All for the good

Article
Accepted Version


It is advisable to refer to the publisher’s version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1111/phin.12070

Publisher: Wiley-Blackwell

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR
Central Archive at the University of Reading
Reading’s research outputs online
1. Introduction

When Elizabeth Anscombe made her now-famous remarks in *Intention* about Milton’s Satan,¹ it would have been hard to predict that the thesis she was there defending, now known as ‘the guise of the good’, would become the subject of such intense debate in ethics and moral psychology.² There is disagreement about whether the thesis – which I will call Guise for short – is one about action, intention, reason and/ or desire. Philosophers disagree as to whether, if true, it is a conceptual truth about agency, a metaphysical truth, or a deep-rooted empirical fact about human nature. All of that before the discussion even begins as to whether Guise is true. This is a striking, but by no means unique, example of how Anscombe’s fertile philosophical imagination could, in but a few observations or comments, spawn an outpouring of debate exponentially larger than its original stimulus. Would that her depth and creativity had been more greatly appreciated in her lifetime.

---

¹ As well as about wanting a saucer of mud and sundry other examples of perverse action or desire. See Anscombe (1963): 71-5.
² For a representative sample of the literature, see Tenenbaum (2010) and the references therein.
I will line up behind the defenders of Guise, or the ‘scholastic view’ as Tenenbaum calls it, echoing Kant. My method will, however, be a little different to the usual. Launching into a to-and-fro about the interpretation of alleged counterexamples is bound to lead to stalemate. Nor can we determine the status of Guise by mere reflection on the concept of agency. Conceptual analysis will not tell us about what is ultimately a thesis concerning the deep metaphysical structure of action. The approach I will take is to locate Guise within an overall metaphysical picture of the nature of moral behaviour that links moral and non-moral senses of ‘good’. If we can ground moral behaviour in a more general account of what is good, then Guise necessarily comes out virtually as a corollary of that overall metaphysical picture. Needless to say, the general theory can only be given a partial defence here, but we will still gain more understanding of how the thesis is philosophically located, and how it might be true, than by treating it as a relatively isolated claim that does not rest on deeper foundations in the metaphysics of the good. It would be pleasant to think that the sort of indirect defence I aim to give is one with which Anscombe herself would have in general agreed.

2. Acting morally


\[\text{\textsuperscript{4}}\text{For further elaboration, see my ‘The Metaphysics of Privation’ (2014b) and ‘Being and Goodness’ (2014a).}\]

\[\text{\textsuperscript{5}}\text{For accounts that are both consistent with and complement my own, see Feser (2014) and Boyle and Lavin (2010).}\]
What do we want when applying metaphysics to ethics? I am now going
to assume a very thin and (at least to realists) unobjectionable version of moral
realism that is committed to the following two claims. First, to act morally\(^6\) is to
act in accordance with moral reality or moral truth: one’s action corresponds
with the way things are, morally speaking. Secondly, that correspondence
cannot be a matter of luck.\(^7\) What this means is that when an agent acts morally
– even if the act is no more than the mental act of willing, deciding or intending
something – this cannot be accidental, just the way things happened to turn out.

Moral behaviour – and I am happy to talk here of paradigmatic moral
behaviour, to allow the notion of ‘moral luck’ – is non-accidental.

How might it be non-accidental? We can easily think of highly implausible
ways. Some sort of pre-established harmony might be true: God, or His Moral
Majesty, or Evolution (with a capital ‘E’) has ensured once for all that all of our
moral behaviour is in accord with moral reality.\(^8\) Occasionalism might also be
proposed: the relevant force or agency always ensures, on each and every
occasion, that when we act morally our behaviour corresponds to the moral
facts. In the latter case it would be hard to see how Evolution could even be a
starter as the requisite force, but perhaps God or His Moral Majesty deigns so to
act time and again.

\(^6\) Where by ‘morally’ in this context I mean the opposite of ‘immorally’, not of ‘non-morally’.
\(^7\) Here I am clearly not talking about ‘moral luck’ as discussed by Nagel (1979) and Williams (1981),
which concerns the moral assessment of a person for something they did or omitted to do where the act or
omission was at least partly the result of factors beyond their control.
\(^8\) For a defence of Evolution’s role in this respect, see Enoch (2011): ch.7.
There is no space to discuss such proposals at length. My general objection to them is that the sort of causation they involve leaves no room for the ontologically substantive (as opposed to compatibilist) free will I am taking for granted. In other words, if some force or agency ensures that moral behaviour is non-accidental, not a mere matter of luck, then there is no space for any contribution from the agent to that non-accidentality. For if there were, the force or agency would not be doing all the work and there would remain something of the non-accidental character of moral behaviour to be accounted for.

It might be objected that pre-established harmony does not exclude free moral action: all it requires is that some force or agency set things up, even before the first-ever instance of human moral behaviour, such that all instances would ever after correspond to reality in the requisite way. But this will not do. Either the pre-establisher is the sole cause of every instance of moral behaviour or not. If it is, free will is excluded. If it is not, then it is something less than a cause – an enabler, or stage-setter, making it possible for free moral action to correspond to reality. If the latter, then the question of whether such correspondence is an accident is still live. The force or agency can set everything up so that moral actions correspond to reality, e.g. by creating that reality in the first place, by creating free human agents, by setting up laws of nature that ensure a certain regularity in worldly affairs, and so on. All of this
can still be true while the question remains open as to how, say, a person who

does what he ought is not merely lucky in so doing; more generally, how a

person can get an action morally right at all without coincidence or miracle.

One might object that a person can in fact be morally lucky in this sense

and still do what’s right. Suppose you promised to bring a cake to lunch with a

friend, then completely forgot your promise, made no effort to check, and so on,

but I – fortunately for you – remembered your promise and slipped a cake into

the boot of your car just before you left for lunch. Voilà: you kept your promise

but through no effort of your own. My reply is: of course, these things happen

often but they are precisely the exception to the fact that moral

behaviour that is far more common; it is this latter in which I am interested,

whatever we might claim about, say, doing the right thing in a thin sense (hence

my allowance for complexity at the margins). Should you be praised for

keeping your promise to bring the cake? In a qualified way if at all, but in

paradigmatic moral behaviour the praise is generally unqualified.

Another objection might go: ‘Well, non-accidental correspondence

between action and moral reality is just what moral action is! So what, exactly,

is the question?’ In reply, note first that not any non-accidental correspondence

will do, as indicated already in the case of occasionalism, pre-established

harmony, and Evolution. A well-meaning but deranged scientist might invent a

drug for universal administration, ensuring pharmaceutically that everyone did
what was morally good; but that would not produce genuinely moral behaviour, and even if you thought – wrongly – that such drug-induced rectitude was quite moral in the full-blooded sense, the fact is that moral action does not and has not taken place under such an influence, and what I am seeking to explain is virtue of the non-pharmaceutically induced kind. The second point is that what the objector indicates is that the question can simply be phrased as: how is moral action possible? That is fine, but it is only my question if it is interpreted as: how is moral action possible inasmuch as such action involves a non-accidental correspondence with moral reality?

The objector might think that my account is simply dispersed with an account in terms of the way things have been antecedently set up, e.g. by God or Evolution. Why isn’t it enough to appeal to the way the world has been arranged even if in the distant past and however this has happened (albeit it would be nice to know exactly how)? We might have additional questions about the mechanism itself and whether it needed further explanation, but as far as explaining moral behaviour in terms of correlation with reality, that such a mechanism exists – if it does – gives us all the explanation we want, so the objection goes. My reply is that the appeal to an antecedent mechanism, assuming now that freedom of the will can co-exist with such a mechanism, suffers a twofold defect: the mechanism is too remote and it is only necessary, not sufficient. Moreover, it is only necessary because it is too remote. By
remoteness I do not mean distance in time. Rather, the antecedent mechanism is not an intrinsic property of the agent. And because it is not intrinsic to the agent, and the agent has free will, it cannot amount to a necessary and sufficient condition of moral behaviour because if it were, it would bypass the agent’s faculty of freedom altogether, which is absurd. That is, it is absurd to suppose that what accounts for a person’s moral behaviour has nothing to do with their free will – more precisely, with the way their faculty of reason engages their free will. If the agent really is free, and their moral actions are not a matter of plain luck, then there must be something else going on to ensure the absence of luck, and this something else must in some way be involved deeply not just in the agent himself, as he is intrinsically constituted.

What I propose, then – and the thought is hardly novel – is that for a person to engage in genuine moral action, where the correlation between what they do and how things are is not a matter of luck or sheer coincidence, the agent himself has to employ a test or criterion. By its very nature, the employment of such a test will engage both the agent’s freedom and his reason. Whatever the ultimate source of human nature and of moral reality, for moral behaviour to be performed there has to be something the agent himself does such that he is able to align his performance with reality. Now morality has associated with it all sorts of tests and criteria, but they are not all equally fundamental. The more specific the case, the more specific the tests involved.
But since we are enquiring into moral behaviour at its most general – moral
behaviour pure and simple – the test must be appropriately general. It must also
be stable, in the sense of being consistently available for use assuming the

faculties needed for employing it are not themselves impeded. No one would
accept a clock as a reliable test of the time if one of its hands moved at a
variable speed; how much more should we not accept a test of morality that
waxed and waned in its applicability? Still, it would be too strong to require
that, as an epistemic matter, the test had to be reliable in the sense of producing
consistently true results. Metaphysically speaking, the test has to produce such
results, but lest we slip into the so-called ‘problem of the criterion’ we should
require only that, epistemically, the test has to produce results that are by and
large plausible across a broad range of moral questions, given common opinion,
and that this should be sufficient enough to make a prima facie case for its
results in more disputed areas. At the risk of being a little disappointing, I will
not have a lot to say here in defence of the plausibility aspect of the criterion, at
least in a concrete way; that task awaits another occasion. But I will defend
plausibility at a more abstract level, making some remarks of a relatively
applied nature along the way.

In addition, the criterion must be universal in the sense of being one that
both applies and is available to all human beings, since all people share a
common human nature and none are exempt from the requirements of morality.
Even though disparities in intelligence, wisdom, background, culture, and so on will produce different levels of knowledge and skill in applying a moral test, still the basic criterion, as a general formula, should not be beyond the reach of anyone. This rules out ab initio any criterion that in its very essence is overly intellectualized, elitist, inherently complicated, or impracticable (think Nietzsche or act utilitarianism, for example).

Now for a very brief sketch of the metaphysic that yields the sort of criterion of morality I defend. Suppose we define the good in terms of fulfilment of appetite, where ‘appetite’ is used in the traditional scholastic sense of natural tendency, inclination, desire, wish, desire, longing, desire, or the appetites of a thing belong to its nature. In this sense, humans have as much an appetite for, say, knowledge as they have for food. In a thing’s nature is the end to which its operations are directed, whether it be the essential behaviour of a water molecule or of a mammal. Everything in the entire universe (deity aside) is essentially restricted in its nature, and hence in its operations. Everything has a range of operation that is limited by its nature. Nothing can become anything, and nothing can do anything. Suppose we accept all of that (a lot to accept, admittedly). Suppose, in other words, that we concede the definition of the good, in general, in such terms. We do not even have to concede that the phenomenon of fulfilment of appetite applies across the board, to every entity that has natural tendencies. All we need to assume is that fulfilment occurs
across the spectrum of living things – that when they act according to their natures, i.e. when their natures are not thwarted, what happens to them in so acting is good. Hence the fulfillment of vegetative appetites – for growth, development, health – is good. So is the fulfillment of sentient appetites such as – with all the usual qualifications – pleasure and the avoidance of pain, the need for proper cognitive function, and the like. For rational creatures, the fulfillment of these and also the specifically rational appetites is good – whether for knowledge, friendship, beauty, and so on through a fairly familiar (albeit contentious) range.

On this highly general conception of goodness, whatever satisfies a genuine appetite – one belonging to the very nature of a thing – is good for that (living) thing. Goodness simpliciter is just fulfillment of appetite. Now if we accept this, could we then add that when it comes to that specific mode of goodness we call moral, we meet an egregious exception? Could we, for instance, declare that moral goodness was a different kind of goodness altogether, such that it resisted the general definition? Might we aver that moral goodness had nothing to do with fulfillment of appetite according to nature, whether because it had all to do with something radically different from this or because, say, it resisted definition altogether? This seems strange on the face of it. The reason is that we would, in either case, have to admit a systematic
ambiguity in the term ‘good’, which is intolerable if for no reason other than that it is one of the fundamental elements of any natural vocabulary, arguably one of the ‘human universals’. How could so general and universal a term be systematically ambiguous? Doesn’t the sceptic here owe us a principled reason why ‘good’ means something fundamentally different when applied to morality, a reason other than that the kind of moral theory one otherwise ends up with has to be resisted? If the theory we end up with is absurd on its face then perhaps we must bite the semantic bullet: otherwise, we should hold to what looks as basic a semantic fact as we can have, unless overwhelming argument forces us to abandon our grip.

Assuming that the meaning of ‘good’ in morality, at least in its most general aspect, is identical to its meaning outside morality, we must appeal to the fulfilment of appetite in defining the fundamental test or primary criterion of moral behaviour. But that cannot be the whole story, since as argued earlier reason and will must be essentially involved in the test. So I propose that what we end up with is the following formula:

The fundamental test of morality is whether an act is directed by reason to man’s ultimate end.\(^{11}\)

---

\(^9\) Even if ‘good’ has no meaning in the moral case – is strictly indefinable, albeit not meaningless; and it is the former that is my concern here – the term still ends up systematically ambiguous, inasmuch as in the non-moral case it does have a definition.

\(^{10}\) Brown (1991); see also the list in the Appendix to Pinker (2002).

\(^{11}\) For a typical scholastic statement, see Cronin (1930): 129ff.
Now the ultimate end is just another way of talking about the ultimate appetite or essential tendency (perhaps tendencies/apptetites in the plural) the fulfilment of which perfects human nature.

To appeal to the ultimate end is, from the ontic point of view, to dismiss the idea that there can be an endless series of appetites, each one such that its fulfilment is at the same time the means to the fulfilment of the next one in the series, where the next one will be broader, more general or all-encompassing.

To countenance the thought is effectively to deny that human beings can ever fulfil their natures, that they can ever be just good. Apart from the intolerable hopelessness this would inject into morality, it would mean attributing a kind of infinite nature to a manifestly finite being, which verges on metaphysical absurdity. From the practical point of view, the appeal to an ultimate end is just to endorse Aristotle’s famous doctrine that all practical reasoning must find a terminus.¹²

To say that there is an ultimate end is not to specify what that end might be. For present purposes it might be simple or complex, difficult or easy of attainment, mental, physical, or some combination thereof. It might be ultimate happiness or something else. More importantly, and sufficient I hope to bypass a certain kind of immediate reaction, the ultimate end may or may not transcend the natural order. In other words, human appetites might be such that their

---

¹² *Nicomachean Ethics/Ross (1925): 1094a19-21.*
complete fulfilment results only in natural perfection. If so, the ultimate end coincides with the natural end. On the other hand, it may be that the natural end is in its turn a means to the ultimate end, which ex hypothesi transcends the natural order.

This way of thinking about natural and ultimate ends bears directly upon how we should understand the primary criterion. Suppose, on one hand, that the natural end just is the ultimate end. Then the criterion is whether reason directs the act to the natural end. But suppose on the other that the ultimate end is not the natural end; suppose it transcends the natural order. The criterion would then be whether reason directed the act to a transcendent, non-natural end. How could anyone, whether the agent or an observer, know that to be the case in a specific instance? There is a real difficulty here. Assume, for example, that there were a transcendent, non-natural end for man: call it beatitude. Then the primary criterion of morality would be whether reason directed the act to the agent’s beatitude. The problem here is not one of knowing whether reason was so directing the act: presumably the source of beatitude would be more than capable of letting the agent know, through some means, whether reason was directing their act to beatitude. The question, rather, is what sort of thing the agent would have to do so as to be in the way of beatitude. What would they have to know about what they do? The sort of analogy I like to think of here is of a person who needs to get to a certain destination as soon as possible but can...
choose only route A or route B. If he asks a bystander which route he should take and is told, ‘Well, you need to take the quickest route’, he wouldn’t consider that much help. It’s not that the reply is vacuous; it’s just uninformative, merely a notational variant of the very question. If an agent wants to know, say, what is morally permissible in the arrangement of his personal life, and if his ultimate end is beatitude, it stands to reason that he needs to arrange it in such a way that he at least does not endanger his potential beatitude. But what, specifically, is he to do so as not to endanger this ultimate end? What has he got to go on?

The answer, I submit, is that he has to go on to his natural end, even if this is not his ultimate end. We can call it his proximate end, since although it is ultimate in the natural order it is not ultimate tout court. All the agent has to go on are the faculties, capacities and appetites of his and others’ common human nature. After all, the source of the ultimate end must have some responsibility for that human nature, that is, the source must have arranged some sort of harmony between the proximate and the ultimate end. In which case the proximate end has to be a guide – the ultimate natural guide – to the ultimate end. For what is the alternative? It would have to be that the proximate end was no means whatsoever to the final end, that the two ends were like non-adjacent jigsaw pieces that essentially did not fit together, or perhaps like
identically charged surfaces that essentially repelled each other. If the moral system has any fundamental coherence, this has to be ruled out.

Suppose, for example, that consideration of the proximate end leads one to consider truth-telling to be, at least on the whole, morally good. One might reason thus: we have the appetite for rational thought and communication. On the whole (at least), use of these faculties perfects human nature: it increases knowledge and understanding, enables the solution of manifold problems of human life, and produces generally harmonious social relations. So it looks like proper use of those faculties will be for the general communication of truth: the proximate end is fulfilled by such use. But it is supposed that the ultimate end, for its fulfilment, required the general communication of falsehood; that lying was the default behaviour tending to fulfilment of the absolutely final end of beatitude. We would rightly regard this as a monumental, cosmically cruel trick. Moral behaviour would be impossible short of an unending series of special revelations to each and every agent, on each and every occasion, informing them that what they needed to do in a specific case was the polar opposite of what their reason, as applied to their appetites, told them to do. The scenario certainly requires wild leaps of imagination to make it even barely comprehensible.

Think of a natural equivalent: reason and observation inform us that proper use of vegetative appetites tends to the increase of physical health. If we
use our mouths to take in food, and the right kind of food, we tend to be healthy; if not, we do poorly and may even die. Since physical health is part of our proximate natural end, it looks like right thinking to examine those natural faculties that work instrumentally to the fulfilment of that part of our end. But now imagine the reality to be that what really produced physical health was refraining from taking in food on a continuing basis. Indeed, in such a world what we call food would not actually be food. In such a world we still would have mouths, and stomachs, and intestines, and all the chemicals and processes that in our world turn food into nutrition. Yet any time we ate anything, we became ill or died; not because only actually poisonous substances existed in that world, but because what was actually nutritious was, in that world, poisonous. Such a world borders on the incomprehensible, but to the extent we can grasp it we would have to ask why, in that world, we still had mouths made for eating, stomachs for digesting, enzymes for breaking down actual food, and so on. How would all that be supposed to work? Certainly there could be no natural selection as it operates in the actual world: spandrels are one thing, suicidal faculties are another. We would need, not a theory of natural selection but a theory of natural deception!

We can play around with such scenarios, but I submit that whatever the ultimate end may be, reason can only direct an act to it via directing it to the proximate natural end. This must be so even if fulfilling the ultimate end
requires assistance outside the natural order, that is, if fulfilling the natural end does not suffice for fulfilling the ultimate end. Perhaps man has to be raised to a higher, spiritual level in order to fulfil the ultimate end. Maybe a liberal sprinkling of spiritual energy by the Moral Majesty is required to convert a purely natural act that for all its natural goodness is not perfectly moral, into one that is as morally good as it can be for a human agent. Still that extra ingredient can only work upon natural acts, acts emanating from human nature and directed by reason at natural perfection.

If this is correct as far as the relation between the ultimate and proximate natural end are concerned, we can apply the same reasoning to the connection between the natural end and the appetites that are instrumental to it. For example, suppose – quite plausibly – that mutual aid and assistance in society is an element of man’s proximate natural end. We can then examine the appetites that subserve that end, such as those involved in the development of social relations – the appetites for family, friends, reciprocity, trust, and so on – and test the moral goodness of acts involving these appetites in terms of whether they are directed to this essential element of the proximate natural end. By this procedure we would conclude that the wilful distortion or damage of these appetites was immoral if it did divert a person from contributing to mutual aid and assistance in society. Such distortion or damage
would be caused by acts harmful to family life, to friendship, and so on: it is not hard to imagine specific examples.

Working further backwards, we can investigate the appetites subserving those for family and friendship, none of which does – or need to – subserve family and friendship exclusively, and any of which can be – and often are – subserved reciprocally by family and friendship themselves.\(^{13}\) Think of appetites involved in communication, intimacy, confidentiality, reliability, and so on. We can apply the test again: does an act divert the agent from fulfilling these appetites such that the diversion also thwarts the fulfilment of appetites directed at family and friendship? Again, we can easily imagine the sorts of act that do frustrate such appetites, and we would generally agree on their moral wrongness.

Suppose, again, that we plausibly count health as an essential element of our proximate natural end. Whether it coincides in part with the ultimate end, or whether it is itself an instrument to a distinct ultimate end, we can test the moral goodness of acts by whether they frustrate the appetites subserving that natural end. We can think of appetites involved in diet, physical activity, work, rest, and the like, applying the test of whether acts involving those appetites thwarted the natural end of health. Needless to say, if I may indulge in caricature, most of us would agree that a person who overeats, does no exercise and sits on the sofa all

\(^{13}\) In other words, we do wrong to suppose that we can map a nice set of independent, linear relationships between appetites.
day watching television is engaging in acts that divert him from the relevant element of his natural end. Now it may be that some people will hesitate in calling such behaviour immoral because they think it is wholly self-regarding and that a person should be free, in such a case, to live their life as they please. There are two responses to this, one more interesting than the other. The first, less interesting, response is that such hesitation reflects, in my view, merely a psychological tendency to associate the term ‘immoral’ with harsh condemnation, judgmentalism, and even potential coercion. Yet we can, if we think about it, quite plausibly call behaviour immoral without engaging in any of these. We can correctly call our pizza-loving sofa-dweller immoral without being harsh or judgmental, let alone believing we or anyone else has the right to compel him to change his slovenly ways. As evidence of this, ask whether any person who refrained from using the term ‘immoral’ in such a context would also refrain, if asked for advice from said sofa-dweller, from advising him precisely to change his ways. If such advice is not moral advice, what other kind of advice could it be? Here the supposed distinction between morality and prudence has little bite. The sofa-dweller needn’t be asking for prudential advice. Suppose him to know full well what he ought to do if he cares about his health; and he may well care about it. He can still ask coherently whether he is living the right way pure and simple: he wants to know whether there is a moral issue here. If you advise him to change his ways on prudential grounds, but
decline to call his way of life a moral issue, then aren’t you telling him that what he is doing is morally permitted? In which case he is within his rights to take on board the prudential advice, admitting his full knowledge of what he needs to know if he wants to be healthy, while declaring that on moral grounds he is permitted to stay right where he is - on the sofa. Is that the sort of reaction you, as an adviser, would accept as proper?

More interesting, however, is the second reply: we should think that at least one distinct end served by health is that it renders us, on the whole, of greater service to others. This is something philosophers tend to overlook, yet it makes plenty of sense. For example, being generally well enabled to be healthy we are the less able we are in carrying out a variety of functions, most of which are required for being of service to others in quite basic ways. There is no paradox here: an appetite can do duty both as an element of man’s natural end and as an instrument to the fulfilment of other appetites that are also elements of that end. Think of how knowledge is plausibly both an end in itself and a means to the achievement of other ends such as beauty and spirituality, both of which should be regarded as parts of the overall end of human beings. Given this, we can apply our test of moral goodness by asking whether an act is directed to the achievement of health and, thereby, to facilitation of one’s being of service to others, where that service too is an element of the natural end, encompassing friendship, family, and mutual aid and assistance more generally. Here it is
probably much easier to see our sofa-dweller as indeed acting immorally, rendering himself not only of no service, but of positive disservice, to those around him.

Working back from the appetites subserving health, we can ask of lesser appetites subserving the former whether they too are thwarted by certain kinds of act. Think of how we might evaluate a person’s use of various parts of their body, or their mind, or for that matter of their time and their opportunities for action. As Aristotle might say, brilliant neologiser though he was, we are very familiar with the appetites associated with these kinds of use but we often have no name for them. SEE ABOVE FOR OFFICIAL

My overall point is simply that we can work back from ultimate appetites to lesser appetites and that in so doing, when we apply the primary moral criterion, we test whether the lesser appetite’s fulfilment conduces to fulfilment of the greater appetite. All along, the test is one of reason directing an act to the fulfilment of an appetite, an essential tendency in human nature.

What I have outlined so far is, to reiterate, a highly general account of the fundamental moral criterion, far less a vindication than a sketch of what it looks like in the light of the metaphysic I have been defending. It is hard to see what other criterion naturally emerges from that metaphysic. Moreover, again speaking at a very high level of generality, it accounts well for the most basic moral truths, by which I mean that those truths are easily converted into the
language of the criterion and are justified by it, where independent

consensus. Think of general duties concerning truth-telling, protection of life, pursuit of knowledge, respect for family and friendship, promotion of beauty, kindness to strangers, the ownership and use of property: the most general moral obligations in these areas are subtended by equally general faculties and appetites that define and delimit human nature, conditioning its fulfilment.

We cannot, however, get carried away with what can be shown at this level of abstraction, and especially within the brief framework outlined here. Two points should be noted. First, the primary criterion applies at the primary level – that of highly general moral truths and principles. It can tell us, for instance, about why we must have justice, why no one can be a mere means to another’s end, and why rights must be respected. It cannot of itself tell us what exactly counts as theft, when a punishment is inherently unjust, or how to handle lost property. For specific problems we need derivative criteria that draw their substance from the primary criterion but only mediately via a careful consideration of specific facts and hypotheses and the application of the primary criterion to relatively more precise appetites. Secondly, I admit to a certain amount of gerrymandering here, in the sense that I am consciously restricting the general truths to ones with implicit ceteris paribus clauses and qualifications. When it comes to truth-telling for instance, it is easy to see how a
consideration of fundamental faculties of understanding and communication receive their fulfilment through the conveyance of truth on the whole, and that on the whole this perfects human nature even in respects that do not directly involve those faculties. It is relatively harder to see whether or how the same considerations could yield the result that lying is always and everywhere to be condemned. In fact I think that careful consideration of vexed questions concerning the use and misuse of faculties and appetites will yield this result; but this must be a task for another occasion. In brief, we need a substantial system of derivative criteria, coupled with a thorough analysis of the strength and limits of the primary criterion, to advance us from the basic level of morality to something that is both plausible and practicable.

3. Acting immorally

What now of immorality? In what does moral badness consist? In sketching what it is to be moral, I set out a theory of orientation: to act morally is to be directed by reason to the final end. If there is such a thing as a ‘moral compass’, then without reason’s hand guiding the needle it is likely to spin out of control. To be morally good is for one’s thoughts and behaviour to be oriented towards pre-moral good. Without further metaphysical consideration, we would naturally add that to be morally bad – to act immorally – one must be
oriented towards pre-moral evil. One can see how the Guise of the Good immediately looks shaky on such a picture: if morality is about directing ourselves either to the good or the bad, it hardly seems a leap to suppose that we can intend the bad, aim at it, desire it – all as bad, i.e. under the guise of the bad.

Yet suppose we assume – as I have argued elsewhere – that evil is a kind of absence, namely the absence of good. It is grounded in positive states: sickness, for example, is grounded in various positive bodily processes. But it is not identical with those positive states: what makes a living thing sick is not the mere being in certain states, but that those states are a departure from, entail the absence of, a certain kind of normal functioning. The evil character of a state, process, quality, action, and so on, is in its departure from some norm, i.e. from the goodness that characterises the fulfilment of the appetites belonging to the thing in that state, undergoing that process, or performing that action. As such, evil is a kind of non-being grounded in, but not identical to, positive being. This view – that evil is a kind of privation – has a long history. I am now assuming it as part of the overall picture I want to sketch.

If we assume that evil is a kind of non-being in the way just outlined, the immediate question is: how can an agent be oriented towards it? The privation theory of evil holds that evil is a kind of non-being. As such, it seems hard to see how an agent could be oriented or directed by reason to non-being. And if

---

14 Where by ‘evil’ I mean the equivalent of ‘badness’, not something especially heinous or wicked.
15 Oderberg (2014b).
the falsity of Guise means that an agent can be oriented to badness, the privation theory of evil looks to be incompatible with Guise’s being false. The opponent of Guise might reply: we can orient our reason to all kinds of non-being – absences such as holes and voids, impossible entities, imaginary things. So why shouldn’t we be able to orient ourselves towards the non-being that we call evil – to think about it, intend it, plan it, will it, as well as bringing it about? In other words, even if the privation theory of evil is true, Guise can still be false.

Let us step back. The privation theory of evil holds that evil is an ontologically dependent reality. There is no evil without good, ontologically speaking. There is a deep and fundamental unity, which is immanent, mixed with some kind of good on which it supervenes, and unrelated to some kind of good in terms of which it must be understood. But this fundamental structure of reality has to be reflected in the way we comprehend both moral and immoral behaviour. If we are realists, we must think of moral behaviour as involving right reason – reason in accord with moral reality. But right reason in this sense just is reason directing the agent to the good. Moral behaviour, being itself a species of goodness, is then understood along the same lines as non-moral goodness – the fulfilment of appetite, only in the moral case it is the rational appetite, the will directed by reason, that is fulfilled by its orientation to the good.
So immoral acts, as a species of evil, must involve wrong reason, and wrong reason has to be understood as a *departure* from the good, just as pre-moral evil itself is a departure from, in the sense of privation of, the good. In other words, the species of evil that is moral evil is itself ontologically dependent on the good whose privation it is. This places limits on how we can legitimately understand the idea of orientation to evil. To go back to non-being in general, we do wisely to consider the different kinds of being and our possible attitudes to them on a category-by-category basis, which is obviously not a task that can be undertaken here. But take the case of holes. These are quite clearly ontologically dependent entities, however we precisely analyse them. There is no hole without a substrate material element subtending it. A hole is no more pure empty space than a privation is a pure absence. You might suppose you are able to think of a hole without thinking of its subtending matter, but is this really possible? You may say you can think of the pure geometrical shape of a particular hole without thinking of any material element; this may be true but then you are not thinking of the hole – by which I mean a real, concrete hole – but only of its shape, a property of it. Still less can you intend to make a hole without intending to alter some portion of matter, or to desire that there be a hole somewhere without desiring that some matter in that region be altered.
Now a hole is not a privation, unless one is thinking of a hole in the heart or in one’s bank account, but we can draw plausible analogies. Can an agent think of pure evil, plan or desire it? I would say – no more than they can bring it about. To cause an evil an agent has to remove a good. So they must, in some way and to some degree, acknowledge or recognise the good that they remove. The ‘must’ here has the force of a metaphysical law of the mind, a kind of psycho-metaphysical necessity. For a person not to think of the good of which they deprive a thing when they cause evil to it, they have to be insane or deluded. You might retort that this is just what psychopaths are – insane and/or deluded, yet prime examples of agents who orient themselves to pure evil. In fact, though, it takes little consideration of the evidence to see that most psychopaths are well aware of the good of which they deprive their victims. Any who do not, if this can even be verified, are in the narrower class of literally insane individuals with no responsibility for their actions; whereas my concern is with agency, not mere efficient causation. So an agent, I would argue, cannot orient themselves to evil per se – to evil in itself with no conceptualisation of the good whose privation it involves. What holds for physical action holds equally well for desire, planning, decision, intention, and the other agential phenomena. How can someone want evil to befall another without ipso facto wanting them to lose a good? How could they plan or intend this either?
Yet this is not the end of what it means, at the most basic level, to act
immorally. If I have established anything, it is only that an agent cannot be
oriented to pure evil in the sense just explained. One might object that
orientation to evil, correctly understood, is still a real phenomenon. If I have
placed any limit on how this should be comprehended, it is not especially
restrictive. I think we need to go further. As I claimed, immorality is a
departure from the good, but how can an orientation be a departure? My
response is that, in a sense, the objection is correct: if we try to explain moral
evil as a departure from pre-moral good constituted by an orientation to pre-
moral evil, even if we establish that there is not an orientation to pure evil, then we
must fail. Here Aquinas’s famous carpentry analogy is highly illuminating.\footnote{De Malo q.1 a.3 resp., Aquinas (2003): loc. 1745-62.}

Suppose (my modified version of the analogy) a builder is instructed to build a
vertical wall, and his measure for doing so is a plumb line. Suppose he then
goes ahead and builds a wall at an acute angle to the ground. Clearly something
has either changed or gone wrong with the situation. To explain why he did
what he did we need to explore the possible explanations, and these are
relatively few in category. One category is that of accident, within which we
can propose, for example, ignorance of fact (he thought it was lead at the end of
the line but it was a lighter and less reliable material) and force of circumstance
(the ground imperceptibly subsided as he was building the wall). Another is
lack of skill: he wasn’t adept at using a plumb line, working with brick, and the like. A third is change of mind: the builder decided to build at an angle after all because he and/or his employer changed the construction plan. The fourth is the hardest to account for: the instruction did not change, nothing intervened, he lacked no requisite knowledge or skill, yet he still built the wall at an angle.

How do we explain that, assuming that the builder did not lose his mind? To suppose that he intended to act perversely in the sense of simply not making the wall vertical stretches credulity. Can he be considered a *rational* agent if his aim was solely to make an angled wall while fully aware of the instruction to make it vertical? In other words, can we say that the agent is oriented towards having an orientation to the non-verticality of the wall?

We all know that there are only two roads to Scotland – the high road and the low road. If Fergus decides not to take the high road, can we consistently maintain him to be both rational and oriented, pure and simple, to not taking the high road? The thought that we can is specious even on a thin conception of rationality according to which all the agent needs is a plan and the ability to execute it. Leaving aside, again, such states as ignorance and carelessness, we must – conceptually and metaphysically – think of Fergus executing his intention of not taking the high road by taking the low road. In other words, there is *something else* he must do in order not to take the high road. Why can’t we describe his taking the low road as a way of being oriented
to not taking the high road, pure and simple? Because for him to be accurately
described in this way, he would have to leave the bounds of even thin
rationality: literally, he would have to be caught staggering about off the high
road, wandering aimlessly, such that when asked what he was doing, he replied
‘I’m doing whatever it takes not to be on the high road’. Does this make any
sense if we are not to account Fergus crazy? If he is in fact taking the low road,
then we can be fairly sure that if we ask him what he is doing, he will reply
something along the lines of, ‘I’m trying to get to Scotland and I’ll be turned
into haggis before I take the high road’. Clearly he has another objective in
mind, one he deems desirable or worthwhile, a further part of his plan.\(^\text{17}\)

To return to our wayward builder, if we discount the usual explanatory
factors in his constructing the wall at an angle, we have to posit some other
objective he had in mind, something else he deemed good even if it was not
really good in the overall context of the situation. Such might be, for instance:
the pleasure of disobedience to instruction; the thrill of working impromptu and
without due attention to detail or employment of all the builder’s available skill;
the good to be obtained by working quickly so as to get the job over with,
enabling him to attend to some other matter; and perhaps other objectives,

\(^{17}\) It need not, however, be a ‘further intention’ in the sense of a ‘forward-looking’
tention, which is the sense of ‘further intention’ that Anscombe sometimes – but not always –
means (1963: 31). Anscombe speaks of the intention ‘with which’ an agent does something,
not the intention for which they act, which generally implies a future-oriented objective. It is
implausible to say that the intention for which Fergus takes the low road is that of not taking
the high road, since by the very act of taking the former he does not take the latter. So his
further intention here is the intention with which he takes the low road.
though it is hard to imagine many distinct categories capable of explaining such perverse behaviour.

An objector might insist that I am conceiving of immorality exclusively in terms of analogies involving alternative actions. Within such limits the analogies might illuminate, but the limits distort the facts. After all, one can be immoral by doing nothing. So why not recast the analogies? Fergus can sit down at the entrance to the high road and read a book, determined never to take that road. The wayward builder can down tools and take a long lunch, determined to clock off before he ever builds a vertical wall. It would be a cheap and irrelevant rejoinder on my part to point to the pleasures of reading or eating as constitutive ulterior objectives on the part of the respective agents. Fortunately, reading and eating have nothing to do with the correct reply, which is that there is inaction and there is inaction. The inaction of the person who freezes like a deer in the proverbial headlights, who is incapable of making a choice or a decision, or who stares blankly into the horizon does not even make it into the voluntary let alone the intentional (assuming for the sake of it that there is even a difference between the two). He might as well stagger about as Fergus could have, whether inebriated or not. Such inaction is of no immediate concern in a theory of immoral action (though it might be in a more specific theory of moral virtue). By contrast, the inaction of the builder who downs tools with the objective of not building the vertical wall as instructed is the inaction
of an agent who might as well have acted in the more familiar sense of intentionally moving his body. From the moral point of view – from the viewpoint of immorality, for our specific purpose – there is not a bit of difference. We still have to cite some other objective – some pleasure or other, some satisfaction, contentment, achievement, and so on – as that to which the physical inaction is oriented.

So to return to the question of how an orientation can be a departure, my answer again is that it cannot if we think a departure from the good can be cast in terms of an orientation to evil. It can, however, be a departure if the orientation is to some other goal which, if it were actioned, according to the relevant moral criteria, a lesser good. In defending the Guise of the Good, ethicists (at least of the Aristotelian-Thomistic school) often speak of the rational appetite always tending to a real or apparent good, as per Aristotle’s dictum. In the specific context of understanding immorality, what I claim is that we should not interpret ‘real or apparent’ to mean that an immoral agent can be oriented to a good that is apparent but not real, in the following sense. An immoral action will always be directed at something thought good – an apparent good. But that good must also be a real, a genuine pre-moral good. Yet for the agent to be immoral, the force of being apparent must be that he thinks this pre-moral good is not only a good, but the good to be done in the situation. He may think it is the morally good thing to be done: witness the cruel dictator who
thinks exterminating millions of his subjects is the *right* thing to do because, say, it maintains the stability of the state, or furthers the primacy of the select racial or tribal group to which he belongs. But the immoral agent may simply think this real, pre-moral good is the overall good he wants to achieve, whether through narcissism, selfishness, lust for power, or any other of the innumerable vices that might be in play. What I want to encourage is the thought that even in the case of immorality the agent must be oriented to something that is a real good, on top of which he must see it as good in some respect. But how could, say, a tyrant’s furthering the primacy of his racial group be in any way a pre-moral good? The fact that furthering the primacy of any group has furthering *some* elements of its members’ well being as an essential part, and furthering anyone’s well being is a good, even if it is not morally good, the *right* thing to do in the situation – or maybe any situation, if we think of racial primacy. In other words, a good may not be *the* good to achieve in a situation. But there must be some pre-moral good underlying the immoral action as its target of orientation, lest the orientation be to evil, which if what I have argued is correct makes the explanation of immoral action all but impossible.

Must we, then, give up on the very idea that evil can be intended? How can an agent intend evil if they cannot be oriented to it? Moreover, if privations, as a kind of absence, have no causal powers of their own, how can they have the power to be caused either? We do not have to yield to thoughts so at odds with
WARNING – AUTHOR COPY ONLY. NOT OFFICIAL PUBLICATION. NOT FOR CITATION. OFFICIAL VERSION IS PUBLISHED IN: Philosophical Investigations 38 (2015): 72-95

common sense and sound moral principle. To take the second question first, privations might not have powers to be affected by anything, but the underlying positive reality on which a privation supervenes does have such powers, and is indeed causally produced by some other positive reality. (Germs cause disease, a hammer can break a leg, etc.) We can properly correctly of a privation’s being caused, but only in this sense. Similarly, privations do not have causal powers to produce anything, but we speak correctly of their causing things only inasmuch the positive states on which they supervene have the relevant causal powers and do the actual causing. (Disease causes unhappiness, a broken leg makes one unable to walk, etc.)

Taking the lead from this idea, we can go on to say that if a person causes evil, this can only be the case inasmuch as the person causes some positive reality on which evil supervenes. But then if a person can truly be said to intend evil, that can only be because they intend to produce some positive reality on which they know the evil supervenes (not that any special philosophical knowledge is required for the agent to apprehend this). So an agent cannot intend evil without intending to produce some positive reality which, by its very nature as positive reality, cannot be evil (since evil is essentially an absence of good). Now on the general view I defend, all positive reality is good in the pre-moral, basic sense of involving fulfilment of appetite –
the actualization of some potentiality, to use the scholastic terminology. It follows that an agent cannot intend evil without intending some good.

So far so quick, however. We need to get clearer on what exactly we mean by intending evil, with the aim of showing that intending evil is consistent with Guise. What we are interested in here is immorality – the intentional departure from the good. It is moral evil I am trying to illuminate. Now since moral evil is a kind of evil, it too is a privation – that privation which is a departure from the good. So moral evil – the departure from the good – must supervene on some positive reality, hence on some good since every positive reality is good in some respect. For the good on which the pre-moral evil involved in the immorality supervenes. It is that pre-moral evil which supervenes on that pre-moral good. If a parent starves her child, the child’s illness supervenes on the positive state of their body – the mixture of actuality and potentiality which we truly call malnutrition. So on what does the departure from the good that is moral evil supervene? On the positive reality – the good – that the agent intends as an alternative to the good they should be causing. So if the parent starves her child, it must be for some good – not the good, to repeat, but a good – she intends to achieve, whether it be making a cry for help, attending to some other need, or simply the perverse pleasure in watching a child suffer. Here, through purely metaphysical considerations
concerning the analysis of moral evil as privation, we come to the same result argued for earlier with the builder and Fergus analogies.

These considerations, though, also help us focus on how evil can be intended. Just as evil is ontologically dependent upon good, so the intending of evil depends ontologically on the intending of good. Just as evil has no nature of its own, no positive essence, so intending evil has no nature or positive essence of its own. One can only intend evil as an indirect effect of the good that one directly intends. What I mean by ‘direct’ here is that one directly intends only what is one’s end or the means to one’s end. This sounds strange since we usually think of ‘direct’ as synonymous with ‘unmediate’ or ‘immediate’, with no intervening steps. In this sense, all but so-called ‘basic acts’, if there are any, are indirect: we perform them by means of other acts. But the sense of ‘direct’ I mean here is well captured by the idea of ‘aiming at’ something, having a target in mind. To be sure we are still in the realm of terms of art, and something more precise has to be said for which there is no room here. I venture to suggest, though, that there may be an irreducibly phenomenological element at least to the identification of direct intentions but possibly also to their existence. That is to say, that there is a direct intention, and what intention it is, depend on what the agent has in mind as the object of his proposed action. This does not mean that an agent intends whatever it is he thinks he intends, or whatever fits the description under which he says he is acting: people can be mistaken about what
they think they are doing, and what they say they are doing can diverge from what they are in fact doing. Moreover, a constitutive phenomenological element to intention is consistent with the ruling out, on independent grounds, of certain putative objects of intention as even possible, metaphysically speaking.

If what I have argued is correct, we have independent reasons for thinking that an agent cannot directly intend evil, in the sense of having it in mind – evil as evil, evil for the sake of evil – as the object of a proposed action, whether as means or end. It is in this sense that I mean an agent cannot have an orientation to evil.\(^{18}\) Yet the truth that agents can intend evil, can do evil intentionally, is non-negotiable; that just is a fact about the world. We would have to say that all cases of apparently intentional evildoing were no more than cases of foresight at best, accident at worst.\(^{19}\) And the strongest degree of culpability for such acts could only be something like gross negligence, never malice. Moreover – and this is a point I want to stress – if evil could never be intended, we would have to reject the Principle of Double Effect, at least in its classical formulation, one of whose essential conditions is that the evil effect must never

\(^{18}\) When Satan says, in Book IV of \textit{Paradise Lost}, ‘Evil, be thou my good’, note that he is not saying, ‘Evil, be thou my intent’, or, in drier terminology, ‘Evil, be though my object’. Lest he be thought of as proposing a straight contradiction, the most plausible interpretation is that he is asking for evil to present something good to him, so that he may achieve \textit{that} and thereby achieve evil. Since there is no good in evil as such, all he can rationally ask for is some good or other, associated with evil, that he may pursue and thereby \textit{indirectly} bring about evil. Such goods are easy to find: pleasure, power, the satisfaction of his own will in disobedience to his creator; and so on. In my view Anscombe had this completely right, whilst her critics have consistently missed the point.

\(^{19}\) See \textit{Summa Contra Gentiles} III.6.7. Aquinas (1956): 46 where St Thomas asserts that moral evil does not occur ‘by chance’ (‘\textit{non sequitur causaliter}’ [fortuitously]).
be intended if an agent be allowed knowingly to permit it when performing some good act. For those like myself who accept PDE, this places an insurmountable barrier to denying that evil can be intended in any sense at all.

So in what sense *can* he intend it? The evil would, I claim, have to be a *by-product*, an *indirect effect*, of what the agent does actually intend directly, which is some good. In this wider sense, we can hold evil to be an objective for an agent insofar as it is within the broader scope of their practical reasoning – not directly intended by the agent. As such it is not to be held as paradigmatically intentional evil, an object of mere foresight, let alone accidental causation. So consider the tyrant who intends to exterminate some oppressed but innocent minority as a means to establishing the primacy of his favoured group (a partial good) or to securing the stability of the state (a complete good). The suffering of the minority is indeed an evil, and causing them to suffer is an intentional evil. But the tyrant, on the view I am defending, does not intend this evil *as evil*. He is not oriented to *this*, but to the securing of his end, which is at least a partial good albeit not *the* good to do. The evil means has some aspect of good to it, even if only the minimal one of being the means that fulfils a partially good end. It is for this reason that the tyrant finds the means desirable, in older terminology appetible – *good* for him to do, and this
even if he is grossly mistaken in thinking that it is the good for him to do. What if he is a plain sadist? Then he is oriented to the pleasure he gets from inflicting suffering, and pleasure is, again, a good – albeit neither fundamental nor the good to be achieved in this situation. What if he is a cold-blooded psychopath?

Then if we are to count the tyrant as minimally rational – not insane, not deluded – he must be oriented to some other good, real or apparent: he might have the sober yet wildly false belief that inflicting this suffering will achieve some greater objective or other, that the Devil commanded him to do it, that the agent himself is superior to all other humans and holds their fate in his hands, and so on.

I contend that all examples of what we would plausibly call intentional evildoing or intentional immorality can be accounted for in these terms. A departure from the good can never be an orientation in the sense of involving the direct intending of evil. It must always involve an orientation toward some good or other, one chosen by the agent in preference to the good that is lost by doing evil. In so choosing, the agent does intentionally perform evil, but does not intend it as evil. This is what Aquinas and the Aristotelian tradition mean by evil as having an accidental, not a per se cause. They do not mean that evil is only ever caused accidentally as opposed to intentionally. What they mean is that evil as such is never the proper object of intentional agency, which of its essence aims only at the good, whether as end or means. The core idea is, I
believe, summed up in one of the most profound passages in all of the writings of St Thomas Aquinas on morality:

Since the nature of the good is the nature of the appetible, …, and since evil is opposed to good, it is impossible that any evil as such be desired by an appetite, either natural, animal, or intellectual, which is the will. Nevertheless, some evil may be desired (appetitur) accidentally insofar as it is consequent upon some good, and this appears in each appetite. … The evil, however, which is joined to some good is the privation of another good. Never, therefore, would evil be desired, not even accidentally, except as long as the good it is joined to is more desired that the good which is taken away by the evil.\textsuperscript{20}

4. Conclusion

I have not attempted a full-blown defence of the Guise of the Good, which would necessarily take us into further realms of moral psychology, the theory of agency, and the interpretation of putative counterexamples. Rather, I have tried to show how the thesis can be seen to emerge from a more general metaphysical theory of good and evil. I have had to make many assumptions, to be sure; if these can be made good – a task for elsewhere – we will end up with a picture that is both plausible and satisfying for its internal coherence as well as

\textsuperscript{20} DeCoursey (1948): 112; \textit{Summa Theologica} I q.19 a.9, resp.
its extrinsic coherence with our deepest ethical convictions. The Guise of the

Good is but one stone in the overall edifice. It is one that Anscombe,

approaching it via her theory of intentional agency, fully grasped and admirably

– albeit laconically – defended.21

Department of Philosophy

The University of Reading

Reading RG6 6AA

d.s.oderberg@reading.ac.uk

REFERENCES


University Press).


____________ (1956) *Summa Contra Gentiles (On the Truth of the Catholic

Faith)*, Book 3: Providence, Part 1 (trans. V.J. Bourke) (Garden City,


161-201.

21 A version of this paper was delivered at SUNY Buffalo in 2013. I am grateful for the comments received there.

Cronin, M. (1930) *The Science of Ethics*, vol. I (Dublin: M.H. Gill and Son, Ltd.).


