage of the distinction is its effect on our understanding of normative luck, both moral and epistemic.

Sadly, sadly, the sun rose; it rose upon no sadder sight than the man of good abilities and good emotions, incapable of their directed exercise, incapable of his own help and his own happiness, sensible of the blight on him, and resigning himself to let it eat him away. Charles Dickens, *A Tale of Two Cities*

The worst cynicism: a belief in luck. Joyce Carol Oates, *Do What You Will*

I. MORAL LUCK

To understand normativity, begin with moral luck. There is something philosophically discomfiting about the possibility of moral luck. Nagel suggested that the problem is one “to which we possess no satisfactory solution” (Nagel 1979: 25) and Bernard Williams thought it might require a significant revision of our concept of morality (Williams 1982: 39). There has been little progress since.

But what is moral luck? Nagel is relatively explicit. “Where a significant aspect of what someone does depends on factors beyond his control, yet we continue to treat him in that respect as an object of moral judgment, it can be called moral luck” (Nagel 1979: 26).

Williams reminds us of the (presumably, Stoic) Sage, who is immune to the impact of “incident” luck. The idea that one’s whole life can be rendered immune in this way has not prevailed since antiquity, but a vestige remains in the idea that moral value *is* so immune: “The successful moral life . . . is presented as a career open not merely to the talents, but to a talent which all rational beings necessarily possess in the same degree” (Williams 1982: 21). Nagel and Williams both see Kant as the leader of the opposition (to moral luck). Here’s a compelling quotation from Kant, “even if it should happen that, by a particularly unfortunate fate or by the stinting provision of a remorseless Nature, th[e] will should be wholly lacking in power to accomplish its purpose . . . it would sparkle like a jewel in its own right, as something that had its full worth in itself.” Kant is thus a rigorous exponent of the view that moral value is “unconditioned.” Such a view “has an ultimate form of justice at its heart,” quips Williams with characteristic irony; it is “only superficially

repulsive – despite appearances, it offers an inducement, solace to a sense of the world’s unfairness” (Williams 1982: 21).

The problem with this Kantian conception of moral value develops, as Nagel appreciates, “out of the ordinary conditions of moral judgment.” Prior to reflection, Nagel says, “it is intuitively plausible that people cannot be morally assessed ... for what is due to factors beyond their control ... when we blame someone for his actions we are not merely saying it is bad that they [the actions] happened, or bad that he exists: we are judging *him*, saying he is bad, which is different from his being a bad thing” (Nagel 1979: 25). There is an important distinction gestured at in passing here, between an agent’s being bad and its being bad that their actions occurred; it’s a distinction to which we will return, emphatically, below.

In any case, however, and upon further reflection as it were, Nagel points out that “what we do depends in many [ways] on what is not under our control ... And external influences in this broader range are not usually thought to excuse what is done from moral judgment, positive or negative” (Nagel 1979: 25). For example, there is a morally significant difference between rescuing someone from a burning building and dropping them from a twelfth-story window while trying to rescue him. “Similarly, there is a morally significant difference between reckless driving and manslaughter” (Nagel 1979: 25). His example of a father’s running up the stairs, realizing that he negligently left the bath running with the baby in it, is terrifying. “[I]f the baby has drowned one has done something awful, whereas if it has not one has merely been careless” (Nagel 1979: 31).

J.J. Thomson shows how Williams and Nagel entertain a certain “Weak Result: Whatever we do, our doing of it is no more to our discredit than are those purely mental acts by which we do it” (Thomson 1989: 199). Effects do not appear to be essential to their causes; whatever a person does, it is either a purely mental act or something done thereby; and finally, the causal consequences of an act do not redound to the discredit of the agent – only the act itself does. This is a line of thought whose ultimate result is to “pare down each act to its morally essential core, an inner act of pure will assessed by motive and intention” (Nagel 1979: 32).

On the other hand, Adam Smith had already noted that “when we come to particular cases, the actual consequences which happen to proceed from any action have a very great effect upon our sentiments concerning its merit or demerit” (Smith 1759). Moreover, the theoretical tendency to reduce the scope of the moral is at odds with the intuitive role not only of what proceeds from an act, but also of what precedes it too: what Nagel calls “constitutive” luck (which we will not treat here) – luck in one’s traits and dispositions – is no less anathema to the Kantian perspective than the more familiar “incident” luck.

Williams famously uses the case of Gauguin to set the scene. According to Williams, whether a given choice is justified, sometimes, can depend on whether one is in the relevant sense *successful*. And this success can depend crucially on factors outside one’s control. For Williams, to the extent it is circumstances external to “what one has chosen” that prevent success, to that extent the justification of the choice is untouched. Failures due to internal factors, by contrast, may leave the agent with “nothing to say” to justify their choice and so do affect its status. But some internal failures can according to Williams nevertheless be due to factors not in the agent’s control.

According to Williams, Utilitarian and broader consequentialist formulations can “offer the thought, ‘it is better (worse) that he did it,’ where the force of that is, approximately, ‘it is better (worse) that it happened,’ but this in itself does not help towards a

characterization of the agent's decision or its possible justification" (Williams 1982: 24). Here we have an echo of Nagel's distinction between its being bad that something happened and a more direct characterization of the agent's decision.

According to Williams, the utilitarian perspective will, to make matters worse, miss an important dimension of our cases: again, that the only sort of failure that matters is internal to the project undertaken. "From the perspective of consequences, the goods or benefits for the sake of which Gauguin's choice was made either materialize in some degree, or do not materialize. But it matters considerably ... in what way the project fails to come off, if it does" (Williams 1982: 25). Summing up, Williams notes, "[w]hat would prove him wrong in his project would not just be that *it* failed, but that *he* failed" (Williams 1982: 25, emphasis added).

So Williams contemplates something Nagel seemed implicitly to reject. For Williams, it seems, whether someone is wrong depends on whether *they* succeed; and while this latter condition excludes what is external to the project, it is nevertheless not entirely in their own control. Nagel by contrast seemed to distinguish evaluation of the subject (whether they are good or bad) from evaluation of the occurrence of their actions; and he seemed at least implicitly set against the idea that external factors – factors outside the subject's control, whether external or internal – could properly affect evaluation of the subject.

Having noted (1979: 35) a connection between the problem of moral luck and the problem of free will, Nagel then has this to say:

The problem arises, I believe, because the self [that] acts and is the object of moral judgment is threatened with dissolution by the absorption of its acts and impulses into the class of events. Moral judgment of a person is judgment not of what happens to him, but of him. It does not say merely that a certain event or state of affairs is fortunate or unfortunate or even terrible. It is not an evaluation of a state of the world, or of an individual as part of the world. We are not thinking just that it would be better if he were different, or did not exist, or had not done some of the things he has done. We are judging *him*, rather than his existence or characteristics. (Nagel 1979: 36)

We judge them, Nagel says, "rather than their desirability or utility" (Nagel 1979: 37) and "[o]nce we see an aspect of what we or someone else does as something that happens, we lose our grip on the idea that it has been done and that we can judge the doer and not just the happening" (Nagel 1979: 38). Nagel emphasizes throughout that what we face is a paradox: Kant's view both "appears to be wrong" but also "arises in response to a fundamental problem about moral responsibility to which we possess no satisfactory solution" (Nagel 1979: 24–5).

I have tried to rehabilitate the problem of moral luck in some detail because I want to recommend an analogy with epistemology. But let me first resist a different analogy, one that Nagel himself recommends, between (i) the claim that a subject's moral responsibility is "undermined" by the ordinary prohibition on external influence, and (ii) the claim that a subject's having knowledge is undermined by the application of ordinary standards for knowledge. In both cases we have a normative status the ordinary conditions for which appear to be effectively unsatisfiable.

But while I too will emphasize continuities between the nature of practical and epistemic normativity, I think the differences between the philosophical problem of

skepticism and that of moral luck are more important than the similarities: *that* is not the best place to locate the analogy.

The temptation to skepticism that can emerge from Cartesian considerations, at least, takes a characteristic form: knowledge, specifically, looks like it makes demands, concerning our having grounds for belief, that we couldn't possibly satisfy. But that temptation is relatively restricted: we don't, in the face of whatever skeptical doubts we have about *knowledge*, automatically also find ourselves doubting whether there might be any such thing as making a good epistemic move, drawing a good inference for example.

By contrast, in the case of moral luck it is good agency itself that is under threat. If the very idea of agency's being apt for evaluation of a certain distinctive sort depends on that agency's being "unconditioned" and immune to external influence, then we face the threat of there being no such thing, nothing that can be assessed in that way. Hume's skeptical doubts may thus mark a better analogy. It's not the claim that we do not know anything or that, philosophically, we should appreciate the impossibility of knowledge, specifically, that, I think, Hume is trying to establish: it is rather the impossibility of our being justified in forming beliefs (of a certain sort, the *a posteriori* sort). And it is that sort of analogy, not the analogy to skepticism drawn by Nagel, that I will pursue here.

2. TWO FORMS OF NORMS

In epistemology and in ethics, *justification* is a central phenomenon. Both subdisciplines are in that way normative. They are concerned with a distinctive kind of status. And in both subdisciplines one important question concerns the degree to which the corresponding feature is in any sense *within* the subject: must internal duplicates be justified alike?

What is often called the "internalism/externalism debate in ethics" is not the focus here. What is often so-called is a debate about motivation: must an agent who *judges* that something is good be *motivated* to pursue it? – to give one of its forms. That's a debate about whether there's an internal connection between ethical judgment and motivation, and about the nature of whatever connection there might be (Smith 1994).

What is incorporated here as an internalism/externalism debate in ethics, by contrast, is the issue that I think most deeply separates committed consequentialists and devoted deontologists from each other: can factors in some sense "beyond" the agent determine relevant ethical matters? Ain't goodness in the head? – to put it derivatively, dramatically, and ungrammatically.

Maybe the analogy with epistemology is not hard to spot. Sometimes, but not always, we have a true belief that has a certain important further epistemic status: it constitutes a bit of knowledge on our part. Must duplicates' true beliefs be alike in respect of this status? Could a pair of such duplicates not know alike? Reliabilists say yes: one's environment might not cooperate. Internalists, on the other hand, of either foundationalist or coherentist varieties, resist the idea of any fundamental normative notion yielding such a hostage to what lies beyond.

So: we will approach issues about normative status in ethics and in epistemology. I'm concerned with whether duplicates must be epistemically or ethically co-justified, with whether either of a pair of duplicates can be ethically or epistemically *better* than the other.

Hume said, "[f]rom this circumstance alone, that a controversy has been long kept on foot, and remains still undecided, we may presume that there is some ambiguity in the

expression, and that the disputants affix different ideas to the terms employed in the controversy” (*Enquiry*; Hume 1748). A first sign that the concept in dispute here may not be univocal appears when we ask what the *bearers* of a normative *standard* might be. What sorts of thing can be justified? Some basic applications of that feature are to subjects in respect of their actions and beliefs. In ethics we are concerned with what it is for *S* to be justified *in doing A intentionally*; and in epistemology, we are concerned with what it is for *S* to be justified *in believing that p*. The agent is justified *in* intending or believing as they do.

Now: need these targets of evaluation be further specified? Surely not, one might think. We evaluate agents *insofar* as they perform individual actions or have particular beliefs. And the beliefs or actions we evaluate are of course held or performed by agents. Normative status is thus instantiated, apparently, by a *complex* constituted by agent *and* their action or belief. The distinction we saw incipient sensitivity to above, in Nagel and in Williams, between an agent’s being bad and its being bad that actions of theirs occurred, is in this way minimized: whether it is bad that they so acted is a function of what they are like and what they are like is, in systematic ways, expressed in what they do.

But these points, seen as favoring the “surely not” response, are not sufficient. The fact that we evaluate agents *insofar* as they perform individual actions does not entail that there can’t *also* be a sort of evaluation of the agent that’s fundamentally different from an evaluation in terms of the actions they perform or of the beliefs they maintain. And in fact we can evaluate the agent’s actions or beliefs directly, independently of our evaluation of the agent.

Ortcutt as a spy is sneaky. And Ortcutt as a pillar of the community is forthright. How’s Ortcutt himself? Maybe it’s hard to say. But maybe not. Maybe Ortcutt’s forthrightness is a put on, a political maneuver. That’s not what he’s really like. Or maybe it’s rather his sneakiness that’s affected – a prerequisite of his spooking. Of course, Ortcutt’s sneaky activity is definitely sneaky, affected by him or no: the *activity* does not affect sneakiness. And his forthright public behavior is just what it is, even if the forthrightness is a put on.

We should think of *S*’s *performing action A* and of *S*’s *believing that p* as each occulting some further structure. In evaluating either of those complexes, there’s the risk of running together an evaluation of *S* themselves with an evaluation of what *S* has done or of *S*’s belief.

Good people at least sometimes do bad things. And vice versa. And good epistemic agents sometimes fail to be justified or to know. And again vice versa. So much is uncontroversial, and relatively insignificant. But a natural further thought is that even otherwise good people who do bad things are, as we might put it, *bad in that* (or *insofar as*, or *inasmuch as*) *they did those bad things*. And otherwise bad epistemic subjects who come by some knowledge are at least, it seems, *epistemically good in that* (or *insofar as*, or ... etc.) *they have that belief*. Affected or not, Ortcutt is sneaky *insofar* as he performs his sneaky activities.

This approach embodies a reductive tendency: good people can do bad things, but their having done that bad thing is a mark against their goodness. And a sloppy thinker can have a fortunate moment of epistemic clarity: the resulting belief counts in favor of their epistemic situation. So doesn’t this approach eliminate any residual significance that might have attached to the distinction between possible objects of evaluation, possible

bearers of standards – the agent on one hand and the agent’s action or belief on the other? Isn’t the goodness of the good agent crucially interdependent with the goodness of her actions? Doesn’t, say, failure to know, given true belief, depend at least in part on an inadequacy of the believer as an epistemic subject?

I think the correct answer is *no*. The further thoughts above need not be denied. But their reductive effect can be isolated from the view being developed here: to say that the good agent was bad in that they did a bad thing is I think just to say that they were the agent of that bad action. That’s a way, but just one way, to be bad: to be the agent of a bad action. And for an epistemically bad subject to be good in having a belief that constitutes knowledge is simply for them to be the subject of that knowledge-constituting state of belief. There is still another, independent, question, I urge, of whether the agent is in such a case good in another important (epistemic) way.

So the view developed does not *deny* the reductive further thoughts that can seem to eliminate the distinction being drawn; those thoughts are correct as far as they go. But they are insufficient for the elimination and do not fulfill the reductive aspirations some may have for them. There are indeed concepts of good agent and good epistemic subject according to which the goodness of the agent depends on the goodness of their actions and according to which good epistemic subjectivity is modally correlated with knowledge. But there can be, compatibly with that, concepts of good agent and good epistemic subject that are *unconditioned* by the notion of good action or of good true belief. And though there are concepts of good action and of good belief that inherit good agency or good subjectivity as a condition, there can be, compatibly with that, concepts of good action and good belief that are independent of the notions of *good agent* and of *good subject*.

It might still seem a platitude that although a generally good epistemic subject might be unjustified in one or another belief, they would thereby be less good. And otherwise good belief that does not reflect the good epistemic activity of the believer seems to be the lesser for the lack. Our concepts of the good believer and good belief might seem inextricably linked. Although exceptions are possible, there should be significant modal correlations, in both directions, between being a good believer and having good beliefs: accidents aside, a subject’s belief *would* not be good unless that subject were good, and the subject *would* be good only if they would have good beliefs.

It is correlations like these that I mean to qualify here. I think even a *perfectly* good subject *could* not (ever) have a good belief. Perhaps this could be only “by a special disfavor of fortune or by the stinting provision of an ungenerous nature” (Kant 1785: 394). Forrest Gump – self-descriptively (if, typically, inadvertently) mangling the proverb “handsome is as handsome does” – intones, “Stupid is as stupid does,” effectively encapsulating the popular view I am challenging here. A large part of the appeal of the Gump phenomenon, as well as that of Mr. Magoo, the notoriously myopic cartoon character from the 60s and 70s, is, however, precisely the *systematic* breakdown between what they’re really like and the status of their beliefs: they’re goofy, but effective. But if the breakdown between their excellence and the status of their states can be systematic, then a correlation between the excellence of a believer and the excellence of their epistemic states should not be presupposed.

Suppose that the presence of Descartes’s demon were “modally robust”: even had the possible world in question, in which the demon operates, differed in any of ever so many ways, the demon would still have operated as they do. Such a world would of course appear to us exactly as our world actually does; but none of our actual perceptual beliefs

would be true with respect to that world. But worse: had *that* world been different, it would still have contained Descartes's deceiving demon, deceiving us in some way or other.

What's surprising is that, relative to that world, epistemic excellence comes with an epistemic cost. Partly *in virtue of* their epistemic excellence, good epistemic subjects would not in that world acquire perceptual knowledge; and that's an epistemic loss. Worse, the failure of good epistemic subjects in that world to acquire perceptual knowledge is not accidental – it is modally robust. Not only *do* they not acquire knowledge, they *would* not acquire knowledge – notwithstanding their epistemic excellence. The deceiving demon produces a misleading course of experience; the good epistemic subject forms perceptual beliefs accordingly. But the bad epistemic subject ignores the misleading perceptual evidence. Strangely, then, subjects who do *not* form perceptual beliefs in accord with their perceptual experiences will not be *deceived* (their beliefs may anyway be true or false, and they anyway won't gain knowledge). The demon will be frustrated (at least in their practice to deceive): only good epistemic subjects can be deceived in that way. Ironically, subjects that are epistemically *bad* are immune to deceit of this sort.

The evil demon is an epistemic analog of Kant's "special disfavor of fortune." A certain sort of *epistemic* excellence thus also "like a jewel . . . sparkle[s] by itself, as something that has its full worth in itself" (Kant 1785: 394). When we focus on the agent as an independent locus of ethical or epistemic status, it becomes hard to conceive what is not in some important sense *in* them as threatening that status.

But – or perhaps *and* – correspondingly, when we focus on the action or belief as a locus of ethical or epistemic status, it is hard to ignore its embedding in the rest of reality. An action that would, say, wipe out all happiness would be, at least in one sense, a very bad action, no matter how well-intentioned. And this is a normative status: such a disastrous action is *to be avoided*. (And whatever intention led to it, however good it might otherwise be, could reasonably be seen as tainted by its effect.) A belief that, were the agent to have it, *would* be true has an important positive evaluative status; it's good in a way, no matter what might be said of its subject. Whatever subjects might aim at, belief by its nature aims at the truth, in at least that sense: there's an important way in which belief states are better if their obtaining is modally correlated with their own truth.

The New Evil Demon problem is the problem of moral luck reframed for epistemology. And Norman and Mr. Truetemp (Gump and Magoo for the epistemic set) raise effectively the same issues (if with the opposite polarity from those raised by the Demon's victim). The approach developed here to these debates in ethics and epistemology, by contrast, proceeds from skepticism about the univocity of the normative concepts in play. There is the status that agents can sometimes have that we describe in saying that they are good or justified. But there is also, quite differently, the status that intentions and beliefs can sometimes have that we describe in saying that an intentional action was a good thing to have done, or that the belief is a good one to have – that it will constitute knowledge.

Must duplicates have the same normative status, ethical or epistemic? Yes and no. There's a status an agent has simply in virtue of being the sort of agent they are, where this very sorting at the same time sets the limits of internality and accordingly gives content to the relevant notion of *duplicate*. *That* status must of course be shared by duplicates. The only question is whether there really is such a status: it's very plausible that there is. The examples recommend it. It is not for lack of appreciation of the robustness of alternative varieties of status that we're committed to this one.

There is, indeed, another normative status agents can have, derivatively, a status they inherit from their actions or beliefs: if you intentionally do something that could hardly but have caused widespread unhappiness, or if you have a belief that could hardly but have been false, then because you are the agent of that action or the subject of that belief, you have come up short in an important way, ethically or epistemically. You are subject to an important kind of censure; and it's an evaluation no other duplicate of yours need share. If you're in a demon world, or even just in barn country, you'll be *worse* off epistemically (than you would otherwise be), even though you may be (indeed, perhaps *because* you are), internally, beyond reproach. This variety of evaluation redounds from your action or belief to you as its agent or subject.

So I approach these debates by distinguishing two subjects of normativity: agents and the states they're in. There are, accordingly, both agent-centered normative statuses and state-centered normative statuses. Let me quickly reiterate a point: intentional action, and belief, *can* be evaluated with respect to agent-centered normative status. And agents *can* be (and often are) evaluated with respect to state-centered normative status. That's partly because our world does not contain systematically ethically- or epistemically-deceptive demons: fortune does not in fact disfavor us so especially, Nature provides rather richly. In our world, whether you are good and whether what you do is good are, as a matter of fact, well correlated (correspondingly for the epistemic case). But your goodness and that of your intentions and beliefs are metaphysically independent matters, as the thought experiments reveal.

I think another possible reason for our tendency to evaluate agents with respect to state-centered normative status (and, to whatever extent, vice versa) is that the expression "Agent S was good *insofar as* they did A" *looks* like an evaluation of the agent; but its content, I've suggested, is effectively the conjunction of *S did A* and *A was a bad action*. The first conjunct is not explicitly evaluative at all and the second involves only a state-centered evaluation of an action. An adequate semantics for "as" phrases is beyond the scope of this essay; but notice that it does not follow from the fact that Orcutt, insofar as he is a spy, does sneaky things that Orcutt is sneaky: his sneakiness is a put-on, professionally affected, remember.

Proposal: objective consequentialism, on one hand, and Kantian deontology and subjective consequentialism, on the other, need not conflict, at bottom, over the effect of external conditions on, say (ambiguously), moral excellence. Kant expressed a deep insight about the status of *agents* as such: morally good agents, agents of a good will, have a kind of excellence that they express in isolation from any possible external contingency. And subjective consequentialism appropriately emphasizes the importance for being a good agent of *aiming at* what, by your lights, would have the best consequences.

Duplicates – in effect by definition, since the relevant variety of duplication is precisely with respect to features of the will – are good or bad alike. But objective consequentialism rightly stresses something different. Some people sometimes have the following morally significant feature: they performed an action that was (and would be) of great value, they made lives better. That sort of goodness (of the good thing done) may well be a consequence of its causal effects (or even of its own causes), features the action has partly in virtue of its embedding in the vagaries of our external contingent circumstance. Duplicate agents might not be evildoers alike. And evildoing is a normatively significant phenomenon; it's *to be* avoided. Still, it's susceptible to outside interference, and to luck. I hope these thoughts were set up by the discussion of moral luck above.

I come to praise externalism, not to bury it. So let it be with epistemology: internalist foundationalism and coherentism need not conflict with externalist conceptions of epistemic normativity. Epistemically good agents, agents with their mind in good epistemic order, have a kind of excellence that they hold with immunity from any possible external contingency. They form perceptual beliefs in accord with their perceptual experience (however exactly this accordance is ultimately understood); they reason well, drawing inferences appropriately; and they develop an elegant and coherent system of belief. Duplicates will be good or bad, in this respect, alike.

That sort of excellence is to be distinguished from the feature on which an externalist perspective focuses (especially if it is broadly reliabilist): some belief states represent and *track* the truth. The belief's obtaining (perhaps: as a result of the process involved) is reliably correlated with the truth of its content. The details of the correlation in question may be controversial (what sort of *tracking* is crucial?): but reliabilists of all kinds are interested in the special status beliefs have when their relation to the truth is modally robust. My epistemic status in fake barn country is problematic precisely in that my belief and its truth are, there, insufficiently intimate.

The relevant intimacy, when it obtains, is a modal articulation of the belief's actual truth; and while everyone is an externalist about truth, the reliabilist, extending this externalism into modal space, finds epistemic excellence in a more stable connection to the truth. It's epistemically good for a belief to be true, and that's an external matter; but it's a different good thing when a belief's reliably true, and this too is an external matter. Then the belief can constitute knowledge. This is a normative status for the belief; the belief is *better* as a belief, as it more reliably achieves its aim.

3. DUAL NORMATIVE-ASPECT THEORY¹

The general picture proposed here should by now begin to come into view. But I should engage a significant challenge to my approach, an issue that derives from the concept of *skill* that flourishes in virtue theories epistemic and ethical. Such views (for one example, see Murdoch 1970) emphasize the role in our ethical thought of "thick" evaluative qualities (e.g. courage) and would oppose the adequacy of a line on which an agent's status can simply shear away from the status of their actions.

Some time back I went to the ballet and saw a skilled performance of Stravinsky's *Rite of Spring*. A skilled performance, I say, meaning to identify a very specific normatively-significant characteristic. Skilled performances are by skilled performers; they are performances in which the skill of the performer is manifest, and they are beautiful. Someone without skill could have moved their body in precisely the same beautiful arcs without the performance having been skilled. And, of course, skilled performers, owing to fatigue or injury for example, sometimes do not manifest their skill. So a skilled performance requires a kind of *co-operation*, of the skilled and the skill.

Aristotle thinks of good living a bit like that, I think. The virtuous agent is, in a way, a *virtuoso*, adept at living – at living a human life, in particular. The best lives are like skilled performances: lives of virtuous agents in which their virtue is manifest and in which other

1 The material in this section overlaps with that in my "The Vice of Virtue Theory" (Sosa 2016).

values are thereby instantiated. Aristotle is right that there is excellence in this sort of virtuous agency. But I think the concept may be, nevertheless, in an important sense incoherent (not: inconsistent). It brings together, as if under the single normative dimension of ethical excellence, what are really two fundamentally different sorts of normative concept. How, in the demon world, to be a virtuoso?

Case (i): An archer takes aim, draws her bow, and releases. Moments later, thanks to the archer's skill, the arrow strikes the bullseye. Compare that with four other cases: case (ii) the skilled archer takes accurate aim, draws, and releases; but the wind whips up unpredictably and the arrow misses. Case (iii): An amateur archer aims badly, draws, and releases; but the wind whips up unpredictably and, no thanks to any skill, the arrow strikes the bullseye. In case (iv) a skilled archer aims well, draws, and releases; the wind whips up unpredictably, but, still more unpredictably, a second, compensating gust arises afterward and the arrow strikes the bullseye anyway, no thanks to the archer's manifest skill. Finally, case (v): owing to fatigue, the skilled archer cannot make manifest her skill. Nevertheless, she shoots, intending to strike the target; and in fact she does. Each of (ii)–(v) have something in common with (i) and something different from it. In (ii) we have a skilled performer, but the arrow does not strike the bullseye. In (iii) we have a striking of the bullseye; but the performer is not relevantly skilled. In (iv), we have both a skilled performer and a striking of the bullseye; but the striking of the bullseye by the archer is not an *action* of the archer's. Finally, in (v), a kind of ethical Gettier case, we have a skilled performer and a striking of the bullseye; and we have it that the striking is an action. But still something's missing – we don't have the right sort of *performance*. The striking is adequately connected to agent's intentions for the occurrences to constitute an action. But because that action does not manifest the agent's skill, we do not have a skilled performance (or something like that; I am not jealous of the terminology).

According to the challenge I'm trying to engage, the first case is the best case; it combines desiderata: we have a good agent, a good action, *and* the goodness of the action manifests or expresses the goodness of the agent. An account that emphasizes that *combination* will conflict with the approach developed here. But is the virtue-theoretic combination in any important sense normatively more than the sum of its parts? I doubt that those parts are sufficiently alike for their conceptual combination not to be, however compatible, crucially incoherent.

Beauty is a value. And for belief, truth is an epistemic value. So consider true beliefs held under beautiful sunsets. That's a feature that some beliefs have: they're true *and* held under a beautiful sunset. The feature in question, moreover, is clearly evaluative: it is constructed from concepts whose instantiation requires the instantiation of values. Any *holding under a beautiful sunset* will entail the existence of some beauty, for example. So, true beliefs held under beautiful sunsets may well excel, exhibiting as they do a specific normative status. But that status is so clearly simply the conjunction of different sorts of excellence that it seems philosophically *hyperopic* not to focus on the individual atoms of the molecular concept – a case of missing the trees for the forest.

Similarly – save for a caveat I'll take up below – with the concepts of virtuous ethical agency and of virtuous epistemic subjectivity. Virtuous agency expresses a kind of normative excellence. To have the sort of mind virtuous agents have is a way for *agents* to be good. And the sorts of intentional action or belief the virtuous agent intends or affirms will, as a matter of contingent fact, instantiate various values. In the best cases, those two kinds of goodness will co-operate: the good subject will produce a good intentional

action or belief. If the subject manifests their virtue, that will all be the case. But the epistemic goodness of good subjects is a far, far different sort from the epistemic goodness of their good beliefs. And production of the latter by the former cannot create a new circumstance on the normative scene.

Being a good agent, and thus sharing the internal constitution of a virtuous agent, is independent of the world beyond the agent and the values that might be instantiated therein. And being a good intention or belief is independent of the goodness of the agent. Of course, *being a good action performed in virtue of a good agent's goodness* is independent of neither. But the goodness of the good agent is internal to that agent in the following sense: nothing can be a duplicate of that agent without sharing her excellence. And the goodness of the good belief is internal neither to the agent nor even to the belief itself: the very same belief, even were it true, held by a duplicate subject, might “wholly lack the capacity to carry out its purpose” (another Kantianism); its *truth* might be in that sense an accident. Remember the barns – and, worse, the demon. Sometimes, Dickens realized, the man of “good abilities” may be unable to avoid “the blight on him.”

The goodness of action or belief, what objective consequentialists and reliabilists are especially interested in making out, is a matter of modally robust relation to certain sorts of value. *Good action* is a teleological and consequently *relational* notion, amounting to counterfactual relations between the agent and the value of the ends intentionally brought about. *Good belief* is a relational notion too: a matter of counterfactual relations between the state's obtaining and its truth. The goodness of the person, by contrast, is a *categorical* notion, specified only with respect to the agent. The goodness of the good agent will necessarily be maintained even across counterfactual variation of their external circumstance. This is where deontological, subjectivist, and internalist intuitions bite down hardest.

I enjoyed that ballet. The performance was beautiful; had the dancers been off their game, the occasion would have been the worse for it. And the dancers, it was clear, were very good dancers. A similarly beautiful event – even an arabesque-for-arabesque reproduction – by an incredibly lucky bunch of mediocrities would have been aesthetically the worse for the absence of the excellence of the dancers that in fact performed. I'm suggesting, however, that to lump together into the notion of a *skilled performance* both of those varieties of goodness is to produce a causally-mediated conjunction with pretensions to coherence that are in fact lacking.

A subtlety remains – the caveat mentioned above: why allow the goodness of an intention's effects to redound to its credit while not allowing the goodness of an agent's intention (and thus of its effects) to redound to the agent? First, unless we achieve a better understanding of the notion of agent-causation (I do hope one day we do), the agent does not appear to stand to their action as actions do to their consequences. And second, even if we did have an adequate account of agent-causation, note that with agents, the idea of being an internal duplicate can acquire a distinctive significance, one that can be revealed in consciousness. (This would properly contrast with the physicalist “molecule-for-molecule” notion of duplication current in debates in the philosophy of mind.) The notion of an internal duplicate according to which good agency is an internal matter isolates that goodness from the causally downstream *in a way for which there's no perfect analog with states of intention and belief*.

Even if duplicates must have intrinsically indistinguishable streams of consciousness, whether they must have the same beliefs and intentions is reasonably controversial.

Should such states be individuated by something like their character or rather by what is now called “content”? This divide, I think, will permit a distinctive variety of agent-centered normativity. If earth and twin-earth twins must share states of *belief*, those states reflect on their excellence as subjects. To the extent twins might differ in respect of another sort of state, for which we might also use “belief,” *those* externally-individuated beliefs may differ in excellence (and will *not* reflect on the excellence of their bearers).

We can distinguish three varieties of value: intrinsic value, instrumental value, and what we might call “coincident” value. Things are intrinsically valuable in virtue simply of the way they are intrinsically. Duplicates have the same intrinsic value; intrinsic value is an internalist feature. Things are instrumentally valuable, on the other hand, in virtue of the value to which they serve as instruments, the value they promote.

But we can consider also, if only as a foil, “coincident” value. All else being equal, true beliefs held under beautiful sunsets are better than those held under darkening skies. They have more coincident value. Neither instrumental value nor coincident value is guaranteed to be shared by duplicates. Now, for entities that have no subjective character, we tend to distinguish between instrumental and coincident value. A hurricane is good or bad, in a way, according as whether it produces bad outcomes. We typically evaluate the hurricane according to the status of phenomena causally and instrumentally dependent on it; but it would be idiosyncratic to evaluate the hurricane according to its coincident value. Similarly, actions and beliefs are normally evaluated with respect, modally, to their instrumental value or reliability; coincident value dangles.

With subjective entities, by contrast, there is a crucial distinction between intrinsic value, on one hand, and *both* instrumental and coincident value on the other. Subjective entities have intrinsic value; the evaluation of the agent merely in terms of instrumental value – if it is even coherent to speak of the *agent* as instrumental (recall the need for an account of agent causation) – is then like an evaluation in terms of coincident value. There is an agent-centered normativity that is independent of both the agent’s instrumental or coincident value.

Mind you, simply in virtue of being more than a mere bodily movement, any action reflects its agent’s intentions. This articulates the position developed here. The agent might have a more or less anodyne intention – to turn on the light, say. We may then get a full-fledged action, reflecting the agent’s state of mind to that extent. That agent turns off a light, which alerts a prowler, which in turn saves the family jewels. The agent’s intervention was crucial: it’s very good that the intentional action happened – it saved the family jewels. But the agent’s still nothing special, just lucky. Hurricanes are different that way. They *aren’t* agents; their status *is* determined by what they produce. Nagel’s compelling distinction between judging the agent and judging the happening has no echo here: there are no bad hurricanes, only hurricanes it would have been better that they not have existed.

We began with moral luck. That perennial philosophical issue dramatizes a contrast between varieties of normativity. The structure of the problem of Moral Luck and of the New Evil Demon exposes an internalist/externalist tension in our normative thought. A good response is to distinguish the *bearers* of normative *standards*. Subjects are apt for a distinctive sort of normative evaluation – to which nothing other than a subject can aspire. Other phenomena (such as states of belief or intention) are still evaluable along another sort of dimension (a dimension that *can* then be exploited to construct yet another, derivative, evaluation of subjects too). We thus achieve a clearer grasp of all the relevant

relations. Much of the clash between competing theories of normativity, in both epistemology and in ethics, can accordingly be quieted. Duplicates are necessarily good or bad alike – that’s a categorical quality. But their states of mind can differ in status – that status is relational. Anyone can *get* lucky; but there’s no such thing as *being* lucky.²

REFERENCES

- Hume, D. 1748. *An Enquiry Concerning Human Understanding*.
 Kant, I. 1785. *Groundwork of the Metaphysics of Morals*.
 Murdoch, I. 1970. *The Sovereignty of Good*. London: Routledge and Kegan Paul.
 Nagel, T. 1979. *Mortal Questions*. New York, NY: Oxford University Press.
 Smith, A. 1759. *The Theory of Moral Sentiments*.
 Smith, M. 1994. *The Moral Problem*. Oxford: Wiley/Blackwell.
 Sosa, D. 2016. ‘The Vice of Virtue Theory.’ In C. Mi, M. Slote and E. Sosa (eds), *Moral and Intellectual Virtues in Western and Chinese Philosophy*. New York, NY: Routledge.
 Thomson, J. J. 1989. ‘Morality and Bad Luck.’ *Metaphilosophy*, 20. Reprinted in D. Statman, 1993. *Moral Luck*, pp. 195–215. Albany, NY: State University of New York Press.
 Williams, B. 1982. *Moral Luck*. Cambridge: Cambridge University Press.

DAVID SOSA is Temple Centennial Professor in the Humanities at the University of Texas at Austin. Beside his work in epistemology, David has published in ethics, metaphysics, philosophy of language, and philosophy of mind, with papers such as “Consequences of Consequentialism,” “Pathetic Ethics,” “The Import of the Puzzle About Belief,” “Rigidity in the Scope of Russell’s Theory,” and (with Bonevac and Dever) “The Conditional ‘Fallacy’.”

2 The material in this paper has benefited from many colloquia, including at the University of Aberdeen, Kansas State University, the Universidad Nacional Autónoma de México, UNC Chapel Hill, Northwestern University, Princeton University, Soochow University, and the University of Wyoming, as well as at the XVIIth Congress of the *Sociedad Interamericana de Filosofía*, the 10th Annual *Midwest Epistemology Workshop*, and the 2016 *Episteme* Epistemology Conference, where Lauren Leydon-Hardy provided excellent comments. While it is impossible individually to thank all who participated in those colloquia, I am grateful.