**The Prescriptive and the Hypological: A Radical Detachment**

*Philosophical Studies*, penultimate draft

Maria Lasonen-Aarnio

**ABSTRACT** A wide range of more objectivist norms appear to leave uncharted an important part of normative space. In the beginning of this paper I briefly outline two broad ways of seeking more subject-directed norms: *perspectivism* and *feasibilism*. According to feasibilism, the ultimate reason why more objectivist norms are inadequate on their own is not that they fail to take into account the limits of an agent’s perspective, but that they are not sensitive to limits on what ways of choosing, acting, and believing are feasible in a given situation. I think of these ways of choosing, acting, and believing in terms of an agent's dispositions. In this paper I focus on a knowledge-first implementation of feasibilism, a view in epistemology that supplements a knowledge norm with a norm urging one to manifest dispositions that are among the most knowledge-conducive feasible ones. A view with different norms that sometimes cannot be jointly satisfied raises questions about the status of these norms. By drawing on two general hypotheses about the relationship between succeeding (e.g. knowing) and manifesting dispositions conducive to success, I argue for a view on which the *prescriptive*and the*hypological* come radically apart. The result is that an epistemic analogue of a thesis that many have assumed to hold in the moral realm should be rejected. This thesis is *Only Blameworthy for Wrongs:* we can only ever be blameworthy for acts that are morally wrong. I argue that on the picture presented, we can be epistemically blameworthy for doxastic states that do not violate any prescriptive epistemic norms. I then generalise the considerations to the moral realm, arguing against *Only Blameworthy for Wrongs*.

## 1. Feasibilism

It is widely agreed among normative theorists that we must distinguish between prescriptive assessments – assessments regarding what is required, permitted, or forbidden – and ones having to do with blameworthiness and praiseworthiness. However, the precise relationship between the *prescriptive* and the *hypological* is a matter of controversy. Most theorists recognise, for instance, that there are cases of morally blameless wrongdoing. But it is much more controversial whether one could be blameworthy without violating a moral prescription. Similar questions could be asked regarding the epistemic domain, assuming there to be something akin to epistemic blame.

My aim in this paper is to introduce and motivate a general normative framework, which I call *feasibilism*, and to sketch a view of the relationship between the prescriptive and the hypological in the epistemic domain by drawing on the theoretical resources provided by this framework. I will give a novel argument for a radical detachment of the prescriptive and the hypological. Though the main focus will be on the normative status of two kinds of epistemic assessments generated by feasibilism, toward the end of the paper I generalise the arguments to the moral domain.

Let me begin by explaining some of the general motivations for feasibilism, and by outlining its bare bones. The aim here is to paint a rather big picture, introducing some general considerations that will be drawn on later.

Consider a range of norms or standards, such as those listed below:

*Choose the best!*

*Keep your promises!*

*Believe p just in case p!*

*Only believe what you know!*

Many would agree that normative theorising cannot stop here. Such standards appear to leave uncharted an important part of the normative landscape. For instance, I might fail to know because I am in a Gettier case. I might fail to truly believe because my evidence happens to be misleading. Which choice is best depends on how things play out in the world much beyond me. And even facts about just what would constitute keeping a promise aren’t always that easy to track. Not all of normativity is so thoroughly subject to the contingencies of how the world and our relation to it plays out.

The push for what we might cautiously call more *subject-directed* norms and assessments is driven by a group of distinct motivations often run together. As already pointed out, sometimes we can only conform to a wide range of norms by fluke or luck. For instance, I may only be able to choose the best option by getting lucky. Here the notion of guidance is often brought in. At least sometimes, many have complained, we cannot be guided by norms like those listed above. And in any case, some normative assessments we want to and need to make are just not captured by these norms. On their own, they provide us with impoverished resources, failing to map out a part of normative space. For instance, the beliefs of victims of evil deceivers can be positively assessable, even if they fail to be true, and fail to constitute knowledge.

Though I am not moved by all of these motivations – in particular, the one appealing to guidance – I agree that a full normative picture needs assessments more attuned to what goes on with a particular subject, and less hostage to the fortunes of how the world plays out. Moreover, I tend to think that the kinds of positive assessment that norms like the above fail to capture cannot be captured just by appeal to blamelessness and excuses. We can all agree, for instance, that a subject in a Gettier case is blameless for not knowing. But something more positive can and should be said about their beliefs.

What more, then, can be said? We should begin by asking exactly why more objectivist norms like those listed above are inadequate on their own. For instance, why do they give rise to the problem of luck? Here is a familiar diagnosis, one that may even seem obvious. The problem with such norms is that their applicability in a particular situation sometimes depends on facts that lie outside a subject’s *perspective*. Either the subject lacks epistemic access to these facts, or they fail to figure in her mental states in an appropriate way. Either way, they are not within her ken. A perspective, as I am using the term, is a kind of representation of the world: it may consist of a subject’s beliefs, her evidence, a select class of mental states such as the seemings of phenomenal conservatives, or even the totality of things she knows, or is in a position to know. The hope, then, is that by perspectivising our norms, we avoid the kinds of problems faced by non-perspectival norms such as those listed above.

As obvious as it might seem, I do not think that the perspectivist diagnosis is correct.[[1]](#footnote-1) The fundamental problem is not that our perspectives are limited, but that there are limits on what ways of choosing, acting, and believing are feasible for us in a given situation. The above norms don’t take into account these limits of feasibility. Non-accidental conformity to a norm requires choosing, acting, or believing in a way that conformity to the norm robustly depends on. The applicability of any norm in a given situation depends on how things stand within some domain of facts. But for just about any domain of facts, in some situation there is no feasible way of choosing, acting, or believing that tracks how things stand within that domain. In such cases non-accidental conformity to the norm is out of reach.

The perspectivist and feasibilist diagnoses may at first sight look deceptively close. After all, the limits of a subject’s perspective often helps explain the limits of feasibility. For instance, I may not be able to make a choice in a way that tracks what is best because I don’t know a range of facts about the world. I may fail to choose the best way to get to work just because I don’t know that my normal cycling route has been closed early in the morning.

But the differences between the perspectivist and feasibilist diagnoses have wide-reaching ramifications for normative theory. Moreover, perspectivism doesn’t solve the problems that it has often been motivated by. In some situations it is not feasible for us to track facts about our own perspectives – our own beliefs, evidence, what we know, or even how things seem to us. That is why perspectivist norms themselves are susceptible to the problem of luck. Just like more objectivist ones, perspectivist norms leave a portion of normative space unmapped: we want to sometimes positively assess acts that fail to conform to a wide range of perspectivist norms.

Hence, my claim is that the correct diagnosis of why more objectivist norms are inadequate on their own, giving rise to problems like the problem of luck, has to do with limits of feasibility, not the limits of our perspectives. When seeking a more subject-directed kind of normativity, we should take into account limits on what ways of choosing, acting, and believing are feasible in a given situation. I prefer to think of these feasible ways of choosing, acting, and believing in terms of specific dispositions of the agent that manifest as acts like the making of a choice or the forming of a belief. Hence, we must take into account what dispositions can feasibly be manifested in a given situation.[[2]](#footnote-2)

The feasibilist does not reject more objectivist standards of assessment, but supplements them. For instance, I am sympathetic to there being a knowledge norm on belief. A more subject-directed level of epistemic normativity is gotten, then, by focusing on ways of believing that manifest dispositions that are conducive to knowledge. I understand the relevant kind of conduciveness in a comparative way, as being good in comparison with *feasible alternative dispositions*. Feasibility is a kind of possibility. There is much more to say about this notion of feasibility, but for our purposes here a close enough approximation is to understand it as taking into account human cognitive limitations. Feasible alternative dispositions are ones that it is not only feasible to have, but to manifest in the situation at hand. Dispositions that are *alternatives* are ones that would manifest in one’s situation as the having of a relevant doxastic state regarding a given subject matter, or as the making of a choice among a menu of available actions.

Faring well by a feasibilist standard of assessment is a matter of manifesting relatively success-conducive dispositions, dispositions that do well in comparison with feasible alternatives. The idea is to compare the manifestations of feasible alternative dispositions across a range of relevant cases, scoring them by using a value function centered on the relevant standard of success. Each disposition is assigned an overall score that is a function of the values of its manifestations across the relevant cases. The higher the score, the more success-conducive the disposition. For instance, in my favoured epistemic implementation of the feasibilist framework I deploy a gnosticist, knowledge-centric value function. A manifestation that is a knowledge-constituting belief gets the highest value, a manifestation that is a false belief gets the lowest value, the values of suspension of judgment and merely true belief being somewhere between the highest and lowest value. On the resulting view a disposition can be among the best feasible ones even if it is not a disposition to know. More generally, a disposition can be among those most conducive to a relevant normative success even if it is not a disposition to achieve that success. As a result, the account applies straightforwardly to states like suspending judgment: suspending judgment can be a manifestation of a disposition that is most conducive to knowledge or true belief in the relevant sense. And choices that knowably result in sub-optimal outcomes can still manifest the best feasible dispositions, dispositions most conducive to success.

In addition to assessments focused on some valuable standard of success and feasibilist ones focused on manifesting relatively good dispositions, the framework outlined can also be deployed to make a third kind of assessment, one having to do with whether a success is creditable to an agent. In this way the feasibilist framework can be deployed to make three distinct kinds of normative assessments:

1. Valuable success (e.g. *true belief, knowledge, morally right action*)
2. Manifesting dispositions that are among the most success-conducive feasibleones.
3. Creditablesuccess

Elsewhere I defend a view of what such creditable normative success involves (Lasonen-Aarnio *forthcoming* b). The right kind of modally robust dependence must obtain between the dispositions that one manifested and the normative success. These dispositions manifest across a wide range of relevant cases as norm-conforming doxastic states, choices, or actions.[[3]](#footnote-3) However, in what follows I will be focus on the first two kinds of assessments: ones tied to a relevant valuable success, and ones having to do with manifesting success-conducive dispositions.

My favoured implementation of the feasibilist framework in epistemology is gnosticist. The relevant standard of success is knowledge. A doxastic state is *reasonable* just in case it manifests dispositions that are among the most knowledge-conducive feasible ones.[[4]](#footnote-4) It looks, then, like I endorse the following distinct norms:

(K) Believe *p* just in case you know *p*!

(R) Be in doxastic state D just in case D is reasonable!

Knowledge-firsters often talk as if (K) is prescriptive: one ought only to believe what one knows, and in the absence of an excuse is blameworthy for believing otherwise. I will for now assume that this is right, though much of what I say about the normativity of manifesting good dispositions will remain intact even given a picture on which there are no authoritative prescriptive epistemic norms. But what about (R): how should we think of the normativity of reasonable belief? I discuss the relationship between success and success-conducive dispositions in more detail below. For now note that manifesting knowledge-conducive dispositions is not sufficient for knowing. Assume, for instance, that my evidence regarding some matter is massively misleading, but I have no inkling of this. Then, it might be reasonable for me to believe *p* – and not reasonable to adopt any other doxastic state – even if I violate (K) in virtue of holding a false belief. In that case I conform to (R), but not (K). And indeed, in so far as it is not feasible in my situation to hold no doxastic state whatsoever regarding *p*, it is then not feasible for me to conform to both norms. In what follows I will outline a picture on which matters of epistemic blameworthiness and praiseworthiness depend on the knowledge-conduciveness of the dispositions manifested by one’s beliefs. Such a picture, I will argue, need not view (R) as a prescriptive epistemic norm.

 In the next section I discuss the general relationship between succeeding and manifesting success-conducive dispositions. I argue for two general hypotheses that will serve as premises in the subsequent discussion.

## 2. Success and success-conducive dispositions

In what follows my focus will be on two central kinds of normative assessments generated by the feasibilist framework. There are, first, assessments tied to a success such as choosing the best, keeping one’s promises, believing the truth, and knowing. And then there are assessments tied to manifesting dispositions conducive to the relevant successes. I have understood such conduciveness as doing well in comparison with the best feasible dispositions.

What is the relationship between success and success-conducive dispositions? The arguments below will rely on the following pair of hypotheses[[5]](#footnote-5):

**Hypothesis 1: Competent Failure**

Manifesting good, success-conducive dispositions is not sufficient for succeeding.

**Hypothesis 2: Incompetent Success**

Manifesting good, success-conducive dispositions, or even failing to manifest deficient ones, is not necessary for succeeding.

I understand success-conduciveness in the feasibilist way outlined above, as being relatively success-conducive in comparison with feasible alternative dispositions. However, I also think these hypotheses are true on a different, non-relative reading of “success-conducive”. On such a reading a disposition is success-conducive just in case it manifests as a successful act (e.g. choice or doxastic state) across a wide range of relevant cases.

For a fairly wide range of candidate successes, these hypotheses are not particularly controversial. Consider, for instance, the successes of believing truly. Manifesting truth-conducive dispositions is not necessary for believing truly. A lucky guess, for instance, can be true, even if it is not a manifestation of dispositions conducive to true belief and indeed, even if it is a manifestation of disposition that tend to manifest as false belief. But neither is manifesting good dispositions sufficient. Sometimes we are unlucky to have deceptive evidence, even if there is nothing deficient with the way in which we came to have such evidence. For instance, if I come to form a belief based on reading a reputable newspaper that just happens to contain a misprint, then my belief is plausibly a manifestation of dispositions conducive to truth both in absolute terms and in relative ones, in comparison with feasible alternatives. Similar points apply to the success of choosing the option that is in fact the best. A random guess can lead to the best choice, and my evidence about which choice is best can be highly misleading.

Consider the success of keeping one’s promises. Assume that I promised to return a book of yours, but have forgotten which book I borrowed. Since I am very bad at keeping these sorts of promises, my bookshelf is full of books that belong to others. I now randomly pick out a book and give it to you, which just so happens to be the one I promised to return. This case falls under *Incompetent Success*. To see that manifesting dispositions conducive to keeping promises does not suffice for keeping a promise, consider the following case. Assume now that I promised to return your book, and have not forgotten which book that was. However, the cover of the book in question was swapped by a trickster with that of another book. I am not in any way negligent: I carefully placed your book in a particular place in my shelf, and check the cover to make sure it is the right one. However, I unwittingly return the wrong book (cf. Ross 1939: 147). Indeed, given my state of ignorance, it may be that I can only do what I promised by some strange twist of luck.

Consider now various perspectivist successes, such as proportioning one’s beliefs to the evidence, or believing or doing what one’s reasons support. *Incompetent Success* is not, I take it, very controversial: one can form a belief that just so happens to be proportioned to one’s evidence or reasons. More generally, one can conform to a perspectivist norm by luck. But more interestingly, the kinds of considerations already mentioned above show that *Competent Failure* also holds for the success of conforming to pretty much any perspectivist norm, a norm the applicability of which in a given situation depends on how things stand with respect to a one’s perspective.[[6]](#footnote-6)

Here is why. According to perspectivism normative facts depend on one’s perspective. Assume, for the sake of argument, that the normative facts are different depending on whether I have Perspective1 or Perspective2: in the former case I ought to φ, whereas in the latter case I ought not to φ. The consideration already alluded to is that sometimes it is just not feasible for us to believe (to choose, to act) in ways that track, across a relevant range of cases, how things stand with respect to our own perspectives. For instance, sometimes it is not feasible to track our own beliefs, our own evidence, what we know, or even how things seem to us. And, in particular, sometimes it is not feasible to track how things stand with respect to those perspectival facts that make a normative difference. Assume, then, that it is not feasible for me to manifest dispositions that discriminate between whether I have Perspective1 or whether I have Perspective2. The claim is that in a case in which I have Perspective2 instead of Perspective1, I might still be manifesting the best feasible dispositions – and indeed, that any of the best feasible dispositions might still manifest as φ ’ing. Hence, I might be manifesting dispositions that are among those most conducive to conforming to the relevant perspectivist norm, without in fact conforming to it. Such cases fall under *Competent Failure*.

 I have also argued that the two hypotheses hold for the success of knowing. The first hypotheses is not very surprising: one can fail to know, despite manifesting dispositions conducive to knowledge. A subject’s environment may be epistemically hostile, whether locally or globally, without this affecting the quality of her dispositions. A subject might fail to know simply because she is in a Gettier case, or because her evidence is misleading. I have argued that the same is true of some cases involving massively deceptive environments.[[7]](#footnote-7) Victims of massive deceit can manifest dispositions conducive to knowledge, even if the deceivers systematically block these good dispositions from manifesting as true belief or knowledge.

 But, one might wonder, how could *Incompetent* *Success* hold for knowing? Whatever knowing involves, I have argued that it does not necessarily involve manifesting dispositions that are conducive to knowledge.[[8]](#footnote-8) Knowing is no exception to the general rule that succeeding does not entail manifesting success-conducive dispositions. This is not to deny that knowledge paradigmatically comes about by manifesting knowledge-conducive dispositions. My preferred diagnosis of many putative cases of knowledge defeat rests on this idea: a subject can continue to know, but in retaining a belief, she is no longer manifesting dispositions that are among the best.[[9]](#footnote-9) This, I argue, does a good job in explaining the pull in two directions that many feel especially when considering so-called cases of defeat by higher-order evidence.[[10]](#footnote-10) In such cases a subject who in fact knows, and whose belief came about in a perfectly good way, acquires misleading evidence that her belief is flawed.

These points related to a distinction I draw between *epistemic* and *dispositional* discrimination. Epistemic discrimination is a matter of epistemic access: I know that I am in one region of logical space rather than in another. By contrast, dispositional discrimination is a matter of a kind of modal tracking, of dispositional responsiveness and sensitivity to differences between the relevant regions of logical space. It requires the feasibility of dispositions that discriminate between the regions by manifesting differently in them. These two kinds of discrimination can come apart. For instance, I know that I am not in the Matrix, even if I cannot dispositionally discriminate between my actual case from a counterfactual Matrix case. This does not, however, make my belief that I am not in the Matrix unreasonable, as such recherché cases are not relevant when assessing the reasonableness of my belief. I have also suggested that we diagnose some putative cases of knowledge defeat as ones in which epistemic and dispositional discrimination come apart. Consider a case in which I originally know a proposition *p*, but then acquire evidence from a reliable source that my belief in *p* is the output of a flawed cognitive process. In so far as I can retain knowledge of *p* in light of such putatively defeating higher-order evidence, perhaps I can also infer that the higher-order evidence is misleading. However, I cannot dispositionally discriminate my case from relevant cases in which my (higher-order) evidence is not misleading. And when assessing the reasonableness of the belief I retain in light of higher-order evidence that the belief is flawed, it *is* relevant to consider cases in which such evidence is not misleading. Such evidence from reliable sources is often not, after all, misleading. Because of this lack of dispositional discrimination, in retaining belief I manifest dispositions that manifest across relevant cases as retaining a botched belief. The dispositions I manifest when retaining belief are not among the most knowledge-conducive ones. These are cases of *unreasonable knowledge*.

I have sketched why *Incompetent Success* and *Competent Failure* hold for a wide range of successes. But in fact, the arguments outlined support a stronger kind of detachment of success and success-conducive dispositions. The arguments support the claim that at least for a very wide range of candidate normative successes, in some situations one *cannot* succeed by manifesting success-conducive dispositions: one’s circumstances are set up so that the best feasible dispositions will fail. One can only succeed by manifesting suboptimal dispositions. This is a stronger claim than *Incompetent Success*.

Consider, for instance, the successes of believing truly or of knowing, and the victims of evil deceivers mentioned above, whether ones plugged into the Matrix or deceived in some other way. Due to the machinations of the deceivers, none of the best humanly feasible dispositions manifest as true beliefs about the surrounding world. Our deceived twins can only believe truths by manifesting bad dispositions. The same is true of the candidate cases of unreasonable knowledge described: a subject with higher-order evidence that her belief is flawed can only continue to know by manifesting a kind of obstinacy, by manifesting dispositions that are not conducive to knowledge.

Similar points apply to perspectivist norms. A general consideration that already came up is that normative facts sometimes depend on differences between cases that it is not feasible for us to track, differences that it is not feasible to dispositionally discriminate. As a result, sometimes normative facts regarding what one ought to do can be changed without changing the facts about what would manifest good, success-conducive dispositions. Assume that according to a norm N whether one ought to φ depends on whether a proposition *p* is true or not. Assuming N to be a perspectivist norm, *p* is a proposition about how things stand with respect to one’s perspective. In case ***c*1** norm N requires one to φ (to believe a given proposition, to make a given choice, to perform a given action, etc.), since *p* is true in ***c*1**. Assume further that in ***c*1** all of the best feasible dispositions manifest as φ‘ing. In another case ***c*2,** however,*p* is false and hence, by N one ought not to φ. However, it is not feasible to dispositionally discriminate between cases ***c*1** and ***c*2**. Moreover, the best feasible dispositions still manifest as φ‘ing in ***c*2**.

Hence, on any candidate norm N, I claim, there will be pairs of cases of the following sort:

***c*1 *c*2**

*p*, and therefore on ought to φ.Not-*p*, and therefore one ought to not φ.

The best feasible dispositions The best feasible dispositions

all manifests as φ’ing. all manifests as φ’ing.

The general claim is that on a wide range of views, including perspectivist ones, there can be pairs of cases that one cannot dispositionally discriminate between and in which the same dispositions are feasible, but in which the normative facts regarding what one ought to do are nevertheless different. Further, in both cases the best feasible dispositions manifest as the same choice, belief, or action.

Because the general argument given is rather abstract, it might be helpful to look at a recently popular perspectivist way of thinking, one that deploys so-called knowledge filtering. The thought is that a range of normative facts depend on a subject’s *possessed normative reasons*, and that a proposition is among her possessed normative reasons just in case she knows the proposition, or is at least in a position to know it. On such a view a perspective is comprised of one’s possessed normative reasons. Let me consider, in particular, a view on which facts about moral rightness depend on one’s possessed reasons.

Consider first the following case:

***Beatrix***

Bill intends to brutally murder Beatrix’s family, including her children. In fact, he is already on his way to do so. Beatrix, however, has no reason whatsoever to suspect any of this. The only way Beatrix can prevent Bill from killing her children is by shooting him now. As she sees Bill on the street, she shoots him out of pure hatred and malice toward men.[[11]](#footnote-11)

On some views, Beatrix is morally permitted – perhaps even required – to kill Bill, for this is in fact the only way to prevent Bill from killing or seriously injuring her children.[[12]](#footnote-12) Indeed, on those views Beatrix’s case is one in which she cannot do what is morally right by manifesting dispositions conducive to doing what is morally right, or by manifesting morally good will.[[13]](#footnote-13) But on the knowledge filtering view envisaged, Beatrix is not permitted to kill Bill, for facts about moral rightness depend on possessed normative reasons, and Beatrix has no inkling of Bill’s intentions.

Contrast the case described above, *Beatrix*, with another case that I will label *Enlightened Beatrix*. In this case Beatrix knows that Bill is about to murder her children, and also knows that the only way to stop Bill is by killing him. Let us assume that since enlightened Beatrix possesses both the reason *Bill is about to murder my children* and the reason *the only way I can stop Bill is by killing him*,morality prescribes – indeed, requires – killing Bill. Doesn’t such a view avoid the detachment of success and success-conducive dispositions I have pushed? It does not, and it is instructive to see why.

Knowledge filtering makes normative facts – in this case, facts about moral rightness – depend on facts about what one knows. But, sometimes it is not feasible to track facts about what one knows.[[14]](#footnote-14) Compare *Enlightened Beatrix* with yet another case, *Gettiered* *Beatrix*. Beatrix is in a Gettier case. Gettiered Beatrix believes, but does not know, that Bill is about to murder her children and that the only way she can prevent this is by killing Bill. This Gettiered twin does not possess the reasons enlightened Beatrix does, though it looks to her as though she does. Indeed, we can assume enligtened and Gettiered Beatrix to be internal duplicates. Gettiered Beatrix may well possess reasons like *I believe that Bill is about to murder my children*, or even *It is likely on my evidence that Bill is about to murder my children*. However, such weakened reasons may not suffice to make taking Bill’s life morally permissible: after all, they do not speak as strongly in favour of killing Bill as the reasons that enlightened Beatrix has do. In any case, there will be some pairs of cases in which knowing allows one to possess reasons that render morally permissible actions which would not be permissible given the reasons possessed by one’s Gettiered counterpart. It simply cannot be that for every proposition *p* and for every possible case, *p* is no stronger a reason than reasons like *I believe p*, or *p is likely*. The kinds of pairs of cases described are bound to arise, and I will assume that *Enlightened Beatrix* and *Gettiered Beatrix* constitute such a pair.

By the knowledge filtering view of moral rightness assumed, enlightened Beatrix is morally required to kill Bill. Moreover, if enlightened Beatrix kills Bill because she knows this is the only way to protect her children, perhaps even lamenting the fact that doing so requires her to take a human life, she is manifesting good dispositions, dispositions conducive to doing what is morally right. By contrast, it is not morally permissible for Gettiered Beatrix to kill Bill. However, the difference between enlightened Beatrix’s knowledge of the relevant facts and Gettiered Beatrix’s lack of knowledge is not a difference it is feasible to track. A dispositional, motivational profile that manifests as killing Bill in the first case, but as not killing Bill in the second, is just not humanly feasible. The very same dispositions that manifest as killing Bill in the enlightened case will also manifest as killing him in the Gettier case, even though in the Gettier case such an action is morally wrong. Nevertheless, the action would sill manifest good dispositions, dispositions conducive to doing what is morally right.

Knowledge is, of course, just one filter that one might deploy. But similar arguments apply to any other filter appealing to some sort of perspectivist condition.[[15]](#footnote-15) The colour of the filter could be changed, but on any such view a fact is among an agent’s reasons just in case it is part of her perspective. Such reasons filtering makes normative facts – in this case, facts about moral rightness – dependent on facts about what is and isn’t part of an agent’s perspective. But no matter how one thinks of perspectives, I claim, sometimes it is just not feasible for an agent to track facts about her perspective.

 I have briefly stated the considerations I see as supporting both the two hypotheses *Competent Failure* and *Incompetent Success*, as well as an even stronger detachment between success and success-conducive dispositions. In what follows I will appeal to these considerations to argue for a radical detachment of the hypological and the prescriptive in epistemology. But before giving the argument, I will say a bit more about the prescriptive and the hypological.

## 3. The evaluative, the prescriptive, the hypological

There is a sense of ‘norm’ and of ‘normativity’ that comes rather cheap, one that I want to set aside at the outset. “Don’t form beliefs on Tuesdays!”. This sentence takes the form of a command: it prescribes not forming beliefs on Tuesdays. But we are not really required to do as it urges. Such a standard is not what we might call *normatively authoritative*.[[16]](#footnote-16) I will simply assume that there are normatively authoritative epistemic standards which include the kinds of epistemic assessments I have been concerned with. Both knowing and reasonably believing have normative weight not had by conforming to the standard of not forming beliefs on Tuesdays.

It is somewhat standard to think that normativity can be partitioned into distinct categories. There is the *deontic* or *prescriptive*, sometimes just referred to as the normative in a narrower sense, and then there is the *evaluative*, also sometimes referred to as the *axiological* (the Latin *valores* and Greek *ἄξιος* meaning ‘that which has worth’). The prescriptive requires, permits, or forbids. By contrast, while the evaluative may involve a kind of approval or disapproval, it has no direct deontic entailments. For instance, it would be good if I supplied the departmental coffee room with a selection of novels today, for such an action would add value the lives of my colleagues. It does not follow from this, however, that I am required to, or that I ought to, do this. Perhaps I am not even permitted to do so, for there are more pressing concerns like living up to my actual commitments. These different categories of normativity have different structural properties. For instance, it is rather widely accepted that whereas axiological properties and concepts are gradable – something can be more or less good – prescriptive or deontic properties are not. To these two normative categories some add a third, the *fitting*, irreducible to either the evaluative or the prescriptive.[[17]](#footnote-17)

I will also be using the term *hypological* – indeed, my main concern is with how the prescriptive and hypological are related, both in epistemology and in ethics. I borrow the term from Zimmerman (2002), who uses it to refer to judgments about moral responsibility. As I use the term, the hypological has to do with responsibility, whether moral or epistemic. Responsibility in this sense is closely connected with credit and blame: to deserve moral credit or blame for something, one must be morally responsible for it. Likewise for epistemic credit and blame, and epistemic responsibility. When we say that an agent has an excuse for acting wrongly and hence, that they are not blameworthy, we are making a hypological appraisal. Similarly when we say that they are creditable for acting rightly and hence, deserve praise for doing the right thing. In so far as an action’s having moral worth is a matter of the agent being creditable or praiseworthy for doing the right thing (e.g. Sliwa 2016, Lord 2017), assessments of moral worth are likewise hypological.

Even though prescriptive and hypological assessments are distinct, certain connections between the two are sometimes taken for granted. In particular, the prescriptive/deontic has often been assumed to be essentially tied to responsibility and blameworthiness in a way that the evaluative (or the fitting, for that matter) is not. At least in the absence of an excuse or exemption, we are accountable for violating prescriptive norms and hence, appropriate objects of blame. But further, blame is only appropriate when a prescriptive norm has been violated: in particular, a necessary condition on being an appropriate target of moral blame is that one has done something morally wrong or impermissible, something one ought not to have done. Indeed, many take this connection with blameworthiness to be a distinctive mark of the prescriptive/deontic, as opposed to the evaluative.[[18]](#footnote-18) On such a view blameworthiness is a mark that a moral prescription or requirement has been violated:

 *Only Blameworthy for Wrongs* (OBW)

S is morally blameworthy for φ’ing (not φ’ing) only if φ’ing (not φ’ing) is morally wrong*.*

As we will see, however, OBW is quite controversial.[[19]](#footnote-19) One could defend a corresponding thesis for praise: S is morally praiseworthy for φ’ing only if φ’ing is morally required.[[20]](#footnote-20) My main focus below will be on blameworthiness. Putting the above thoughts regarding blameworthiness and moral prescriptions together we get:

 *Blameworthy for Wrongs*

S is morally blameworthy for φ’ing (not φ’ing) just in case φ’ing (not φ’ing) is morally wrong and S lacks an excuse or exemption*.*

I am here assuming that something is morally wrong just in case moral norms prescribe against it. On this view, then, blameworthiness is an exclusive mark of the prescriptive: it is sufficient, and not merely necessary, for prescriptiveness. Negative prescriptive/deontic appraisal is a necessary condition for blame, and in the absence of an excuse or exemption, it is sufficient.

Below I sketch a picture of the relationship between epistemic responsibility and epistemic prescriptions that rejects the epistemic analogue of *Only Blameworthy for Wrongs*:

S is epistemically blameworthy for being in doxastic state D only if D violates a prescriptive epistemic norm.[[21]](#footnote-21)

Some are sceptical about the very existence of epistemic blame. I briefly address such skepticism below. Note that *Blameworthy for Wrongs* is compatible with the hypological and prescriptive coming apart in various ways. Indeed, I take it to be rather widely accepted even in the moral case that the hypological and prescriptive are distinct in that there are no entailments from wrongness to blameworthiness, or from rightness to praiseworthiness: there are both cases of *blameless wrongdoing* and of *praiseless rightdoing*.

Though there are many ways to argue for the existence of blameless wrongdoing and praiseless rightdoing, here is one that I find convincing. Strawson (1962) famously argued that our responsibility responses – on Strawson’s view, our reactive attitudes – are reactions to the quality of will or regard manifested by the actions of others.[[22]](#footnote-22) I won’t assume the full package of views defended by Strawson, remaining neutral, for instance, on the nature of blame itself. But I do find the following, rather minimal, claim to be plausible: we are only ever morally blameworthy or praiseworthy in so far as our actions manifest a certain quality of will or regard. One can only be morally blameworthy for an action if one’s action manifests ill will, or at least insufficient good will. Similarly, one can only be morally praiseworthy for an action if one’s action manifests good will and regard for others. Whatever else it may depend on, deserving credit or blame depends on the quality of one’s will.

This amounts to a very minimal version of a so-called *quality of will* view of moral responsibility. What I will assume is that the quality of one’s will is a necessary, though not sufficient, condition on blame- and praiseworthiness. I certainly don’t take this to entail that the domain of our responsibility is in our heads, or that we are only ever responsible for the quality of our own will. Our responsibility reaches far out into the world, but on the view I find plausible, it reaches out into the world only through the quality of our will.

My talk of quality of will here is a placeholder for something playing a given role in hypological assessments. The quality of will manifested by an action is a matter of the source of the action within the agent, and in particular those morally relevant properties or states of the agent that impact the action’s blame- or praiseworthiness. The substantive assumption I am making here is that the kind of valenced responsibility at issue in blame- and praiseworthiness is always a function of certain properties of the agent from which a morally assessable action stems. I here leave largely open just how we should think of quality of will.[[23]](#footnote-23)

What is the relationship between facts about the quality of one’s will and facts about moral rightness or permissibility? It is rather widely accepted that acting out of good will does not suffice to make actions right or permitted: facts about moral rightness at least sometimes depend on facts outside one’s will.[[24]](#footnote-24) This is on most implementations of the major approaches to moral rightness. Whether, for instance, an action maximises utility does not supervene on the will of the agent performing that action. Nor does whether an action constitutes killing an innocent person or breaking a promise.

Good will is not sufficient for doing the right thing, but neither is it necessary. It is widely accepted that it is possible to do the right thing “for the wrong reasons”, without one’s action manifesting good will or regard. The claim that not all morally right actions are morally worthy, where moral worth places constrains on one’s mental states not placed by rightness, is widely accepted.[[25]](#footnote-25) An action has moral worth just in case one deserves credit for doing the right thing, which in turn requires that it not be merely due to luck or accident that one did what was right. There will be cases in which one fails to manifest good will or proper regard – perhaps even cases in which one manifests ill will – but still does what is right. Hence, there are cases of praiseless rightdoing.

It is important to see, however, that the case for this detachment of the hypological and prescriptive – for cases of both blameless wrongdoing and praiseless rightdoing – can be made independently of such a quality of will view of responsibility. Indeed, a wide range of views agree on the possibility of both kinds of cases. Hence, I take a somewhat standard view of the relationship between the hypological and the prescriptive to be represented by

*Figure 1*:



 *Figure 1*

Note that ‘Permitted’ labels cases that are morally permitted but not required. The double check marks indicate categories that are wholly uncontroversial. As an example of a case in which one does something merely morally permitted that one is not praiseworthy for (bottom right corner of the figure), think of morally neutral choices like which route around the park to take whilst walking to work. The single check marks indicate widely recognized categories of cases in which the hypological and prescriptive are not perfectly aligned. These include cases of blameworthy wrongdoing and praiseless rightdoing discussed above. I also include cases in which one is praiseworthy for doing something that is merely permissible, but not morally required. Supererogation falls into this category. The categories with question marks are more controversial. *Only Blameworthy for Wrongs* entails that it is not possible to be blameworthy for doing something that is morally right – that is, either morally required or permitted. This, as will become clear below, rules out so-called *suberogation*. A thesis entailing that one can only be praiseworthy for doing what is at least morally permitted, in turn, rules out cases in which one is praiseworthy for doing something that is wrong.

In what follows I argue that my epistemological picture entails that the prescriptive and hypological come apart in a more radical way, allowing for epistemic analogues of all of the categories of cases represented in *Figure 1*. I will then revisit moral normativity, giving a structurally analogous argument for a radical detachment of the hypological and prescriptive in the moral domain as well.

## 4. A radical detachment of the prescriptive and hypological

We have practices of epistemic criticism that involve holding others responsible for their doxastic states, and these resemble our practices of holding others morally responsible. In what follows I will talk about epistemic blame and blameworthiness. One might object on the grounds that all blame proper is moral. But even if this is right, we need to be able to draw distinctions within the epistemic domain analogous to those that arise in the moral domain. For instance, some things might be epistemically good to do, but one is not in any way criticisable for not doing them. Instead of listening to some music to relax after work, it might be epistemically good to spend the hour refreshing my memory of 12th century history, triple-checking my tax returns, or working out some logical entailments of my beliefs. However, I am not epistemically criticisable for failing to do these things. Perhaps these examples even involve epistemic analogues of morally supererogatory action. Similarly, we need to draw a distinction between justifications and excuses, between conforming to epistemic norms and blamelessly flouting them (see Brown 2017, 2020a, 2020b).

However, I concede at the outset that epistemic blame does not take many of the forms that moral blame does: it is not clear, for instance, whether resentment or indignation are ever fitting responses to violations of epistemic norms, though other affective states, such as frustration and even varieties of anger, might be. Below I discuss the possibility that not only blame, but the nature of epistemic responsibility itself, is importantly different from moral responsibility. Indeed, here I am leaving the nature of epistemic responsibility largely open.

I will now argue that the epistemological framework presented supports a rather radical detachment of the prescriptive and hypological, one that entails the falsity of the epistemic analogue of *Only Blameworthy for Wrongs*. The argument will rest on two assumptions. The first is the pair of hypotheses already laid out above, *Competent* *Failure* and *Incompetent Success*. The second is what I will call a *quality of dispositions* condition on epistemic responsibility, a condition that is analogous to the quality of will condition on moral responsibility mentioned above. The thought is that we are only ever epistemically responsible for our doxastic states through the quality of the dispositions that those states are manifestations of. Moreover, good dispositions are dispositions that are relatively conducive to beliefs that are epistemically right: they are amongst the most success-conducive feasible dispositions. One might take such rightness to consist in truth, knowledge, being proportioned to the evidence, or something else. Hence, the argument does not rely on a particular understanding of epistemic rightness.

On my view there is an important connection between what doxastic states (choices, actions, etc.) are manifestations of an agent’s dispositions and facts about what she is responsible for, at least in a sense of attributability. Such attributability is widely taken to be necessary for responsibility in a fuller sense of accountability.[[26]](#footnote-26) It follows that one is only responsible for beliefs or other doxastic states in the first place that are manifestations of one’s own dispositions: it is only then that one can deserve credit or blame for them. Hence, one way to be blameless for a doxastic state is for it to fail to manifest one’s own dispositions. However, in so far as epistemic blameworthiness involves responsibility in a sense that goes beyond such attributability, I concede that more conditions may need to be imposed on the relevant dispositions, conditions that go beyond their being an agent’s own. These plausibly include conditions on how the agent came to have the dispositions in the first place. For instance, dispositions that have their source in neuropsychiatric disorders plausibly don’t satisfy these conditions. And perhaps dispositions acquired in closed fundamentalist communities sometimes don’t satisfy these conditions either. In such cases an unreasonable belief may be a manifestation of an agent’s dispositions, but the agent is not epistemically blameworthy for the belief.

Assume that we have settled what further conditions the dispositions manifested in an agent’s doxastic state must satisfy in order for the agent to be appropriately responsible for the doxastic state. The claim I am making is that the valence of the agent’s responsibility depends on the quality of the dispositions she manifests when forming, revising, and retrieving the doxastic state from memory. One can only be blamed for a belief, or for violating an epistemic norm such as a truth or knowledge norm, if the dispositions one manifested are deficient. Blameworthiness requires at least that one failed to manifest good dispositions, and it may require that the dispositions one manifested are positively bad, being significantly worse than the best feasible ones. Similarly, one can only be praised for a belief, or for knowing, if one’s belief is a manifestation of good, knowledge-conducive dispositions. The claim is that when one’s dispositions do satisfy the conditions for responsibility, the valence of one’s epistemic responsibility depends on the quality of the dispositions manifested.

Now consider a candidate array of normatively authoritative prescriptive epistemic norms:

(T) Believe *p* just in case *p*!

(K) Believe *p* just in case you know *p*!

(E) Proportion your doxastic states to the evidence!

What I want to now show is that the dispositions one manifests in one’s doxastic revision are in the domain of epistemic normativity *even if* we do not assume that there is also a prescriptive norm along the lines of (R):

(R) Be in doxastic state D just in case D is reasonable!

On my favoured view a doxastic state is reasonable just in case it is a manifestation of dispositions that are relatively knowledge-conducive when compared with feasible alternatives. However, there are alternative epistemic implementations of the feasibilist framework. Indeed, for each of the norms listed, we can interpret (R) in a corresponding way, understanding reasonableness instead as conduciveness to a different standard of epistemic success: to truth, or to proportioning one’s doxastic states to the evidence.

 This list of candidate prescriptive epistemic norms does not, of course, exhaust the options. But it is an instructive starting point. Irrespective of which epistemic standard of success one takes as a starting point, the picture I have presented predicts epistemic analogues of cases of blameless wrongdoing and praiseless rightdoing. But, in line with a more radical detachment of the hypological and prescriptive, it also predicts cases in which one is blameworthy for a doxastic state that is epistemically right in the sense of being permitted by prescriptive epistemic norms. It also allows for cases in which one is praiseworthy for a doxastic state that is wrong in the sense of violating epistemic requirements.

Let me begin with cases of *blameless wrongdoing*, and the norms (K) and (T). By *Competent Failure*, manifesting truth- or knowledge-conducive dispositions does not entail knowing or truly believing. Such cases plausibly arise, for instance, in mundane cases in which one’s evidence is misleading, supporting a falsehood. A false belief can be a manifestation of the most knowledge- and truth-conducive feasible dispositions. I have argued, in connection with the New Evil Demon Problem, that even victims of massive deceit can have beliefs that manifest dispositions that are conducive to knowledge and true belief. If a belief is reasonable – if it is formed and retained by manifesting dispositions that are among the most success-conducive ones – then one cannot be rightly blamed or criticised for not knowing. This follows from the quality of dispositions condition on epistemic responsibility. Manifesting knowledge-conducive dispositions provides a kind of excuse, a way to be blameless for false belief. Hence, victims of massive deceit provide us with epistemic analogues of blameless wrongdoing, of blameless false belief and hence, blameless violations of the (K) and (T) norms.

However, manifesting dispositions that are among the best feasible ones doesn’t merely provide one with an excuse. If we were merely looking for blamelessness or excuses, we could perhaps have considered subjects who are jetlagged or under the influence of drugs. Unlike other subjects with excuses, the kinds of victims of deceit under discussion are positively praiseworthy for at least some of their beliefs. This follows from the quality of dispositions view, assuming the relevant dispositions pass the required conditions for responsibility. We have here not merely a case of blameless wrongdoing, but one in which an agent is praiseworthy for doing something that is in fact wrong.

What about (E)? Sometimes there is no feasible way of tracking facts about one’s perspective and hence, no feasible way of tracking the application conditions of perspectivist norms. As a result, even the best feasible dispositions will sometimes manifest in choices, beliefs, and actions that fail to conform to the relevant perspectivist norms. So, for instance, sometimes manifesting the dispositions most conducive to doxastic states proportioned to one’s evidence or reasons does not entail succeeding in proportioning one’s beliefs to the reasons or evidence. Hence, there are also blameless – and, indeed, praiseworthy – violations of (E). The same points apply to other perspectivist norms, such as norms formulated in terms of possessed epistemic reasons.

My framework also predicts cases of *praiseless rightdoing*. According to the hypothesis *Incompetent Success,* manifesting success-conducive dispositions is not necessary for succeeding. In so far as this also holds for the success of conforming to the above norms, then it is possible to conform to the norm in a praiseless manner. Consider (T) or (E). A lucky guess or belief formed by reading tea-leaves can be true, proportioned to the evidence, and proportioned to one’s reasons. In such cases one conforms to the relevant norm without being praiseworthy for doing so. Further, one’s belief may be positively blameworthy. Indeed, it is possible to conform to these norms while manifesting very bad dispositions. Hence, in line with the more radical detachment of the hypological and prescriptive discussed above, we also have cases in which one is blameworthy for a doxastic state that conforms to these norms.

 (K) is different from the other norms listed in that it is possible to conform to (K) without knowing. One trivially conforms to (K) by suspending judgment on *p*, or by failing to form any doxastic state whatsoever regarding *p*. As a result, a disposition can be conducive to conforming to (K) without being conducive to knowledge. Consider, for instance, a Pyrrhonian sceptic who suspends judgment on almost everything. By systematically opting out of believing, the Pyrrhonian manifests dispositions that are among the most conducive to conforming to (K): by not believing one trivially conforms to (K). It does not follow, however, that the sceptic manifests the most knowledge-conducive feasible dispositions. Indeed, she would do better, given the value of knowledge, by forming beliefs about a range of ordinary matters, for those beliefs would constitute knowledge across a wide range of relevant, somewhat normal cases.[[27]](#footnote-27) On the gnosticist implementation of the quality of dispositions condition on epistemic responsibility that I have outlined, praise and blame depend on the knowledge-conduciveness of one’s dispositions, not on their conduciveness to conformity to (K). A sceptic who systematically suspends judgment on a wide range of matters despite the fact that knowledge is right at her fingertips is not praiseworthy for her suspension of judgment. Assuming (K) to be a correct prescriptive norm, then, some instances of suspending judgment should on my view be classified as cases of praiseless rightdoing – praiseless because the dispositions one manifests are not conducive to knowledge. And indeed, perhaps these are even cases in which one is blameworthy, despite doing what is right. Skeptical dispositions may fall very short of the most knowledge-conducive feasible ones.

Further, I find it plausible that there are cases of unreasonable knowledge. Some cases of unreasonable knowledge plausibly provide instances not only of praiseless rightdoing, but of doing something that is right but nevertheless blameworthy. I see no reason to think that knowledge is an exception to the general hypothesis that success does not entail manifesting success-conducive dispositions.[[28]](#footnote-28) Paradigm cases of knowledge are ones in which one knows precisely in virtue of manifesting dispositions conducive to knowledge, but it is not a condition on knowledge itself that one’s belief be formed by manifesting the most knowledge-conducive feasible dispositions. I have argued that at least some putative cases of knowledge defeat are in fact cases of unreasonable knowledge: once we see that subjects can be criticized on the basis of the kinds of dispositions they manifest – dispositions that, across a wide range of relevant cases, result in retaining a belief in a falsehood – we can no longer argue from criticizability to lack of knowledge.

On the picture that emerges, the normative significance of the kinds of dispositions one manifests when forming and revising doxastic states is evaluative and hypological, but not prescriptive. There is no need to view (R) as a prescriptive norm.

I now want to consider an objection to this picture, which is that if it is right, there is a structural disanalogy between moral and epistemic normativity. My picture entails that the epistemic analogue of *Only Blameworthy for Wrongs* is false: one can be blamed for being in doxastic states that do not violate epistemic prescriptions. But one might worry that this creates an objectionable structural divergence between the epistemic and moral domains, since in the moral domain the prescriptive and hypological are not thus radically detached.

First, we cannot simply assume that the moral and epistemic domains are structurally isomorphic when it comes to the relationship between the prescriptive and the hypological. For instance, many moral philosophers have tied together prescriptiveness with a kind of direct control, but it is widely agreed that we don’t have such direct control over our beliefs. Moreover, several authors have argued that epistemic blame is importantly different in nature from moral blame. For instance, Kauppinen (2018) draws a distinction between *reactive blame*, connected with Strawsonian reactive attitudes, and *relational blame*, connected with Scanlonian relationship modification, arguing that epistemic blame is relational, not reactive. Boult (manuscript)[[29]](#footnote-29) follows suit, arguing that epistemic blame is a distinctive kind of modification of epistemic relationships.[[30]](#footnote-30) If the nature of epistemic and moral blame is different, the nature of moral and epistemic blameworthiness, and the kind of responsibility involved, will likewise be different, and moral blameworthiness might be connected to prescriptiveness in ways that epistemic blameworthiness is not.

Even more ambitiously, however, I want to argue that a structurally analogous line of argument to that given above supports a radical detachment of the prescriptive and hypological in the moral domain.

## 5. The independence of moral rightness and moral blameworthiness

The moral analogue of my picture in epistemology allows both cases in which one is morally blameworthy for an action that violates no moral norms, as well as cases in which one is morally praiseworthy for an action that is morally wrong. Since I have focused more on the relationship between blame and wrongness, and less on that between praise and rightness, the discussion below will be focused on the former, and in particular, on *Only Blameworthy for Wrongs* (‘OBW’). If this principle is correct, then a subject cannot be morally blameworthy for doing something that is not forbidden.

Though *Only Blameworthy for Wrongs* is often taken for granted without argument, it is controversial, and rightly so. Indeed, many who explicitly discuss the principle reject it. Consider, for instance, the category of the *suberogatory*. Suberogatory actions are morally bad, but not morally wrong or forbidden. Numerous authors recognize the category of the suberogatory, arguing that in such cases agents can be blameworthy without doing anything that is morally wrong.[[31]](#footnote-31) While I am very sympathetic to these views, my case against OBW won’t rest on presenting putative examples of suberogatory actions.

I will now give what I see as the most convincing argument against the principle, which I will call the *Independence Argument.* The argument appeals to a kind of independence between facts about the quality of will manifested by an action and the prescriptive status of the action. This argument bears a structural analogy to the one presented above for a radical detachment of the hypological and prescriptive in the epistemic domain. In particular, it draws on premises that are structurally analogous to *Competent Failure* and *Incompetent Success*.

The argument will draw on a weak quality of will condition on moral responsibility of the sort already mentioned above. I do not assume, as some quality of will theorists do, that failing to manifest good will or regard, or even manifesting ill will, is a sufficient condition on blameworthiness. I only assume that it is necessary. There may be cases in which an agent manifests ill will, but is exempted from being held responsible. And perhaps there are even cases where the agent is excused for manifesting ill will. Psychopaths, sociopaths, and children can arguably manifest ill will, even though they are not thereby blameworthy (Cf. Watson 2004: 228). All that is needed is the assumption that in some cases an agent’s ill will, or lack of good will, can render an otherwise blameless act blameworthy. The act is rendered blameworthy when the other conditions for moral accountability are in place.

The Independence Argument relies on the following analogues of *Competent Failure* and *Incompetent Success*:

***Goodwilled Wrongdoing***

Manifesting good will is not sufficient for doing the morally right thing: morally wrong actions can manifest good will.

***Illwilled Rightdoing***

Manifesting good will, or even failing to manifest ill will, is not necessary for doing the morally right thing: morally right actions can manifest defective or even ill will.

As previously, by ‘morally right’ I mean either morally permitted or required. These two theses entail that there is a kind of modal independence between facts about the quality of one’s will and facts about moral rightness, just as *Competent Failure* and *Incompetent Success* entail that there is a kind of modal independence between facts about the quality of one’s dispositions and facts about whether one’s doxastic state is successful by various standards built into candidate prescriptive epistemic norms. The argument against *Only Blameworthy for Wrongs* will rely on *Illwilled Rightdoing* in particular. *Illwilled Righdoing* entails thatthe fact that an action manifests lack of good will, or even the fact that it manifests ill will, is not a sufficient condition for its being morally wrong. As pointed out above, according to a wide range of views, good will is not necessary for doing what is morally right: even an action that is morally required can fail to manifest good will. The claim now is that one can do what is morally right even while manifesting ill will.[[32]](#footnote-32) Note that the case against *Only Blameworthy for Wrongs* being made here is compatible with the claim that *some* actions are morally wrong in virtue of manifesting ill will; what is denied is just that manifesting ill will *always* suffices to make actions morally wrong.[[33]](#footnote-33)

Just what a case of illwilled, and thereby blameworthy, rightdoing would look like depends on one’s first-order moral views. I do not here want to take a stance on normative ethics beyond the claim that the correct first-order theory of moral rightness should allow for the kind of modal independence of facts about an agent’s will and facts about moral rightness or wrongness expressed in *Goodwilled Wrongdoing* and *Illwilled Rightdoing*. Moral rightness and wrongness are not just matters of the will: good will is neither necessary nor sufficient for the rightness of an action, and ill will is neither necessary nor sufficient for its wrongness.

It is instructive, however, to have a few candidate cases on the table. I urge the reader to bear in mind that my argument does not rest on intuitive verdicts about cases, but on general theoretical considerations. Nevertheless, here is one candidate case:

***Trolley***

James can divert an out-of-control trolley away from 500 trapped track workers and onto Paul. He diverts the trolley, however, not to save the 500 people, but solely in order to see Paul, whom James hates, die.**[[34]](#footnote-34)**

We can, of course, make the number of trapped workers arbitrarily high. I find it plausible that at least if there are enough people in the way of the trolley, James is not only permitted, but morally required, to divert the trolley. James very well knows that by diverting the trolley he would inevitably be killing Paul. Perhaps, then, he is morally required to kill Paul. Someone with a more perspectivist or subjectivist view of moral rightness can concur: after all, he James knows all the relevant facts about his situation. James’s motives can be further embellished. Perhaps, for instance, James has a deep-seated hatred of persons of a certain social or ethnic group that Paul is a member of. By diverting the trolley and killing Paul, James does what morality requires. But he is nevertheless blameworthy for doing so. Or, to give a different variant of the case, assume that James is convinced that a certain deontological moral theory is correct, perhaps reasonably so. The theory prohibits diverting a trolley onto one person in order to save the lives of even 500 people. However, James is malicious, and set on doing whatever is morally wrong. Due to his moral befuddlement, James does not succeed in his intention of wrongdoing, but his malicious intention nevertheless makes him blameworthy.[[35]](#footnote-35)

This is, of course, just one candidate case, and my argument does not rest on intuitions regarding particular cases. The general considerations appealed to by the Independence Argument are *Illwilled Rightdoing*, which amounts to the claim that manifesting ill will is not a sufficient condition on moral wrongness, together with the claim that in some cases of illwilled rightdoing the quality of one’s will – or, more generally, the mental states such as intentions, concerns, motives, desires and so forth leading to one’s action – renders the action blameworthy. I find the case for *Illwilled Rightdoing* so strong that I won’t discuss the possibility of rejecting it. However, the defender of *Only Blameworthy for Wrongs* could accept this, while rejecting the claim that in some cases of illwilled rightdoing the ill will manifested by an action renders the agent blameworthy for the action itself. In the above case, for instance, they could claim that James is not blameworthy for diverting the trolley onto Paul.

It is difficult to see how any such view could be reconciled with the thought that at least in the absence of an excuse or an exemption, an action that manifests ill will is morally blameworthy. For instance, James’s action of diverting the trolley does manifest ill will: after all, he diverts the trolley out of nefarious motives and intentions, and he has no excuse or exemption. In order to retain *Only Blameworthy for Wrongs*, it looks like we must reject the claim that at least in the absence of an excuse or exemption, an agent is blameworthy for actions that manifest ill will.

The proponent of *Only Blameworthy for Wrongs* might try to argue that James is not blameworthy for diverting the trolley, but that he is blameworthy for doing so out of nefarious motives. Alternatively, one might simply argue that what James is blameworthy for is the ill will itself manifested by his action. Either way, James is blameworthy at least for manifesting ill will. Note, however, that if James is blameworthy for his will or motives, and if *Only Blameworthy for Wrongs* is true, then it follows that James is morally required to manifest good will, or at least required to not manifest ill will: if James can only be blamed for what is morally wrong, and he is to be blamed for his will or motives, then morality must prohibit manifesting or acting out of such will or motives. It would be *ad hoc*, however, to accept this in response to my argument without accepting that the prescriptions of morality generally extend to an agent’s will.

In the following section I discuss a view on which morality, in addition to everything else it prescribes, prescribes manifesting good will, or at least not manifesting ill will. Part of my reason for doing so is to respond to the outlined attempt to save *Only Blameworthy for Wrongs*. But the discussion has independent interest in that it outlines a view of the normative status of good will in ethics that parallels my view of the normative status of good dispositions in epistemology.

## 6. Doing the right thing out of ill will

In the previous section I gave an argument against *Only Blameworthy for Wrongs*. If there are cases of illwilled rightdoing, and the quality of will condition on blameworthiness assumed is right, then in some cases an agent is blameworthy for an action that is morally permitted, perhaps even required. I now want to discuss the response that in such cases agents are not blameworthy for the morally right actions, but instead, for manifesting ill will. So, for instance, while James is not blameworthy for diverting the trolley, he is blameworthy for manifesting ill will, and perhaps for diverting the trolley from ill will. But then, by *Only Blameworthy for Wrongs*, it is morally wrong for James to manifest such ill will. To avoid an *ad hoc* view, the defender of *OBW* must accept that morality makes prescriptions concerning one’s will or motives: morality universally requires manifesting good will, or at the very least, it requires not manifesting ill will.

However, it is far from clear how such prescriptions concerning one’s will are to be reconciled with other moral prescriptions. Indeed, I will now argue that if the requirements of morality thus extend to one’s will, then its prescriptions will sometimes come into conflict: there will be cases in which one cannot both satisfy the moral requirements concerning one’s will and satisfy other moral requirements. The relationship between manifesting good will and doing what is morally right is structurally analogous to that between manifesting dispositions conducive to a valuable epistemic success and conforming to a prescriptive norm tied to the success in question. Just as success-conducive dispositions can have epistemic normative relevance even if there are no prescriptions to manifest such dispositions, one’s will can have moral normative relevance even if morality does not make universal prescriptions concerning one’s will.

Why would a universal prescription to manifest good will, or to not manifest ill will, come into conflict with other moral prescriptions? Because in some situations morality prescribes doing things that one just cannot do – that it is not feasible to do – while manifesting good will. And conversely: because in some situations morality prohibits doing the very things that manifest good will. Sometimes one is doomed to either goodwilled wrongdoing or to illwilled rightdoing.

Why so? *Incompetent Success* states that certain kinds of cases are possible: it is possible to succeed by various candidate prescriptive norms whilst failing to manifest good, success-conducive dispositions, and perhaps even by manifesting bad ones. But as pointed out above, the arguments supporting the principle support a stronger claim, which is that in some situations one *can only* succeed by manifesting suboptimal dispositions. In some situations none of the most success-conducive humanly feasible dispositions would manifest in successful choices, actions, or doxastic states. For instance, victims of evil deceivers cannot form true beliefs about the world by manifesting dispositions conducive to true belief. The same points, I argued, apply to perspectivist norms. For instance, sometimes one can only proportion one’s beliefs to the evidence by manifesting dispositions that are not conducive to proportioning one’s beliefs to the evidence.

The Independence Argument against *Only Blameworthy for Wrongs* appealed to *Illwilled Rightdoing*, which claims that it is possible to do what is morally right while manifesting ill will. I now want to make a stronger claim: there are situations in which one cannot do what is morally required by manifesting good will. Compare: there are situations in which one cannot succeed while manifesting success-conducive dispositions. Whatever the facts are about one’s mental life that enter into determining facts about the quality of will manifested, only certain combinations of those facts are feasible in a given situation. We could talk of feasible *motivational profiles*, talk of motivation being a placeholder for the states or properties of the agent that facts about blameworthiness and praiseworthiness depend on. In any given situation, only a limited range of motivational profiles are feasible. We can evaluate these profiles as good or bad, or better or worse, depending on the quality of will that they exemplify.

My hypothesis, then, is that whichever first-order normative theory is correct, it will give rise to cases in which the best humanly feasible motivational profiles manifest as actions that are morally wrong. Sometimes a subject just cannot do what is morally right while manifesting good will. She is doomed to either doing what is morally wrong, or to doing what is right while manifesting a defective will. And in some such cases, I claim, one’s will is defective enough to make blame appropriate.

For now, let the morally required actions be those prescribed by the moral norms that do not concern an agent’s will. The claim is that sometimes the facts that settle what one is morally required to do are independent and separable from the facts that settle what manifests good will. That is, the prescriptive facts can be changed without changing facts about how good will manifests itself. The hypothesis is that on any plausible view of moral rightness there will be pairs of cases of the following sort:

***c*1 *c*2**

*p*, and therefore one is required to φ.Not-*p*, and therefore one is required to not φ.

Good will manifests as φ’ing and no other Good will manifests as φ’ing and no other action. action.

In case *c*1one is required to φ, and whether or not one is required to φ depends on the truth of a proposition *p* that does not concern one’s will. While *p* is true in case *c*1, in an otherwise similar case *c*2 in which *p* is false, it would be morally wrong for one to φ. Whether morality requires φ’ing might depend, for instance, on facts about another person’s intentions, facts about exactly where and when one promised to meet a friend, facts about what the desires and needs of a loved one are, or even facts about the expected utilities of available actions. Assume further that in *c*1 good will manifests as φ’ing, and it is not possible for the agent to perform any of the other available actions while manifesting good will. The claim is that there is a possible case *c*2 in which *p* is false and hence, in which morality prohibits φ’ing, but in which good will nevertheless still manifests in φ’ing. That is, it is just not feasible to manifest good will while doing anything other than φ’ing in *c*2, even though morality requires not φ’ing. If there is, in addition, a norm that prescribes manifesting good will, then that norm will prescribe φ’ing. Morality will make prescriptions that one cannot jointly satisfy.

Could the problem be evaded by perspectivising moral norms? It cannot, for reasons similar to those already rehearsed. As an instructive example, consider again a perspectivist view that deploys knowledge filtering. On such a view facts about moral rightness depend on the normative reasons a subject has, and a proposition *p* is among the subject’s reasons only if she knows *p*. Knowledge filtering makes normative facts – in this case, facts about moral rightness – depend on facts about what one knows. But sometimes there are no feasible motivational profiles that track facts about one’s own epistemic position.

Recall the two cases described above, *Enlightened Beatrix* and *Gettiered Beatrix*. Gettiered Beatrix believes, but does not know, that Bill is about to murder her children and that the only way she can prevent this is by killing Bill. Hence, Gettiered Beatrix does not possess the reasons that enlightened Beatrix does, though it looks to her as though she does. As a result, though enlightened Beatrix is morally permitted and perhaps even required to kill Bill, Gettiered Beatrix is not morally permitted to do so. If enlightened Beatrix kills Bill because she knows this is the only way to protect her children, perhaps even lamenting the fact that doing so requires her to take a human life, her will is certainly not defective, and she is not blameworthy for killing Bill. Indeed, if anything, she would be blameworthy for not taking action to protect her children. By contrast, it is not morally permissible for Gettiered Beatrix to kill Bill. However, the difference between enlightened Beatrix’s knowledge of the relevant facts and Gettiered Beatrix’s lack of knowledge is not a difference that it is feasible to track. A motivational profile that manifests as killing Bill in the first case, but as not killing Bill in the second, is just not humanly feasible. In so far as enlightened Beatrix would be morally blameworthy for failing to protect her children, so would Gettiered Beatrix. It is just not feasible for Gettiered Beatrix to do what morality requires except by manifesting a defective will. Note, once again, that my argument does not rest on appeal to intuitions about a particular case. A view that deploys knowledge filtering is bound, I argued, to give rises to pairs of cases with the structure described.

Let me wrap up my discussion of a view that retains *Only Blameworthy for Wrongs* by adopting moral prescriptions that universally require manifesting good will, or at least ones that universally prohibit manifesting ill will. Since morality does not merely concern matters of will, such prescriptions will, I argued, sometimes come into conflict with other moral prescriptions: the two kinds of prescriptions will make requirements that it is impossible to satisfy. In some cases it is just not feasible to do as morality requires while manifesting good will, and it may not even be feasible to do as morality requires without manifesting ill will.

But why think that the requirements of morality cannot come into conflict? First, it is good to recall the broader dialectical context. It was argued that my view of epistemic normativity is problematic since if I am right, the hypological and deontic come apart in the epistemic domain in ways that they cannot come apart in the moral domain. *Only Blameworthy for Wrongs*, it was objected, is a widely accepted principle amongst moral philosophers. As pointed out above, numerous authors who consider the principle reject it. But further, an *Ought Implies Can* principle is itself widely accepted.[[36]](#footnote-36) And yet, I have argued, given plausible assumptions, the principle is in conflict with *OBW*. But my case here does not ultimately rest on assuming an *Ought Implies Can* principle. For every new layer of moral prescriptions that sometimes cannot be jointly satisfied with other, existing prescriptions, strikes me as a theoretical cost. I have argued that facts about the quality of will can have relevance for moral responsibility without there being universal moral prescriptions that concern matters of will. Quality of will can have normative relevance of both an evaluative and hypological kind even if morality does not issue in general prescriptions concerning the quality of one’s will.

I have sketched a broad view of how and why *Only Blameworthy for Wrongs* fails in the moral domain. The argument is structurally similar to the argument I gave for why the epistemic analogue of *OBW* is wrong. Nothing that has been said casts doubt on the thought that in the absence of an excuse or exemption, we are blameworthy for violating prescriptive norms. But if *Only Blameworthy for Wrongs* is false, then blameworthiness is not an exclusive mark of the prescriptive. We can no longer use being an appropriate object of blame as evidence that a prescriptive norm has been violated.[[37]](#footnote-37)

**References**

Arpaly, Nomy (2003). Unprincipled Virtue: An Inquiry Into Moral Agency. New York, Oxford University Press.

Berker, Selim (2022). ”The Deontic, the Evaluative, and the Fitting”. In Howard, Chris and Rowland, R. A. (ed.), Fittingness: Essays in the Philosophy of Normativity. New York, Oxford University Press**:** 23–57.

Capes, Justin A. (2012). "Blameworthiness Without Wrongdoing." Pacific Philosophical Quarterly **93**(3): 417–437.

Christensen, David (2010a). "Higher-Order Evidence." Philosophy and Phenomenological Research **81**(1): 185–215.

Fassio, Davide (2020). "On the generality argument for the knowledge norm". Synthese 197, 3459-3480.

Graham, Peter A. (2010). "In defense of objectivism about moral obligation" Ethics, 121(1): 88-115.

Graham, Peter A. (2011b). "'Ought' and Ability." The Philosophical Review **120**(3): 337–382.

Graham, Peter A. (2014). "A sketch of a theory of moral blameworthiness" Philosophy and Phenomenological Research, 88(2): 388-409.

Greco , John(2010). Achieving Knowledge: A Virtue-Theoretic Account of Epistemic Normativity. New York, Cambridge University Press.

Haji, Ishtiyaque (1998). *Moral Appraisability*. Oxford: Oxford University Press.

Harman, Elizabeth (2011). "Does Moral Ignorance Exculpate?" Ratio **24**(4): 443–468.

Harman, Elizabeth (2016). "Morally Permissible Moral Mistakes." Ethics **126**(2): 366–393.

Hirvelä, Jaakko (2019). "Knowing Without Having the Competence to Do So." Thought: A Journal of Philosophy **8**(2): 110–118.

Kamm, F. M. (2001). "Toward the Essence of Nonconsequentialism’".

A. Byrne, R. Stalnaker and R. Wedgewood (eds.), *Fact and Value: Essays on Ethics and Metaphysics for Judith Jarvis Thomson*. CambridgeMass.: MIT Press.

Kauppinen, Antti (2018). "Epistemic Norms and Epistemic Accountability”. Philosophers' Imprint **18**(8): 1–16.

Khoury, Andrew C. (2011). "Blameworthiness and Wrongness." Journal of Value Inquiry **45**(2): 135–146.

Lasonen-Aarnio, Maria (2019a). “Higher-Order Defeat and Evincibility”. In Skipper, Mattias and Steglich-Petersen, Asbjørn (ed.), Higher-Order Evidence: New Essays. New York, Oxford University Press**:** 144–172.

Lasonen-Aarnio, Maria (2019b). "Guidance, Epistemic Filters, and Non-Accidental Ought-Doing." Philosophical Issues **29**(1): 172–183.

Lasonen-Aarnio, Maria (2020). "Guidance and Epistemic Filtering." Acta Philosophica Fennica **96**.

Lasonen-Aarnio, Maria (2021). ”Dispositional Evaluations and Defeat”. In Brown, Jessica and Simion, Mona (ed.), Reasons, Justification, and Defeat. New York, Oxford University Press**:** 93–115.

Lasonen-Aarnio, Maria (Forthcoming a). ”Competent Failure and Victims of Deceit”. In Dutant, Julien and Dorsch, Fabian (ed.), The New Evil Demon Problem. Oxford, Oxford University Press.

Lasonen-Aarnio, Maria (Forthcoming b). "Perspectives and Good Dispositions." Philosophy and Phenomenological Research.

Lasonen-Aarnio, Maria (Forthcoming c). *The Good, the Bad, and the Feasible: Knowledge and Reasonable Belief*, Oxford University Press.

Liao, S. Matthew (2012). "Intentions and Moral Permissibility: The Case of Acting Permissibly with Bad Intentions." Law and Philosophy **31**(6): 703–724.

Macnamara, Coleen (2011). "Holding Others Responsible." Philosophical Studies **152**(1): 81–102.

Markovits, Julia (2010). "Acting for the Right Reasons." The Philosophical Review **119**(2): 201–242.

McHugh, Conor (2012). "The Truth Norm of Belief." Pacific Philosophical Quarterly **93**(1): 8–30.

McKenna, Michael (2012). Conversation and Responsibility. New York, Oxford University Press.

Parfit, Derek 1984. *Reasons and Persons*. Oxford: Clarendon Press.

Ross , W. D (1939). *The Foundations of Ethics*, Oxford: Oxford University Press.

Scanlon, T. M. (2008). *Moral dimensions: Permissibility, meaning, blame*. Belknap Press of Harvard University Press.

Shoemaker, David (2013). "Qualities of Will." Social Philosophy and Policy **30**(1–2): 95–120.

Smith, Michael (2007). “Meta-Ethics”, in Jackson & Smith (eds.), *The Oxford Handbook of Contemporary Philosophy*, Oxford University Press.

Strawson, Peter (1962). "Freedom and Resentment." Proceedings of the British Academy **48**: 1–24.

Thomson, Judith Jarvis (1971). "A Defense of Abortion." Philosophy & Public Affairs **1**(1): 47–66.

Thomson, Judith Jarvis (1991). "Self-Defense." Philosophy & Public Affairs **20**(4): 283–310.

Watson, Gary (2004). Agency and Answerability: Selected Essays. Oxford, Clarendon Press.

Whiting, Daniel (2020). "Epistemic Worth", Ergo 7(3): 105-129.

Zimmerman, Michael J. (1997). "A Plea for Accuses." American Philosophical Quarterly **34**(2): 229–243.

Zimmerman, M. J. (2009). *Living with uncertainty: The moral significance of ignorance*. Cambridge University Press.

Vranas, Peter B. M. (2007). "I Ought, Therefore I Can." Philosophical Studies **136**(2): 167–216.

1. See Lasonen-Aarnio (*forthcoming b, c*). [↑](#footnote-ref-1)
2. Within my epistemic picture dispositions play a theoretical role similar to the role played for process reliabilism by belief-forming processes, for some by the rules or methods one follows when forming and revising beliefs, and for others by talk of basing beliefs on evidence or reasons. The basic feasibilist idea could be developed, however, given some alternative view of ways of forming and retaining doxastic states. [↑](#footnote-ref-2)
3. My favoured view of this dependence is explanationist: creditable success requires that there be a robust explanatory generalisation linking the manifesting of the relevant dispositions and the success, a generalization that explains the success on this occasion. [↑](#footnote-ref-3)
4. It is worth emphasizing that I think there is knowledge that is not creditable to the agent: knowing can in this sense be a kind of fluke. One can in certain cases know even if the success of knowing does not robustly depend on one’s manifesting relatively knowledge-conducive dispositions. [↑](#footnote-ref-4)
5. These labels are easy to recall and capture the essence of the hypotheses, though the reader should bear in mind that my dispositions should not be thought of as competences à la virtue epistemologist: they are not dispositions to believe truly, or dispositions to know. [↑](#footnote-ref-5)
6. See Lasonen-Aarnio (forthcoming *b*, *c*). [↑](#footnote-ref-6)
7. See Lasonen-Aarnio (forthcoming *a*, *c*). [↑](#footnote-ref-7)
8. See also Hirvelä (2019). [↑](#footnote-ref-8)
9. See Lasonen-Aarnio (2010, 2021, forthcoming c). [↑](#footnote-ref-9)
10. Christensen (2010a: 193) nicely describes this. [↑](#footnote-ref-10)
11. The case is very close to one described by Capes (2012). [↑](#footnote-ref-11)
12. See e.g. J. J. Thomson (1991). [↑](#footnote-ref-12)
13. Cf Capes (2012). [↑](#footnote-ref-13)
14. For more detailed arguments, see Lasonen-Aarnio (2019b, 2020). [↑](#footnote-ref-14)
15. See Lasonen-Aarnio (2019b, 2020). [↑](#footnote-ref-15)
16. Cf. McHugh (2012) [↑](#footnote-ref-16)
17. See Berker (2022) [↑](#footnote-ref-17)
18. See e.g. Smith’s (2007: 5) and McHugh’s (2012: 9) characterization of the prescriptive/deontic. As Zimmerman (2009: 170-171) points out, such a thesis is often implicitly endorsed. [↑](#footnote-ref-18)
19. See e.g. Capes (2012), Fassio (2020), Graham, P. A. (2010: 93-94), Scanlon (2008: 124-125), and Thomson (1991: 295). The principle is often rejected for reasons quite distinct from those that I give below. For instance, Haji (1998) and Zimmerman (2009) argue that a subject can be blameworthy in virtue of acting against their judgment about what is morally right, even if what they do is not objectively morally wrong. But I do not, for instance, think that Huckleberry Finn is blameworthy in virtue of doing something he believes to be morally wrong. Parfit (1984: 25), for instance, distinguishes between subjective and objective rightness, claiming that an agent can be blameworthy for certain acts that are subjectively right but objectively wrong. My argument, however, does not rest on such a distinction. [↑](#footnote-ref-19)
20. Someone who wants to recognize the category of the *supererogatory* might want instead to defend the thesis that S is morally praiseworthy for φ’ing only if φ’ing is morally right, where being morally right is being either required or merely permitted. Which thesis for praiseworthiness one opts for won’t make much difference to the discussion to follow: ultimately I think that both should be rejected. [↑](#footnote-ref-20)
21. Whiting (2020) argues that there are epistemically right but criticisable beliefs, though he does not talk about blame. [↑](#footnote-ref-21)
22. Strawson (1962). [↑](#footnote-ref-22)
23. See Shoemaker (2013) for a discussion of different ways of understanding quality of will. I should note that commitments I make elsewhere (see Lasonen-Aarnio forthcoming *c*) entail that manifesting good will requires manifesting dispositions that are at least relatively conducive to doing what is morally right. Note that thinking about will in terms of dispositions manifested in one’s choices and actions does not rule out other accounts, such as thinking of quality of will in terms of an agent’s motives. [↑](#footnote-ref-23)
24. Cf Khoury (2011: 136). Of course there are those who disagree – perhaps, for instance, the twelfth-century philosopher Pierre Abélard, some Kantians, and some virtue ethicists (e.g. Slote 2001: 14). [↑](#footnote-ref-24)
25. See e.g. Arpaly (2003) and Markovits (2010). [↑](#footnote-ref-25)
26. See Watson (2004) for the distinction. [↑](#footnote-ref-26)
27. This is to make some assumptions about how the gnosticist value function works. Assume knowledge to have a positive value $k$, false belief to have a negative value $-f$, and suspending judgment to have a neutral value. I am here assuming that $f$ is not so much greater than $k$ that dispositions that invariably manifest as suspending judgment get a higher score than ones that often manifest as knowledge, but sometimes as false belief. [↑](#footnote-ref-27)
28. Lasonen-Aarnio (2010), see also Hirvelä (2019). [↑](#footnote-ref-28)
29. *The Puzzle of Epistemic Blame*, book manuscript. [↑](#footnote-ref-29)
30. See also D. Greco (2021), who draws on Kauppinen (2018). [↑](#footnote-ref-30)
31. For defences of the suberogatory see e.g. Driver (1992), McKenna (2012, Ch 8), Macnamara (2011, 2013), Schoemaker (2013), Mason (2017), Harman (2016a, b), Thomson (1971). [↑](#footnote-ref-31)
32. Perhaps one can be blameworthy without manifesting ill will, just by manifesting insufficiently good will. Indeed, cases of this sort were pointed out above in connection with the suberogatory. If this is right, then one can argue against OBW without appeal to *Illwilled Rightdoing*. [↑](#footnote-ref-32)
33. According to what Liao (2012) calls the *Intention Principle*, sometimes an agent’s intention can render an otherwise permissible act impermissible. Those who subscribe to the Doctrine of Double Effect, for instance, are committed to some such claim. Graham (2014) also discusses cases in which an agent seems to be violating a moral prescription in virtue of manifesting ill will. [↑](#footnote-ref-33)
34. The case is adapted from ones given in Graham (2011: 71) and Kamm (2001: 156). [↑](#footnote-ref-34)
35. According to Zimmermann (1997) acting freely while believing that what one does is morally wrong suffices for blameworthiness. I disagree with this general claim. [↑](#footnote-ref-35)
36. See e.g. Vranas (2007). [↑](#footnote-ref-36)
37. I am very grateful for discussions with Jessica Brown, Giada Fratantonio, John Hawthorne, Jaakko Hirvelä, Nick Hughes, Antti Kauppinen, Chris Kelp, Benjamin Kiesewetter, Aleks Knoks, Max Lewis, Giulia Luvisotto, Jack Lyons, Ram Neta, Mona Simion, Keshav Singh, Justin Snedegar, Ninni Suni, Teru Thomas, Teemu Toppinen, Barbara Vetter, Ralph Wedgwood, Alex Worsnip, Tim Williamson, audiences at the 5th Annual Chapel Hill Normativity Workshop (April 2023), Timfest (July 2023, University of Oxford), a workshop on my book manuscript *The Good, the Bad, and the Feasible* at the Human Abilities Center in Berlin (March 2023), another book manuscript workshop at the Cogito Epistemology Research Center at the University of Glasgow (May 2023), and participants at a graduate seminar at USC in October/November 2023. Special thanks to detailed comments from Jessica Brown, Jaakko Hirvelä, Antti Kauppinen, Benjamin Kiesewetter, Matt Vermaire, Tim Williamson, Alex Worsnip, and an anonymous referee. This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme under grant agreement No 758539. [↑](#footnote-ref-37)