

## Society for Applied Philosophy 2022 Annual Lecture Symposium

### Three Dogmas of Normativity

RUTH CHANG 

**ABSTRACT** *In this article, I identify and critically examine 3 dogmas of normativity that support a commonly accepted ‘Passivist View’ of rational agency. I raise some questions about these dogmas, suggest what we should believe in their place, and moot an alternative ‘Activist View’ of what it is to be a rational agent that grows out of rejection of the 3 dogmas. Underwriting the dogmas and the Passivist View, I suggest, is a deeply held but mistaken assumption that the normative domain is fundamentally akin to the nonnormative domain. Once we allow that the normative may be fundamentally unlike the nonnormative in certain key ways, a shift in our thinking about what it is to be rational becomes possible. I end by considering some implications of this paradigm shift in rationality from the passive to the active for various applied matters.*

In this article, I identify and critically examine 3 dogmas of normativity. These dogmas are so deeply entrenched in mainstream thinking about normativity that highlighting them as such – let alone raising some questions about them – may strike some as bizarre or misguided. But they deserve scrutiny for two reasons. First, unlike some well-established beliefs that win their place in the philosophical canon after considered argument and debate, these dogmas are little more than deeply held but unreflective assumptions that, while widely deployed in philosophical discussion, are not so obviously correct as to warrant acceptance without examination. Second, and more importantly, these dogmas support a widely held, similarly deep-rooted view of rationality that underwrites much of our current thinking about important matters not only in philosophy, but also in other domains of inquiry. By highlighting these dogmas, raising some questions about them, and mooting some alternatives, a different way of thinking about rationality comes into view. As I will suggest, investigation of these 3 dogmas leads the way to a fundamental shift in our understanding of what it is to be a rational agent, one that puts active, creative human agency at the center of rational thought and action. And with this alternative way of thinking about rationality in place, different ways of thinking about old issues, ranging from affirmative action to business ethics to technological design, come into view.

The 3 dogmas are as follows.

First, the claim that normativity is ‘absolute’, that is, first-order normative truths need not be relativized to a substantive normative consideration, like justice or beauty or empirical warrantability, but can be unqualifiedly true. According to this dogma, it could be true that human life is more valuable than cockroach life, simpliciter; that you ought to believe that the sun is 109 times the diameter of the earth, period; and that you have most reason not to shoplift at the corner bodega, full stop, without qualifying any of these claims by reference to some substantive normative consideration, such as the value of conscious

life, warrantability-on-the-basis-of-evidence-E, or the prudential good of staying out of jail. We might call this the ‘normativity is absolute’ dogma.

Second, the claim that two items can be normatively related in one of only three basic ways – one of the normative equivalents of ‘greater than’, ‘lesser than’, or ‘equal to’ – and that if none of these three relations holds, then the items are normatively incomparable. Thus one item must be better, worse, or as good as another; one belief must be more supported, less supported, or equally supported by the evidence; one reason must be weightier, less weighty, or as weighty as another, and so on, else they cannot be normatively compared. We might call this the ‘normativity is trichotomous’ dogma.

Third, the assumption that the reasons we have to think, feel, and do, and the value or merits of propositions for belief or options for action, are given to us, that is, as exclusively a matter for recognition and discovery and not also as a matter for our creation. The reason to stay perfectly still when confronting a snake in the grass, to take an example from Parfit, is something to be discovered, not invented by us. Our reasons and values are given to us, not made by us. We might call this the ‘normativity is given’ dogma.

Thus, according to the 3 dogmas, normativity is absolute, trichotomous, and given. My aims are threefold. First, to describe each dogma, diagnose why it has been assumed, raise some difficulties for it, and suggest what we might believe in its stead. I am under no illusion that the brief remarks I make here will convince anyone that these dogmas are mistaken or that the alternatives I propose are correct (although I have attempted more detailed arguments in other work<sup>1</sup>). The aim is the more modest one of shining a light on some hidden and unreflective assumptions that lurk in the background while playing a significant role in our thinking about normativity that, I believe, deserve further examination.

Examination of each dogma naturally leads to a different way of thinking about rationality and what it is to be a rational agent. As it is widely and commonly understood, rationality is, I will suggest, a fundamentally passive business; our job as rational agents is to discover normative and nonnormative truths that are given to us, not to create such truths. This ‘passivist’ model of rationality works well for inquiry in science and other nonnormative domains, but its success there has misled us, I believe, to assume that rationality writ large is to be understood in these passive terms. Rejecting the 3 dogmas and accepting the claims proposed in their place opens the way to a shift in understanding rationality as centrally an active capacity, a capacity to quite literally create reasons for oneself. My second aim is to sketch this alternative, ‘activist’ paradigm of what it is to be a rational agent.

Finally, with this alternative view of rationality in hand, I end by listing some implications it might have for a range of topics: decision theory and economic approaches to choice, governmental regulation, war and conflict, religious and scientific belief, egalitarianism, affirmative action, law and legal reasoning, business ethics, bioethics, and AI development and design. I suggest that by accepting the Activist View of rationality, we can potentially make some progress – or at least put a salutary spin – on some difficult and long-standing debates in these and other fields of inquiry.

Some ground clearing. The normativity of interest is what Derek Parfit (2011) has called ‘the normativity of normative reasons’, and not the normativity of mere norms or rules that are not, at least at first blush, normative in Parfit’s strong sense. Put another way, the 3 dogmas concern what is sometimes called ‘substantive’ rationality, the normativity of doing, feeling, and thinking in the light of normative reasons, as opposed

to 'structural' rationality, the rules or standards governing movements of mind or movements of mind to action. Although the dogmas hold fast in both the practical and theoretical spheres, I am going to focus here on practical normativity, the normativity of noncognitive attitudes, intentions, and actions, since theoretical normativity, the normativity of beliefs, raises special problems and complications, but I hope the reader will be able to see analogues (though admittedly more controversial) in the theoretical realm. It will also be convenient to frame the arguments in terms of concepts that are generally accepted in current debates since nothing will turn on exactly which concepts are employed. Sometimes I will talk in terms of values or goodness and sometimes in terms of reasons because, I believe, normativity can be understood in either of these terms. But the points made can be modified to fit any preferred understanding of the normative, whether it be in terms of oughts, choiceworthiness, obligations, permissions, justification, rationality, or some plurality of notions.

## 1. First Dogma: Normativity Is Absolute

The first dogma is a bit tricky to state because it can be assumed in a number of guises. We might start with some relatively uncontroversial normative claims:

1. Human life is more valuable than cockroach life.
2. *Ceteris paribus*, you ought to save five lives over one.
3. It would have been better if the COVID-19 pandemic had not happened.
4. Justice as a value or ideal is more important than gustatory pleasure.
5. Rationality requires you to always do what you have most or sufficient reason to do.
6. A taxation policy that improves the lot of the worst off at the expense of the best off, so long as the best off are extremely well off, is preferable to one that takes from the worst off in order to benefit the best off.
7. A life of contemplation is more choiceworthy than a life of physical pleasure.
8. It is morally wrong to cause suffering for its own sake.
9. Morality requires you to lead your life so as to maximize the best consequences, or to act in accordance with your duty, or to be virtuous, etc.

There are various ways to understand these claims, but my focus will be on three common ways according to which the claims are 'absolute', i.e. ways that presuppose, if only implicitly, the first dogma of interest. I will suggest a different way of understanding such claims in due course.

Suppose you think that human life is more valuable than cockroach life, *simpliciter*. It's not that human life can be more valuable than cockroach life only with respect to, say, the quality of conscious life or the purpose of populating a planet with maximal value, although these relativized claims may also be true. Instead, human life can be more valuable than cockroach life, *simpliciter*, full stop, period, absolutely.

Now there are three ways you might understand such *simpliciter* claims. First, you might believe that human life is more valuable than cockroach life, *simpliciter*, because you believe simply and straightforwardly that normative claims can be absolute. Just as there are absolute claims in science, unrelativized to any substantive physical

considerations, so too, there are absolute claims in normativity, unrelativized to any substantive normative considerations. The sun is 94.4999 million miles from Earth, period. Similarly, human life is more important than cockroach life, period. Absolute claims are just some of the kinds of claims within the normative domain. End of story.

You might instead believe that human life is more valuable than cockroach life, simpliciter, because like (early) G.E. Moore, you believe that there is some special superconsideration, such as being simply good (or being simply valuable, or simply what one ought to do, or simply choiceworthy, etc.) that all good things have in common and in terms of which simpliciter claims can proceed.<sup>2</sup> Human life is more valuable than cockroach life, simpliciter, because with respect to, say, the superconsideration of being just plain valuable, human life has more of it or instantiates it in a more significant way.<sup>3</sup>

A third way to understand simpliciter claims is as proceeding with respect to all the substantive normative considerations there are. Sometimes absoluteness is equated with holding ‘all things considered’, understood to refer to all the normative elements there are. Human life is more valuable than cockroach life, simpliciter, because, taking into account all the values there are, human life scores better than cockroach life. *Ceteris paribus*, you ought to save five rather than one, all things considered, because considering all the reasons there are for doing anything, saving five is, *ceteris paribus*, better supported than saving one.

Thus we might say that absolute normative claims can be understood either straightforwardly, or as relativized in an unusual way, namely, either to a superconsideration or to all the substantive normative considerations there are. So the first dogma can be stated as follows:

*First Dogma: Normativity Is Absolute*

First-order normative truths about what one should do or about the goodness of things (and so on) – simpliciter – can proceed straightforwardly, or as relativized either to a superconsideration or to all the substantive considerations there are.<sup>4</sup>

In short, those who accept the first dogma think there are first-order normative claims, simpliciter, that hold in at least one of the three ways mentioned above. (Henceforth, by ‘normative claims’, I will mean first-order claims). Those who reject the first dogma believe that all such normative claims must be relativized to substantive normative considerations that are neither superconsiderations nor all the substantive normative considerations there are – what we might call ‘ordinary’ substantive normative considerations. Human life cannot be more valuable than cockroach life absolutely but must be more valuable relative to an ordinary substantive normative consideration like ‘quality of conscious life’, ‘populating a planet that maximizes value’, and ‘being deserving of legal protection’ or some proper subset of ordinary considerations. Similarly, it cannot be true that, *ceteris paribus*, you ought to save five lives over one absolutely, but only that you ought to save five over one with respect to morality. There are no absolute normative claims about whether it is more choiceworthy to tax the rich over the poor, whether it would have been better if the pandemic had never happened, or indeed about whether one value, such as justice, is more important than another, such as gustatory pleasure. All of these claims, according to those who would deny the first dogma, must be relativized to ordinary substantive considerations such as ‘justice’, ‘the social good’, or ‘what makes for a flourishing human life’.

Why is the first dogma widely accepted? I suspect that many who assume it have never reflected on why they do so. I suggest three general diagnoses of why it is so easy to assume that cut across the different ways a claim might be thought to be absolute.<sup>5</sup>

First, without absolute normative truths, it is hard to see how we can 'keep track' of normativity over time and combine or aggregate the value or reason-givingness of a normativity bearer. Each of the actions you have performed today have some value or dis-value or are supported or dis-supported by reasons. How can there be a normative truth about whether your day has been good or lousy overall or whether you were consistently and stunningly irrational this morning unless there are absolute normative truths that report some kind of aggregation or combinatorial function of the individual values or rationality of each of your actions, absolutely and without qualification? On your deathbed, you might conclude that your life has been a good one, where that thought seems to be about the goodness of your life, simpliciter, unrelativized to any ordinary substantive normative consideration. More broadly, we might say that human history has been one mostly marked by rationality, enlightenment, and progress, but that it now seems that we are headed for a dark period of irrationality, fear, and decline. How are we to understand such claims if not as 'tracking' normativity in some absolute, unqualified sense?

Second, if the first dogma is mistaken, then it follows that every normative claim must be relativized to some ordinary substantive normative consideration. But what are these considerations in terms of which seemingly absolute normative claims proceed? Taking the claims at the outset of this section, to which ordinary substantive normative considerations is each relativized? Since it is so difficult to name substantive considerations to which such claims are relativized, it is natural to suppose that they are not relativized to any such consideration. They are just as they seem: absolute.<sup>6</sup>

The deepest diagnosis of why the first dogma is so easily assumed lies, I believe, in another long-standing assumption that underwrites all 3 dogmas: that the normative domain is of a piece with the scientific domain. In particular, many thinkers treat normativity as a justificatory force akin to a physical force like gravity or electromagnetism: normative elements are justificatory pushes and pulls that contribute to an overall absolute, unqualified justification for thinking, feeling, or doing something. Just as the earth, moon, and stars are physical elements that contribute to absolute truths about the gravitational force of an object, so too the ways in which something is valuable or supported by reasons come together to comprise its overall, absolute goodness or value, or the overall, absolute normative or justificatory force for choosing it. Sometimes the analogy is extended to necessities. Derek Parfit and Thomas Scanlon argue that just as there are physical, mathematical, and logical necessities, all of which are absolute, so too there are absolute normative necessities.  $2 + 2 = 4$  states an absolute truth and so does 'It is wrong to cause suffering for its own sake'.<sup>7</sup>

With this underlying, 'scientific' view of normativity in place, it is natural to suppose that, just as there are absolute truths about how things are, there are absolute truths about how things ought to be: about what we are justified in doing and believing. Indeed, it is this unqualified, absolute justificatory force in terms of which we can in principle 'keep track' of the normativity in the world and avoid having to locate obscure substantive values to which those claims might otherwise be relativized.

Why think the first dogma is mistaken? Here are three considerations for doubt.

First, if normativity were absolute, then normativity in its comparative form would also be absolute. That the comparative form is problematic is easiest to see in the case of value.

If there is goodness, simpliciter, there is its comparative, betterness, simpliciter. But betterness, simpliciter, does not appear to have unambiguous content. If I say ‘You are better than me’, I have not expressed a complete thought.<sup>8</sup> We can only understand what I say by imputing some ordinary substantive respects in terms of which you are better: you might be better than me as a moral agent, as a speller, in achieving work-life balance, and so on, and the relativization is required in order for there to be a complete thought and, correspondingly, something that is truth-evaluable. The fact that comparative forms of goodness, choiceworthiness, ought-to-be-doneness, and so on do not permit of absolute claims gives us some reason to think that it is a mistake to think that their noncomparative forms. Indeed, Judith Thomson has argued that the noncomparative goodness, simpliciter, is mistaken and that there is only goodness-in-a-way.<sup>9</sup> Even if Thomson’s arguments are not accepted, we would need at the very least some explanation as to why ‘You are good, simpliciter’ supposedly expresses a complete thought while ‘You are better than me, simpliciter’ does not.

Suppose that when I say ‘You are better than me’, I mean that with respect to all the values there are – sometimes captured by the locution ‘all things considered’ – you are better than me.<sup>10</sup> In this case, we would need an account of how all the values there are hang together such that it can be true that you are better than me with respect to all of them taken together. To my knowledge, no such account has been attempted. One suggestion might be to interpret being better, ‘all things considered’, in terms of dominance or Pareto superiority; you are better than me with respect to all the values there are because you are at least as good as me with respect to all of them and better than me with respect to some. But this suggestion is problematic in two ways. It narrows absolute claims to cases of dominance where those who accept the dogma purport to make absolute claims in a much wider range of cases. And in any case, dominance – if there are organic unities or parity – is a poor formula for betterness.<sup>11</sup>

Of course, ‘all things considered’ is not usually taken to mean literally all things considered but only some subset of ordinary substantive considerations that are relevant to the case. So understood, normative claims employing ‘all things considered’ would not assume the first dogma. Later, I will suggest a unified way of understanding all apparently absolute claims, including ‘all things considered’ claims.

It might be worth pointing out that if the above arguments are correct, a lacuna in modern-day pluralist consequentialism comes into view. Bentham posited utility as playing the role of goodness, simpliciter, which meant that he had the requisite theory of value required to assess consequentialist claims about outcomes as true or false. Modern-day consequentialists, such as Philip Pettit, have improved on Bentham in many ways, in particular by being pluralists about the values relevant to assessing the goodness of outcomes, but they insist that there is such a thing as the most valuable option, simpliciter.<sup>12</sup> But this improvement brings with it an axiological burden; before any normative claim about the goodness of an outcome can be evaluated as true or false, a theory of all the values there are and how they normatively relate to one another is required – that is, if consequentialists both embrace the first dogma and purport to provide a general theory of right action. But no such axiology has been offered. If, instead, modern-day consequentialists reject the first dogma and allow that every normative claim about the goodness of outcomes is relativized to some specific set of ordinary values, we need instead an account of which values are relevant to any such claim, what determines how those values normatively

relate, and how claims relativized to different values allow for aggregation across those different values.<sup>13</sup>

Another reason to doubt the dogma concerns how the diversity of values, reasons, and normative elements in general can be properly accounted for in absolute claims. As Aristotle pointed out long ago, good things are often good in different ways, and it is a mystery how it can make sense to compare diverse goods in absolute terms. Which is better, simpliciter, the utility of a doorstep or the pleasure of a bowl of Rocky Road ice cream, the pleasure of reading Kant or the abstract beauty of the number 4, the justice of repaying a debt or the goodness in having a pet dog? If there were absolute truths about goodness, it should at the very least make sense to compare any good thing with any other good thing. This is not a point about ranking all good things but the intelligibility of asking which of two good things is better, simpliciter. Sometimes it makes no sense to ask how two good things compare.<sup>14</sup> The same goes for reasons. Could the reason provided by the fact that she is your cousin be put together with the reason provided by the fact that she is a good philosopher in thinking about whom to hire in the philosophy department? Once we allow that only certain substantive considerations are relevant to determining a normative truth, that truth is no longer absolute. Absolutists might respond by appealing to the restriction that absolute comparative normative truths only compare like with like.<sup>15</sup> But how can we determine whether one item is like another without appeal to substantive normative considerations? Apples are unlike oranges with respect to being a deliciously crunchy snack, but they are alike with respect to being a deliciously healthful snack.

A third reason to doubt the dogma is that by rejecting it, various putative puzzles disappear and the possibility of unifying explanations of a fundamental phenomenon opens up. Some philosophers, for example, have alarmingly suggested that 'better than, all things considered' is nontransitive, but their arguments appear to be about a relation that, if we reject the first dogma, does not exist.<sup>16</sup> If 'better than, all things considered' does not stand for 'better than, with respect to all the substantive normative considerations there are', it most plausibly stands for 'better than, with respect to ordinary substantive considerations *u*, *v*, *w*'. An intuitive appeal to judgments with respect to 'betterness, all things considered' cannot establish the nontransitivity of 'betterness, all things considered' without ruling out equivocation among the judgments that supposedly lead to a nontransitive cycle. Perhaps *A* is better than *B*, all things considered, and *B* is better than *C*, all things considered, and *C* is better than *A*, all things considered, but there is no nontransitivity in being better all things considered if the things considered in each judgment are not the same. Nor can the nontransitivity of 'betterness with respect to all the substantive normative considerations there are' be established without some account of how all the substantive normative considerations there are hang together such that judgments that lead to intransitive cycles are correct. Rejecting the first dogma helps dissolve what otherwise might be puzzles about the transitivity of what might be thought to be definitionally transitive relations like 'better than'.<sup>17</sup>

Rejecting the first dogma also helps solve a puzzle about exclusionary reasons. Take the case of law. Joseph Raz (1990) famously argued that law provides exclusionary reasons, that is, second-order reasons not to consider first-order reasons in determining what to do.<sup>18</sup> But exclusionary reasons run into difficulties in certain cases. Suppose you are driving on a desolate stretch of desert road. You come upon a red light. The legal rule, 'You ought to stop at red lights', provides you with a second-order exclusionary reason not to take the fact that there is no one for miles around as a reason not to stop. In the



absence of any further reasons, you now have most reason not to stop at the red light. This is an odd result. If we understand legal norms, not as absolute but as relativized to ordinary substantive considerations, then the norm, 'You ought to stop at red lights' can be understood as the norm 'You ought to stop at red lights with respect to achieving the value of solving coordination problems' or 'You ought to stop at red lights with respect to the value of achieving better compliance with the first-order reasons to stop at red lights across typical circumstances' or so on. Exclusionary reasons as relativized norms – in the law and beyond – can avoid problematic cases because once relativized, they do not offer reasons to disregard first-order reasons in the problematic cases.

While there are other matters in normativity that can be illuminated by denying that normativity is absolute,<sup>19</sup> I will end by noting that denying the dogma makes space for a simple, unifying explanation of a fundamental feature of the normative theorizing: a consideration can determine what is good or what you ought to do in a specific way, in one set of circumstances, but determine what is good or what you ought to do in a completely different way, in a different – even slightly different – set of circumstances. This phenomenon has been variously flagged in the service of axiological organic unities, reasons holism, Jonathan Dancy's particularism,<sup>20</sup> Shelly Kagan's 'additive fallacy',<sup>21</sup> and Frances Kamm's 'principle of contextual interaction',<sup>22</sup> but it is a fundamental feature of any plausible account of normativity. One consideration may count in favor of an action, A, when what matters is V, but count against or have different normative significance vis-à-vis A when what matters is W. If normative truths are relativized to ordinary substantive normative considerations, what looks like the nonadditivity of values or variation in normative significance depending on context can be explained in terms of the relativization of normative truths to different ordinary substantive normative considerations.

If the first dogma is false, what should we believe in its place? In particular, what should we make of seemingly absolute normative claims?

I suggest that 'goodness', 'choiceworthiness', 'what one has most reason to do', simpliciter, 'all-things-considered-better-than', and the like are placeholder or category concepts that hold the place of and collect substantive normative considerations but are not themselves substantive normative considerations.<sup>23</sup> Consider, for instance, 'value'. There is no substantive consideration of being valuable; there is only the concept of being valuable that collects substantive evaluative considerations like 'beauty', 'justice', and 'utility as a doorstep' as falling under it. Or, equivalently, being valuable is not itself a substantive consideration but a formal, schematic consideration that requires filling in if it is to deliver a truth-evaluable normative truth. When we say that something is 'valuable, simpliciter' then, we are saying that it bears some substantive value, to be inferred from the context or to be determined by further inquiry. 'Simpliciter', 'full stop', 'all things considered', and the like are placeholder concepts or signify formal schematic considerations that hold the place of ordinary substantive considerations to which all normative claims are relativized.

Thus there is no such thing as being good or valuable or better than or choiceworthy or rational or obligatory or what we ought to do, simpliciter. Human life is never good absolutely; it can be good only with respect to certain substantive values.<sup>24</sup> Nor can human life be more valuable than cockroach life, simpliciter; it can only be more valuable relative to a substantive consideration. Indeed, cockroach life may be better with respect to the value of surviving nuclear holocaust. Similarly, you never have most reason to do one thing rather than another, full stop. Instead, you can have most reason to do A over B with



respect to some ordinary substantive normative consideration. You are never obligated to do something, *simpliciter*. Your obligations arise relative to some ordinary substantive normative consideration. And you have no absolute rights. Your rights are always relativized to some substantive normative consideration, such as justice. If there are no absolute rankings of goods, values, reasons, oughts, and rational courses of action, then ‘keeping track’ of the normativity in our lives will be a nuanced affair, with the most general normative claims relativized to the most comprehensive of substantive considerations but no totally comprehensive claims that aggregate the value or normativity in our lives literally overall.

I have elsewhere called the ordinary substantive normative considerations with respect to which normative claims proceed ‘covering considerations’ since they must ‘cover’, i.e. apply to, the subject matter of the normative claim. While ‘justice’ covers governmental policies, even bad ones, it does not cover numbers or doorstops.<sup>25</sup> Covering considerations give what matters in a normative truth. I suggest that we reject the first dogma and accept instead the claim that all normative truths are relativized to covering considerations.<sup>26</sup>

None of this is to say that we cannot sensibly talk of one thing being better than another, *simpliciter*, or of there being something which you are all-things-considered rationally required or ought to do, or how much value we’ve had or how rational we have been in our lives overall. Philosophers including Moore and Wittgenstein have noted that we regularly use normative language to make *simpliciter* claims.<sup>27</sup> But we must be clear that such talk is elliptical for talk of relativized, rather than absolute, claims. We can say that you ought to save five rather than one, *ceteris paribus*, but we need to roll up our sleeves and investigate the substantive normative considerations in terms of which such a claim is true. Talk of what one has most reason to do, full stop, is shorthand for talk of what one has most reason to do relative to a value, purpose, policy, or other ordinary substantive normative consideration.

When we ask what all normative or good or choiceworthy or rational things have in common, the answer is not that there is some absolute consideration of being normative, good, choiceworthy, or rational, *simpliciter*, that they each instantiate or to which they contribute, but rather that each falls under a *simpliciter* concept or is a candidate specification of a formal, schematic *simpliciter* consideration. If this is right, then Moore’s insight needs adjusting: there is no property of being good, *simpliciter*, that runs like a thread through all good things. What all good things have in common is that they are picked out by the concept or the schematic consideration of being good, *simpliciter*.

Note that this ‘placeholder’ way of understanding normativity avoids the three objections to the first dogma while making sense of our talk of goodness, what we ought to do, choiceworthiness, obligations, preferability, and so on, *simpliciter*. We make *prima facie* absolute claims as shorthand for claims relativized to ordinary substantive considerations that must be imputed in order for us to express complete, truth-evaluable thoughts. Putting together disparate normative elements is no longer a problem since the question of which values and reasons are put together by substantive considerations is a matter of substantive debate without any presumption that any two values or reasons can be put together. And there is no need to puzzle over what looks to be the nontransitivity of ‘better than, all things considered’ since there is no such relation. Nor need we puzzle over the fact that a reason can have different normative significance when what matters is given

by different covering considerations. Holism about reasons can be recast as the relativization of reasons to various covering considerations. And by understanding exclusionary reasons as relativized reasons, we have a more nuanced way of determining when they block underlying first-order reasons and when they do not.

The placeholder view understands normativity as a collection of truths relativized to ordinary substantive normative considerations – i.e. covering considerations. When we ask, ‘Is this the best/right/most rational way to live?’, we are asking how we should live relative to some ordinary substantive consideration, such as morality, harmony with the ecosystem, or being a basis for a riveting TV documentary. Every normative claim, however abstract and grand, must be relativized to a substantive covering consideration.

If this is right, then normative theorizing might be cast in a different light. If there are no absolute truths about what is good, best, or choiceworthy, or about what we ought to do, how we have most reason to live our lives, what is most rational, and so on, then the critical questions of normative theorizing become: (i) what determines which ordinary substantive considerations, i.e. covering considerations, matter in different circumstances, and (ii) what normative relations hold among the various substantive considerations to which normative truths are relativized in those circumstances?<sup>28</sup> Normative comparisons, on this view, lie at the heart of understanding the normative domain, and it is to those that we should turn our attention when answering the question, how should we live?

## 2. Second Dogma: Normativity Is Trichotomous

Pick a value, any value. If we try to compare two things with respect to it, it seems that only one of three possible ‘basic’ value relations could hold between them: one must be better than the other, worse than it, or the two must be equally valuable.<sup>29</sup> If none of those relations holds, then the items are incomparable with respect to that value.

The same is thought to hold for basic normative relations generally. When we compare two items with respect to any normative consideration, we assume that if they can be compared with respect to that consideration, they must stand in one of the normative analogues of ‘greater than’, ‘lesser than’, or ‘equal to’. A reason for believing that *p*, for instance, must be weightier than, less weighty than, or just as weighty as a reason for belief that not-*p*; a principle for organizing human affairs must be more just, less just, or equally as just as another; an obligation must be more stringent, less stringent, or equally as stringent as a competing impossible obligation, and so on. For ease of discussion, we can stipulate that ‘better than’, ‘worse than’, and ‘equally good’ are ‘master relations’ that include all the other possible trichotomous relations of normativity. Hence we can state the second dogma as follows:

### *Second Dogma: Normativity Is Trichotomous*

If two items are comparable with respect to some normative consideration, one must be better than, worse than, or equally good with respect to that consideration.<sup>30</sup>

I have elsewhere called this dogma the ‘Trichotomy Thesis’, the thesis that normative considerations, like many nonnormative considerations such as length, weight, and

volume, are trichotomous in structure. This dogma is ubiquitous; it underwrites thinking across a wide range of normative domains in the humanities and social sciences, in epistemology, and in the philosophy of science.

Here is a diagnosis, admittedly speculative, of why it is believed. When we first came to populate the earth, our overwhelming concern was survival and the avoidance of suffering imposed by a mysterious and terrifying external world. We bumbled around in more or less this state for hundreds of thousands of years until around 4000 BC when the Sumerians discovered number, which they used to count items for trade. A mere 1000 years later, Egyptians used number to represent abstract quantities like length to build the great pyramids. Some 24 centuries ago, the Pythagoreans posited that all dimensions of physical reality could be represented by a ratio of integers. Today we use real numbers to build spaceships, manage pandemics, and connect with one another through social media. In short, the discovery of numbers has allowed us to tame a dangerous nonnormative reality and to manipulate it for our purposes.

While real numbers – with a suitable interpretation – can in principle be used to represent anything, they are by their nature trichotomous in structure. One real number must be greater, lesser, or equal to another, and something does not even count as a real number unless it stands in one of these relations to some other number. Since assuming a trichotomous structure in the nonnormative world – especially in scientific inquiry – has been so successful in allowing us to measure, make predictions about, and understand the nonnormative world, it is perfectly natural that we import the same assumption to our thinking about the normative world. Again, a ‘scientistic’ assumption about the nature of normativity can explain why we unreflectively assume the second dogma. Just as one length must be greater, lesser, or equal to another, so too one good must be better, worse, or equally as good as another, and one action must be better supported by reasons, less supported by reasons, or equally supported by reasons as another, and so on. This is not to say that we assume that values and normative elements are quantities; indeed, it is clear that both normative and nonnormative properties have qualitative features. Rather, the second dogma supposes that normative elements can be represented as quantities, that is by numbers, even if they are not themselves quantities.

In other work, I have proposed arguments for thinking that the second dogma should be rejected. Here, my aim is the more limited one of suggesting some considerations for thinking that we should not be so cavalier about assuming its truth, considerations that I hope make a sufficient case for concluding that this dogma deserves scrutiny in its own right.

We might start with a naïve, pretheoretical worry. From our current path-dependent, theory-laden perspective, it may seem odd to think that the second dogma could be false. After all, if two things can be compared with respect to some consideration, surely one must be better than the other, worse than it, equally good, or else they cannot be compared. But from a naïve, pretheoretical perspective, it seems even more odd – indeed, outrageously implausible – simply to assume that normative elements such as love, beauty, justice, tawdriness, kindness, duty, and rights in the normative realm have the same structure as nonnormative elements such as length, weight, volume, electrical current, time, temperature, and luminous intensity in the nonnormative realm. We would not naturally assume that the structure of, say, a bridge, is the same as the structure of a family since bridges and families are such different sorts of things. Why should we

be so confident that the structure of the normative is the same as the structure of the nonnormative?

Many thinkers, especially in the social sciences, regard the second dogma as a conceptual truth, and indeed define 'comparability' in terms of it. So, for example, many economists define the rankability of items in terms of being to be preferred (better than), to be dispreferred (worse than), or to be indifferent between (equally good). If none of the trichotomy of relations holds, then we are left with a partial ordering, where items not so ordered cannot be compared.

But we can show through a thought experiment that our ordinary notion of comparability does not presuppose the second dogma, and thus that it is a mistake to think it is a conceptual truth. Imagine a community of dichotomists who think there are only two distinct ways items can be evaluatively compared. If two things can be compared, one must be better or worse than the other, otherwise they are incomparable; dichotomists do not recognize that what trichotomists call 'equally good' a distinct relation from being incomparable. Now suppose that a trichotomist meets a dichotomist at a conference. The trichotomist says to the dichotomist, 'I really enjoyed your talk. It was equally as interesting as my own'. The dichotomist will be flummoxed. What is this being 'equally as interesting as'? Is he being insulted or praised? By his lights, if one talk is not better or worse than another, they are incomparable. The relation of being equally good is not one that he recognizes.

We might think that the dichotomist cannot understand what the trichotomist says because each operates with a different stipulation of what it is to be comparable. They are just two ships passing in the night, operating with different concepts of comparability. But – crucially – we can also understand the case as involving a mistake: the dichotomist makes a substantive mistake about the range of basic values relations that exhaust the conceptual space of comparability between two items. He thinks there are just two, but he has overlooked a possible third basic relation, 'equally as good as'. We might try to convince him of the error of his ways by providing him with substantive arguments that, for example, distinguish between two kinds of case he calls 'incomparable', viz., those of an item and its twin, say, and some other cases. We employ arguments to show him that his category of incomparability fails to distinguish different kinds of case worthy of distinction, and that one kind of case is closer to a shared notion of comparability than to its denial.

We think that the dichotomist has made a substantive error because we make the substantive assumption that there are three – and not two – such relations. The second dogma is not built into our ordinary notion of being comparable; if it were, we would simply regard the dichotomist as having a stipulative understanding of 'comparability' different from our own. Thus, insofar as economists and decision theorists aim to provide us with models that reflect our ordinary notion of comparability, they should not define their functions in terms of the usual trichotomy of relations but construct models that leave open the question of what relations exhaust the conceptual space of ordering relations among items.<sup>31</sup> At the very least, we are owed an argument in defense of trichotomy.

It might be thought that defense of the dogma lies in the fact that it serves us well. We do just fine researching normativity and leading our lives on the assumption of the dogma; the work in economics that assumes the dogma has led us out of recessions and bolstered the world economy; psychologists and neuroscientists who assume the dogma uncover interesting truths about mechanisms of decision-making; and philosophers who assume it are not prevented from delivering insights into normativity, practical reason, and agency.

There is no need to fix something that is not broken. But notice that the same appearance would hold in a world of dichotomists. In a world in which everyone believed that in the typical course of things, one action was more choiceworthy than another, one belief more warranted than another, one good better than another, and in exceptional cases, items for which this was not true are incomparable, nothing would seem amiss. People would make choices and lead their lives in unexceptional ways, and there would be a dichotomy-based academic industrial complex churning out theories, models, and insights with dichotomy as a dogma. Thus the mere appearance that the second dogma serves us well arguably gives us little reason to think that the dogma is correct. In any case, it is disputable whether the dogma does serve us well. As I have argued elsewhere, problems such as explaining the phenomenology of choice and evaluation,<sup>32</sup> special partial reasons in committed relationships,<sup>33</sup> hard choices,<sup>34</sup> the distinctiveness of normative powers,<sup>35</sup> how to avoid various puzzles in population ethics,<sup>36</sup> why it is sometimes okay to be value pumped,<sup>37</sup> and how we can be justified in moving through life through one set of choice situations rather than another,<sup>38</sup> cannot happily be explained in ways that are consistent with the second dogma. If these arguments are correct, we have grounds for investigating whether the unreflective assumption that normativity is trichotomous in structure is correct.<sup>39</sup>

Finally we can offer a story as to how the second dogma could be false. Note that normative elements typically have both quantitative and qualitative dimensions. A normative reason, for instance, might have weight or strength, understood quantitatively, but also a quality or significance. Or consider the value of being pleasurable. Two pleasures – say the sharp, thrilling pleasure of bungee jumping and the languorous, luxuriating pleasure of taking a whirlpool bath – each involve both quantitative and qualitative aspects of pleasure that together determine the overall pleasurable-ness of the respective experiences. Assuming that they can be compared, must one pleasure be greater, lesser, or equal to the other?

If the pleasures can be trichotomously related, there must be some numerical representation of them – if they can be ordered, they can be represented on a list by the numbers 1, 2, and so on, or by the same number if they are equally pleasurable. Let us assume for the sake of argument that the contribution of the quantitative aspects of each pleasure can always be represented numerically because quantitative aspects already have built into them a coordinating function so that the number representing the contribution of the quantity of bungee-jumping pleasure ‘coordinates’ with the number representing the contribution of the quantity of Jacuzzi pleasure. So we grant (though controversial) that we can numerically represent the contributions of the quantitative aspects of both pleasures.

Can the contribution of the qualitative aspects similarly be represented numerically? There are two possibilities. Suppose that the quality of the pleasure of each experience is the same. In this case, we can again grant, for the sake of argument, that the same number can represent the contribution of these qualities of pleasure to the overall pleasurable-ness of each experience. Suppose that the quality of the pleasures is different, as they manifestly are in the present case. How could we numerically represent the contribution to the overall pleasurable-ness of each experience in this case? A natural route is first, to establish a rate of trade-off between the quality of bungee-jumping pleasure and the quality of Jacuzzi pleasure – perhaps the quality of pleasure from soaking in a 112F jetted tub is 1.246 times the quality of pleasure from leaping 134 meters off the Nevis Highwire in New Zealand – and second, to establish a rate of trade-off between quantity and quality of these pleasures – perhaps quantity and quality each contributes 50% to the overall

pleasurableness of an experience. With these two rates of trade-off in place, we can numerically represent the normative relation between bungee-jumping pleasure and Jacuzzi pleasure.

But as many philosophers have pointed out, the idea that there are such mathematically precise rates of trade-off for amorphous, complex values is highly implausible; and for any posited rate, counterexamples abound. It is in the nature of an organic unity, for example, that there is no rate of trade-off between contributory elements that determine the normativity of the whole; if the contribution of these contributory elements can be compared, then it will take something other than a rate of trade-off to establish their putatively trichotomous relations. Plausibly there are precise rates of trade-off in only the most *recherché* of cases, and even (some) economists recognize that the assumption that such rates of trade-off accurately reflect the relations among normative elements is at best an idealization. But there is a hidden assumption here. Although precise rates of trade-off are, many believe, a fiction, they are, many also suppose, an 'approximation' of the truth. All we need to do is to 'roughen' up our rates of trade-off to capture the genuine normative relations, and that is a detail we can worry about down the line. Indeed, this strategy is suggested by Parfit;<sup>40</sup> by thinking of the normative relations among normative elements as 'imprecisely cardinal', and presumably thereby allowing 'imprecise' rates of trade-off among normative elements, we can maintain trichotomy.<sup>41</sup> But why think that what undermines the possibility of precise rates of trade-off among normative elements is not also something that undermines the assumption that they can be only trichotomously ranked? Aristotle recognized the 'incommensurability', i.e. lack of (trichotomous) cardinal comparability, of values but thought that the imposition of cardinal rankings on items was a useful fiction to make the trading among incommensurable goods feasible. But he did not suppose that commensurating goods approximated the true normative relations among them.<sup>42</sup> We should, I suggest, follow Aristotle's caution in this regard.

If the second dogma is false, what should we believe instead? I have argued elsewhere that normativity is not trichotomous but tetrachotomous in structure. Two items can be normatively related if one is better than the other, worse than it, equally good, or on a par. There are four, not three, basic ways two items can normatively relate. Parity holds in the normative realm because qualitative differences in normativity give rise to a fourth basic way in which items can compare. Two qualitatively diverse items that are nevertheless in the same 'neighborhood' of value are neither better than one another nor equally good. They are on a par. The pleasure of bungee jumping is qualitatively very different from the pleasure of a hot bath. And yet both pleasures, suitably specified, may be in the same 'neighborhood' – the same nonhierarchically ordered category or cluster – of value overall. These are experiences that are on a par with respect to pleasurable-ness.

### 3. Third Dogma: Normativity Is Given

According to the third dogma, if we ask, what makes something normative? – that is, what makes something valuable or a reason to do something? – the answer will always be found in something beyond our direct volitional control. It has long been thought that the source or foundation of what we ought to do is God's will, the dictates of a supreme leader, facts



of normative reality, conventions created to achieve a set of goals, the aims or desires of an agent constrained by formal procedures, etc. To be clear, these are not views about which reasons or values we have but views about that in virtue of which something is a reason or is valuable. So when God commands that we honor our father and mother, this command is that in virtue of which we have a reason to undertake actions that constitute honoring them; when it is healthful to eat more vegetables, the healthfulness is that in virtue of which we have a reason to eat more vegetables; if you want to work from home rather than go into the office, your desire is that in virtue of which it would be a good thing for you to work from home. The source or grounds of normativity – that in virtue of which something is a reason or is valuable – have traditionally been understood as given to us and not created by us as a matter of our own wills.

*Third Dogma: Normativity Is Given*

Normativity is always given to us and never created by us; that is, the ground of a reason or value is never under our direct volitional control.

This dogma seems so obviously true that it is hard to imagine how it could be mistaken. One diagnosis of why we believe it calls again upon the assumption that the normative is akin to the nonnormative. In science, our reasons for believing that the Omicron BA.5 variants of the COVID-19 virus elude the immunity provided by current vaccinations are given to us, to be discovered by us, and certainly not created by us. Since it seems that our reasons to believe claims about the nonnormative domain are given to us, it is natural to suppose that the same holds for reasons to feel and do in the normative domain. Another diagnosis follows from what denial of the third dogma would seemingly entail. If we could, as a matter of will, create reasons by having direct volitional control over their grounds, then presumably we could create reasons to engage in all sorts of unsavory or immoral actions by willing their grounds. Denying the third dogma seems to wed us to a highly unattractive and implausible view of the reasons we could have. Moreover, it seems that reasons are not the kind of thing that we could create. After all, their purpose is to guide us in thinking, feeling, and doing. In order to do this, reasons must bind us. But how could reasons that we conjure up as a matter of will bind us?

Kant, sometimes interpreted as the progenitor of the idea that normativity has its source in a (nontheistic) agent's willing, tried to solve the extension and bindingness problems by insisting that a rational agent must will in a particular way: she must will so that the general principle describing her action conforms to the Categorical Imperative, which thereby makes her action morally permissible. Since a rational agent can only will in ways that are morally permissible, willing can ground reasons while avoiding the extension problem since the rational agent will perform only morally permissible actions. Neo-Kantians like Christine Korsgaard, who sometimes treat willing as a mental state or activity, have likened the Kantian requirement on willing to a requirement of structural rationality, the rationality of movements of mind.<sup>43</sup> If the Kantian requirement on willing is a rational requirement, the question arises: what normative reason do we have to conform to such a requirement? If there is an independent normative reason to conform to the Categorical Imperative, then the source of normativity does not lie in willing. If instead there is a further structural requirement to obey this structural requirement, then the question becomes why we should comply with that bundle of structural requirements, even if it constitutes our agency, rather than a different bundle of structural requirements, which

may constitute not agency but ‘schmagency’.<sup>44</sup> The Kantian line of thought leaves us in need of a reason to think that requirements to will one way rather than another are binding on us. So neither problem is solved.

Sartre and the existentialists also link the idea of willing and normativity. They take a rather different tack; instead of trying to solve the extension and bindingness problems, they avoid them altogether by denying, wholesale, the existence of normative reasons. There is nothing normative antecedent to choice that can normatively guide and bind us in choosing, and thus no problem of getting the right extension of normative reasons. And there is nothing to bind one, to choose one way rather than another. Hence the existential ‘nausea’ we face when we come to realize that there are no normative anchors given to us by the world and that we float free through life. We can talk of ‘reasons’ after choice only by overlaying such talk as if there were reasons and values in the world, but this is just a convenient way to talk about the radically free choices we make. In short, existentialists avoid the problem of how we can create reasons by being normative nihilists. But this is not to solve the problems raised by created reasons but to hold views – deeply problematic in other ways – according to which the problems do not arise.

These sketches of two existing views that might be thought to support the idea that we can create reasons through willing seem to leave the extension and bindingness problems in place. Thus it appears that the third dogma is on firm ground.

Are there reasons to doubt it? Elsewhere I have described how we could have reasons that are not given to us but are instead created by us through an activity of our wills.<sup>45</sup> I will have a bit more to say about this view shortly. But here I want to suggest a modest reason to doubt the dogma. Again, my aim is not to show that the dogma is false but only to nudge those not antecedently committed to it to be open-minded about the possibility.

If all reasons are given, then we should have some account of their grounds, of what makes them reasons. So if the third dogma is true, it is plausible that one of the leading metanormative views about the grounds of given reasons is correct. But the debate between source internalism and source externalism about reasons is in a state of gridlock, suggesting that neither view is true.<sup>46</sup> A thumbnail sketch of the main contours of the debate illustrates the point.

Start with source internalism, which maintains that something is a reason or value in virtue of a relation – usually ‘satisfying’, ‘fulfilling’, or ‘constituting’ – between the consideration that is the reason and one’s desires, goals, or aims. Parfit called such reasons ‘desire-based’.<sup>47</sup> Since we lack directional volitional control over our desires, all desire-based reasons are given to us, not created by us. (As Anscombe pointed out long ago, we cannot directly will ourselves into having a desire, but we can only undertake steps to cause ourselves indirectly to have a certain desire.<sup>48</sup>) The trouble with source internalism, as Parfit and others have argued, is extensional; it allows agents with certain desires to have no reason to want to avoid future agony for its own sake. But everyone, regardless of their desires, Parfit urges, has a reason to want to avoid future agony for its own sake. Some source internalists bite the bullet and allow that they must accept extensional oddness,<sup>49</sup> but they point out that agents with the mental states required to have such extensional oddness in their reasons will be so bizarre that we are in no position to slander them with the label ‘irrational’.<sup>50</sup> Other internalists deny that their view has any such extensional oddness; by appealing to the unity of agency or connections between present and future mental states, they argue that internalism can block the problems of extension.<sup>51</sup> In response, externalists might insist that even bizarre folks, such as those

with future Tuesday indifference, are irrational, and that complex arguments about the unity of agency and the connection between present and future mental states cannot succeed without illicit appeal to substantive normative considerations inconsistent with internalism. There are internalist responses to these charges and the debate continues.

According to source externalism, reasons are grounded in normative facts, such as the fact that something is good or, indeed, in the fact that the consideration is a reason (in which case we might more accurately say that the reason is 'ungrounded'). Parfit calls such reasons 'value-based'. Since external reasons are grounded in values and normative facts over which we have no direct volitional control, they are given, not created, reasons. Here, too, there are standard objections, Mackie's metaphysical and epistemic queerness objections, and the more recent objection from Korsgaard that externalists cannot explain how external reasons 'get a grip' on agents.<sup>52</sup> In response to Mackie, externalists have pointed out that other truths we have good reasons to accept are also metaphysically queer, that the epistemology of normative truths need not entail a *sui generis* faculty but simply the faculty of reason and argument, and that the queerness charge in general depends on an illicit assumption that empirical truths provide the default model for understanding the nature of other truths and how we come to know them. In response to Korsgaard, externalists have insisted that being motivated by reasons is in part what it is to be rational, and so the question, why should an agent be motivated by externalist reasons? becomes the arguably otiose, why should a rational agent be motivated by the reasons she has?<sup>53</sup> But surely queerness is a nontrivial cost, even granting partners in crime, and perhaps Korsgaard has in mind by her 'normative problem', not a problem about motivation but a problem about volitional engagement and a broader view about the standpoint from which we must understand such reasons. The debate continues.

Some version of the debate between internalism and externalism has existed for centuries. One might be forgiven, then, for thinking that, despite the increasingly nuanced and sophisticated argument on both sides, the debate is in a stalemate, gridlock, or, at any rate, not making progress on the question of what makes a normative element normative. Indeed, the arguments proposed by one side, even when considered to be decisive refutations of the opposing view, make nary a dent in the convictions of those holding the opposing view. A gridlock in long-standing philosophical debate is often a sign that the framing or assumptions of the debate need revisiting and more careful examination. I suggest we examine the assumption that all reasons must be either desire based or value based. By entertaining the possibility that the third dogma is false, we revisit the assumption that all reasons are given to us and never created by us.

Suppose for the sake of argument that the third dogma is mistaken. What should we believe in its place?

In other work, I have suggested that by abandoning the assumption that all reasons are given to us, we can arrive at a 'hybrid' metanormative view that recognizes not only given reasons – whether desire based or value based – but also, centrally, 'will-based' reasons that we create.<sup>54</sup> If this view is correct, it provides a way to break gridlock about the grounds of normativity by injecting a different view about how its possible grounds are related and rejecting a key assumption shared by traditional positions in the debate, namely, that all reasons are given.<sup>55</sup>

We create reasons through the volitional activity of 'commitment'. Commitments, as I understand them, are the volitional activity of taking something to be a reason by putting one's very self behind the consideration. By putting our very self behind some

consideration, we can endow that consideration with the normativity of a reason. A commitment is that in virtue of which that consideration is a will-based reason. We thus 'create' reasons by engaging in an activity of willing that is their ground. Willing is something over which we have direct volitional control since it just is the exercise of direct volitional control. So will-based reasons, unlike given reasons, are creatures we make, not discover.

For present purposes, the details of such a view need not concern us. Instead, I want to explore a few initial thoughts about how created reasons can be binding on us. Elsewhere I have argued that if we understand the relation between created and given reasons in the right way, we have a Goldilocks solution to any extension problem.<sup>56</sup>

The bindingness problem, recall, is that it seems implausible that reasons whose grounds are matters over which we have direct volitional control could bind us as ordinary reasons do. After all, if we have direct volitional control over the ground of a reason, we can just will that ground away and the reason thereby disappears. So how could will-based reasons bind?

We might begin by pointing out that, as far as phenomenology goes, the ground of a reason does not appear to determine whether a reason is binding. It seems that (i) we can make self-promises, self-vows, and resolutions; (ii) we can, in principle, understand these phenomena as commitments in the technical sense proposed, i.e. as putting oneself behind a consideration or plan of action that confers normativity on that consideration or plan by being the ground of its normativity; and (iii) the reasons thereby generated are binding on us as reasons. If I commit to exercising and eating right, if I put my very self behind these activities and, by hypothesis, endow those activities with value, then the reasons I have to exercise and eat right seem to bind just as reasons I would have to engage in those activities on my doctor's recommendation. Indeed, the reasons that are generated by my commitments may seem even more binding since I put my very self behind them.

One way to 'test' the bindingness of a reason is to ask whether the existence of its ground is subject to the vagaries of contingent circumstances. If the grounds of both given and created reasons can be contingent, and indeed contingent on the very same change in circumstances, the mere fact that committing can be contingent is no reason to think that the will-based reasons it grounds cannot be binding.

Suppose you are contemplating whether to pursue a career in philosophy or in architecture. You might be passionate about philosophical questions and thereby have a desire-based given reason to pursue philosophy. Exploring philosophical questions might also make your life go well, thereby grounding a value-based given reason to be a philosopher. And you might commit to the satisfaction of thinking about complex philosophical problems and thereby create a will-based reason to pursue philosophy. These are some ways you might come to have both given and created reasons in thinking about which of two careers to pursue.

Now suppose you wake up the next morning with what your doctor has diagnosed as long-term depression. This contingent change in your circumstances can destroy both your given and created reasons. If you are depressed, you lose your passion for philosophy and, along with it, your desire-based reason to become a philosopher. Long-term depression may also make philosophical work bad for you, and so you lose your value-based reason to pursue philosophy. The same holds for your will-based reason; your new mental condition now makes it impossible for you to sustain your commitment to the satisfactions of philosophical work since all such work seems pointless. Your will-based reason to

pursue philosophy disappears. Depression, winning the lottery, world war, and so on are contingent circumstances that can destroy the grounds of both our given and created reasons. The ground of reasons can be contingently destroyed in the blink of eye. But this does not make them nonbinding as reasons.

Of course, created reasons are distinctive in that they can be created or destroyed by an activity of will. But notice that you can also change the ground of your given reasons through volitional activity. By undertaking to spend more time with the local Save the Planet chapter, you can replace your passion for philosophy with a new passion and thereby make your desire-based reason to pursue philosophy disappear. And by resolving to live the high life, you can change the evaluative facts about what makes your life go well; you may no longer have the same value-based reason to pursue the pecuniary life of philosophy. The difference between given and created reasons is that you can destroy your created reasons through volitional activity directly, and you can destroy your given reasons through volitional activity only indirectly. Is this difference one that can sustain a distinction in bindingness? If there were a readily available pill that you could swallow, or an app you could launch on your smartphone, or a type of meditation you could undertake that would instantaneously cause you to lose certain desires or have new ones, would it follow that desire-based reasons were not binding? Each of these ways of destroying the grounds of your desire-based reasons involves an intermediate causal step and thus only indirect volitional control, but it is hard to believe that this slight difference in the causal chain could make the difference as to whether the corresponding reason was binding.

The third dogma supposes that the grounds of normativity must be beyond our direct volitional control for reasons to be binding. But why should we assume this? Indeed, will-based reasons may be more binding than given reasons since they involve putting our very selves behind the consideration that is the reason and help constitute our rational identities.<sup>57</sup> Without an account of why the directness of control makes a consideration nonbinding, especially in light of the similar contingency of given reasons, thinking that it must do so may amount to sheer prejudice against will-based reasons.

#### 4. Two Views of Rationality

For many centuries, thinkers have understood ‘rationality’ roughly in terms of two core capacities: first, the capacity to recognize reasons (or values or other normative elements), and, second, the capacity to respond appropriately to what is recognized by thinking, feeling, or acting accordingly. There are other capacities, too, on which these core capacities depend, but it has been almost universally supposed that the job of a rational agent can be broadly described as recognizing reasons and responding to them appropriately. Leading contemporary philosophers, such as Joseph Raz and Susan Wolf, explicitly gloss rationality in these terms.<sup>58</sup> Indeed, some neo-Kantians think that rationality so understood provides a counterpoint to worries about freedom of the will: the freedom that matters is the freedom to recognize and respond to normative reasons.

This orthodox view sees rational agency as fundamentally passive in that it is no part of rational agency to create reasons by grounding them in volitional activity. We might call such a view of rationality the *Passivist View*. According to the opposing *-Activist View*, rationality centrally includes the capacity to create reasons by willing their grounds.

The Passivist View is supported by the 3 dogmas. The relation is not deductive, but the dogmas and the Passivist View are mutually supporting. If normative truths can be absolute, and normativity does not fragment into only relativized truths, it is natural to think of normativity as a single, absolute force like gravity whose manifestations we must discover and respond to accordingly. If the structure of normativity is trichotomous, as is the structure of many nonnormative considerations like length, it is easy to think that normativity, like length, is something to be recognized. And this trichotomous structure permits a pleasing isomorphism between the three apparent basic normative relations, on the one hand, and the three apparent basic rational responses to reasons, on the other, viz., (i) favoring/preferring/choosing, (ii) disfavoring/dispreferring/choosing the other, and (iii) being indifferent or flipping a coin/picking between items. When none of the three basic relations holds, the matter falls outside the realm of rational agency since the items cannot be compared and the agent is left to nonrational ‘plumping’. Finally, if normativity is given to us rather than created by us, it is natural to think that it is something to be recognized and responded to. Given the 3 dogmas, it is no wonder that the orthodox view of rational agency has reigned supreme. The support is mutual. Starting with the idea that practical rational agency is a matter of recognizing and responding to reasons, it is natural to adopt a view of normativity that parallels views in the sciences in which rational inquiry is primarily a matter of recognizing and responding to reasons for belief. Just as scientific truths can be absolute, so can normative truths; just as comparisons of scientific elements are trichotomous, so too are normative comparisons; just as the domain of scientific inquiry is a matter of given reasons, so too is the domain of normative inquiry.

I believe we should reject the 3 dogmas and the Passivist View of rationality that goes along with it. Instead, we should adopt the Activist View that grows out of the claims suggested to take the place of the 3 dogmas. First, all first-order normative truths are relativized to ordinary substantive normative considerations, i.e. covering considerations. When we ask what is choiceworthy or rational or what we have most reason to do, we need to specify what matters in the choice or evaluation before we can properly answer the question, where what matters is given by a covering consideration. Normative theorizing tends to assume the logical form of normative truths is unrelativized, but the proposed alternative view of rationality maintains that this is a mistake. Being rational is not a matter of maximizing some absolute consideration; instead, it fragments according to the most general ordinary substantive considerations to which normative truths about what you have most reason to do can be relativized. Thus, we cannot rely on absolute, abstract rankings of values or goods to determine which of two courses of action is rational. Nor can we ‘keep track’ of the rationality in our lives as some kind of property or force that aggregates with each thought, feeling, or action we perform. Being rational is a fragmented affair.

Second, normativity is tetrachotomous, not trichotomous, in structure; items can be on a par. We need to expand the range of possible rational responses to our reasons. Instead of thinking that there are only three basic responses – favoring, disfavoring, or being indifferent – we add a fourth possible response of the utmost importance: commitment. Commitment is a volitional activity, a putting of our very selves behind some normative element and standing behind it, thereby endowing it with normativity grounded in our will.

This leads us to a third important feature of the Activist View. Normativity is not always given to us; our reasons and values are not always to be discovered, grounded in considerations over which we lack direct volitional control. Instead, we can quite literally create



reasons through the volitional activity of commitment and make it the case that we have most reason to do one thing rather than another. Some of our reasons are desire based or value based, that is, grounded in considerations over which we have only indirect volitional control, and some of our reasons are 'will based', that is, grounded in the volitional activity of commitment. When we make a commitment, we provide the grounds for will-based reasons and thereby create them.

The critical difference between the Passivist and Activist Views of rationality concerns the role of agency in determining what reasons we have. On the Passivist View, we exercise our agency in discovering given reasons and figuring out how appropriately to respond to them. This can involve 'active' agency in a broad sense since recognizing and responding to reasons requires the exercise of agency and is often difficult, focused, and deliberative work. Our interest, however, is in 'activity' of a specific sort; the activity of determining directly through one's own agency what reasons one has.

The Activist View allows that we can create reasons for ourselves by willing their grounds.<sup>59</sup> This view is active in two dimensions: it understands rational agency as centrally involving the capacity to create reasons through volitional activity, which provides their grounds. A rational agent can thus make it true for herself that she has most reason to pursue one course of action over another. The Activist View also explains how agents can constitute their own rational identities. Through our commitments, we can make it true that we have most reason to be academics rather than architects, where 'making true' is something we do directly, by creating reasons in favor of one option over another. Moreover, we can create reasons to be in one kind of choice situation as opposed to another and thereby carve out one path through life rather than alternative available roads not taken. The Activist View of rational agency gives rational agents the normative power to determine what they have most reason to think, feel, and do.

## 5. Some Applications

Abandoning the 3 dogmas and adopting their suggested substitutes has some possibly interesting implications for a number of practical matters. I end by outlining some thoughts about how the Activist View of rationality might affect our thinking in certain domains and about a number of practical issues.

### 5.1. *Decision Theory and Economic Approaches to Choice and Valuation*

Decision theory, including rational and social-choice theory, insofar as they claim to be normative, assume all 3 dogmas. If claims about what is to-be-preferred are erroneously assumed to be absolute, and if it is a mistake to take as definitional of an ordering that there are only three possible basic ways items can be ordered, the very foundations of such approaches to rational choice and valuation are mistaken. Moreover, if rationality involves not just the discovery of reasons but also their creation, standard forms of decision theory, such as expected utility theory, fail to model rationality as it is properly understood. Decision theory would at best be a fundamentally mistaken model of human choice, and we would need to rethink its most basic assumptions in order to accurately represent values and reasons as they truly are – tetrachotomous in structure and not contributors to a single consideration of being what is best or choiceworthy or rational. At the very least,

decision theory, were it to reject the 3 dogmas, would need to distinguish two different phenomena within the gaps of a partial ordering, parity and genuine incomparability, and to add a fundamental attitude to choice beyond preference, dispreference, and indifference, e.g. commitment.

### 5.2. *Governmental Regulation*

Governmental regulations in the United States, the United Kingdom, and many other countries rely not only on the assumption that value and reasons are trichotomous in structure but often go further and assume the commensurability, i.e. cardinal comparability, of all goods involved in a regulatory trade-off. Cost-benefit analysis is widely employed and indeed mandated as the approach to formulating policy at the FDA and other governmental agencies. But cost-benefit analysis assumes all 3 dogmas and assumes further that absolute truths about correct trade-offs between values can be discovered through the common cardinal metric, usually of dollars. If the 3 dogmas are false, then the standard methods of making regulatory decisions are flawed and would require reassessment at best and arguably a fundamental overhaul from root to branch. Of course, the difficulty in overhauling such approaches is finding the sweet spot between accurate reflection of the values at stake and ease of mathematical representation. With a tetrachotomous framework in place, I believe that models could be developed that do better than current models.

### 5.3. *War and Conflict*

Suppose we could put a chemical in the world's drinking water that caused everyone to reject the 3 dogmas and accept the proposed alternatives instead. If, for example, everyone in the world believed that many ways of life are on a par, without there being a best, perhaps the impulse to righteousness that leads to so much conflict in the world would be undermined or at least slightly diminished. If people around the world truly accepted that, while at the same time rejecting the unattractive relativity often thought to be the only alternative to intolerance, they would recognize that different ways of doing things are often on a par, and in such cases, an agent's or group's commitment could make a way of life best for them, even though, given our commitments, we think that is an inferior way of life. Perhaps then the default outlook might be one of cautious tolerance rather than reflexive judgement, denigration, and fear of the different. Perhaps war would be less ominously just over the horizon.

### 5.4. *Religion and Science*

Perhaps the evidence for belief in the existence of a god may be on a par with evidence for suspension, withdrawal, or even disbelief. If rational agency is creative, not only in the practical but also in the theoretical sphere, then an agent may create reasons for having one attitude over another for herself. In this way, creative rational agency can help explain why rational agents can have very different attitudes towards matters of such significance while those matters are supported by qualitatively very different sorts of evidence. Moreover, scientific inquiry need not be understood as a wholly passive affair; there may be multiple standards of justification appropriate for a proposition that are on a

par, and it may be that by committing to features of one hypothesis (such as its higher statistical probability), an inquirer can ground reasons that thereby justify that hypothesis over others. In this way, in hard cases, rational agency can contribute not only to what we have most reason to do but also, more controversially, to what we have most reason to believe. Laura Callahan is exploring how much our believing might be ‘up to us’.<sup>60</sup>

### 5.5. *Egalitarianism*

Richard Arneson writes in the *Stanford Encyclopedia of Philosophy* that ‘an egalitarian favors equality of some sort: people should get the same, or be treated the same, or be treated as equals, in some respect’.<sup>61</sup> Equality of opportunity is the ideal that people with the same merits should be given the same opportunities to compete for advantages. Equality of resources maintains that people should have the same resources or, more plausibly, that they enjoy real or effective freedom to the same extent given and thus that resources be distributed accordingly. Marxist equality is the ideal that everyone has the same right to receive what she needs relative to her contribution according to her ability. Relational equality holds when people relate as equals and have the same moral status. In short, egalitarian ideals tend to be underwritten by a notion of equality as amounting to sameness in some respect. This suggests that egalitarians are interested in equality in the strict, logical sense, that is, in a relation holding between things with respect to some covering consideration that is reflexive, symmetric, transitive, and, most importantly, such that if you improve one of two equals, it follows that the improved item is now better than the other and not equal. You treat two children equally if you treat them the same, by giving them the three Smarties each; and if you treat one a bit better, by giving her an extra Smartie, you no longer treat them equally. The logical form of equality rules out ‘equality’ as parity.

If normativity is tetrachotomous and not trichotomous in structure, we might understand egalitarianism more expansively; the egalitarian (or ‘paritarian’) ideal does not demand sameness in rights or resources or opportunities but allows, for example, rights to religious expression, resources, and opportunities for betterment to be on a par. An expansive egalitarian (or ‘paritarian’) recognizes that there are plural, qualitatively different ways in which a right to, say, free speech, can be had, qualitatively different kinds of resources that may make various distributions on a par, and qualitatively different opportunities or moral statuses that are not the same or equal but instead are on a par. This more expansive form of egalitarianism (or ‘paritarianism’) reflects, I believe, the deep motivation for traditional egalitarian principles while eschewing the undefended second dogma that drives its traditional form.<sup>62</sup> In short, what a just society owes us, then, is strictly speaking parity, not equality.

### 5.6. *Affirmative Action*

The debate over affirmative action is arguably still stuck between the Scylla of Lyndon-Johnson-style affirmative action, which concerns getting more underrepresented minorities to apply for opportunities, and the Charybdis of narrow understandings of merit according to which it is unfair to award opportunities to the ‘less qualified’ simply on the basis of their gender or other diversity characteristics and the attendant stigma and stereotype threat that may follow. With the Activist View of rational agency in hand,

we can understand affirmative action in a new way: often candidates are on a par with respect to the substantive considerations that matter in the choice between them, which may or may not include their contribution to diversity. In this case, an institution can commit to some feature – like the diversity characteristic of a candidate – and thereby confer normativity on that characteristic, which may then make that candidate best for that institution. So ‘affirmative action’ can be understood instead in terms of the affirmative commitment of an institutional decision-maker that grounds a will-based reason to choose a minority candidate who is on a par with other nonminority candidates with respect to given reasons. By making such a commitment, the decision-maker, via their decision-making authority within the institution, helps to constitute the rational identity of that institution. Some institutions faced with candidates that are on a par may choose to commit differently or not to commit at all. Their institutional identity will then be one in which affirmative action does not figure but, say, social cohesiveness with existing personnel does. The important shift in understanding affirmative action is that the issue is no longer centered on the unproductive question of how ‘merit’ should be understood but is instead a matter of the rational identity of the institution making the decision. At an institution that takes the affirmative action of committing to diversity, a diversity candidate in a case in which the given reasons are on a par will in fact be best, all things considered, no matter how narrow and conservative a notion of ‘merit’ is employed. This way of thinking about affirmative action may then remove any stigma associated with being ‘less qualified’. Chosen candidates are in fact not less qualified but all-things-considered best.

### 5.7. *Law and Legal Reasoning*

The Activist View of rational agency has interesting implications for legal adjudication and legal reasoning. If we reject the 3 dogmas and adopt their suggested substitutes, we can see that two giants of jurisprudence, H.L.A. Hart and Ronald Dworkin, were both right about what the law is.<sup>63</sup> Hart was right that legal reasons often run out and that the judge or legislator has to step in and do something. Dworkin was right in that there is always a right answer in law. How can they both be right? The reasons to find for the plaintiff against the defendant are often on a par because normativity is tetrachotomous in structure. In such ‘hard cases’, the judge can commit to some consideration in favor of the plaintiff thereby creating a will-based reason to rule for the plaintiff. Thus there is always a right answer, even in hard cases. At the same time, the given legal reasons ‘run out’ – not because of vagueness or uncertainty à la Hart, but because as a substantive matter, the options are on a par. The judge does not ‘create’ law, but she creates will-based reasons that make her ruling in favor of the plaintiff correct as a matter of law. On the Activist View of law’s rationality, will-based reasons are a central part of law’s empire.

### 5.8. *Business Ethics*

One of the leading concerns about the rise of the corporation in the world’s economies derives from the fact that in many countries, such the United States, corporations are mandated by statute to have shareholder interests as their primary concern. That is to say, a corporation’s purpose is to maximize profit for its shareholders. Given the relatively significant power corporations wield in many societies, this results in an uncomfortable distortion of the place of wealth and monetary gain in human life. One response to this

distortion is to propose that corporations are beholden to ‘stakeholder’ rather than ‘shareholder’ interests, where stakeholder interests include factors that contribute to a shareholder’s well-being beyond the accumulation of wealth. But stakeholder interests are, arguably, still too narrow since they do not encompass public goods. One interesting further response, proposed by Colin Mayer, is to reconceptualize the purpose of a corporation as ‘solving problems profitably’, that is, making wealth maximization play an adverbial role rather than being the aim of corporate decision-making and allowing that corporations should take on social problems, such as global warming, not just problems of wealth maximization.<sup>64</sup> Mayer’s important and admirable suggestion seems subject to two worries: first, practical difficulties of implementation, e.g. those who lead the charge by changing their purpose will likely lose out in market share to the sharks who continue to see their sole purpose as the maximization of profit, and, second, democratic worries about nonelected CEOs having authority to solve social problems, e.g. should Elon Musk determine free speech conditions on Twitter?

Mayer’s argument also assumes the second and third dogmas and the Passivist View of rational agency. If we took seriously the idea that rational agency involves the creation of reasons, we might consider an alternative approach to blunting the distortions of the rise of corporate power. Instead of allowing businesses to wade into the treacherous waters of attempting to solve social problems, we might require them to ‘stay in their lane’ of profit making but to recognize that a vast number of decisions made within a firm – from which paper to order to how to invest earnings – are between options that are on a par. In such cases, on an Activist View of rational agency, corporate decision-makers have the normative power to commit to some consideration, such as the importance of contributing to reductions in greenhouse gases, and create will-based reasons to choose the option – the bamboo-based paper and the green investment portfolio – that their commitment supports. In this way, we can blunt the distortion that wealth maximization causes when corporations become increasingly powerful in a society while both preventing abdication of democratic authority to nondemocratic business leaders and harnessing – albeit conservatively – the power of businesses in helping to promote the social good.

### 5.9. Healthcare Ethics

The ethics of healthcare apportionment has typically followed a prioritarian principle; the worst off should be prioritized in the distribution of healthcare resources. The traditional debate has then focused on who is the worst off: is it those who are now suffering the most or those who have the worst health outlooks looking toward the future? In the late 60s and 70s, social scientists offered a third interpretation that has gained traction in policymaking, such as the evaluation of the cost-effectiveness of drugs. Those who should have priority in healthcare benefits are those with the worst Quality-Adjusted Life Years (QALYs), where a QALY is a measure (a number between 0 and 1) representing the quality of health a patient has enjoyed over the year (where 0 represents death and 1 perfect health). Those with the lowest QALYs, summed from the past and projected into the future, are the worst off and deserve priority in the distribution of healthcare.<sup>65</sup> But QALYs are a technical measure of health that make not only trichotomous assumptions about the evaluative components of health but cardinal ones that rule out organic unities among components of healthfulness. There is no room for parity in either traditional or the newer QALYs approach to healthcare distribution. There is also the assumption that

there is one best distribution to be discovered. If the Activist View of rationality is correct, rational choices about healthcare provision will often be a matter of commitments to certain components of health over others. This would allow for a hospital, for instance, to create reasons for distribution that prioritize a patient's indigent circumstances, say, over the painfulness of the disease in deciding to whom to distribute scarce health benefits. Anders Herlitz has been pursuing this line of research.<sup>66</sup>

#### 5.10. *AI Design and Development*

The two main problems in AI research are the control problem – how we prevent AI from becoming our overlords – and the alignment problem – how we get AI to perform functions that align with our human values. The two problems appear to be related because if we can get AI to align with our values, presumably we will have achieved some *de facto* control over it. Some computer scientists have proposed ideas for solutions to each of these problems while others have argued that, at least if AI achieves general intelligence, we are lost. Philosophers are now entering the fray. Peter Railton has recently offered an arresting and novel approach to tackling these problems by suggesting that we build AI with the same bundle of premoral linguistic, epistemic, social competencies that empirical data suggest is the basis on which animals, small children, and adult humans learn to cooperate to gain rewards that rely on that cooperation.<sup>67</sup> His suggestion is that, rather than being circumspect about building AI, we need to build significant numbers of AI, each with this bundle of premoral capacities and motivations, so that they can create cooperative communities with one another and with us. These cooperative communities will, he suggests, provide a check on AI outliers that try to dominate other AI or humans in the same way that hunter-gatherer communities would check hunters who would refuse to share their spoils even with the rest of the community. And as collaborators, their values would more or less align with our own.

Central to his provocative proposal is reliance on reward-based learning and the model of expected utility functions upon which such learning depends. But expected utility functions, at least as they are currently employed in AI deep learning, assume all 3 dogmas. If we accept their substitutes, we might propose an alternative approach to confronting the control and alignment problems. What if we design all AI so that tetrachotomy and not trichotomy is the structure against which an AI must make a decision? When an AI determines that options are on a par, there is room for human commitment to resolve the choice. This simple design hack would guarantee that human commitment was in the loop whenever options were on a par. Peter Eckersley, Brian Christians, Bryce Goodman, and I have noted the problem posed by the incommensurability of values and hard choice,<sup>68</sup> and some computer scientists have begun to explore solutions of the sort in line roughly with the view of rationality favored here.<sup>69</sup> While bringing humans in the loop in this way could engender inefficiency, it would go some way towards making more manageable the problems of alignment and control. If, instead, we blithely continue to assume the 3 dogmas, we may be on the brink of building a new digital world on what could be a fundamentally mistaken view of what it is to be rational.

*Ruth Chang, University of Oxford, Oxford, UK. [ruthechang@gmail.com](mailto:ruthechang@gmail.com)*



## Acknowledgements

This article is based on the annual Society for Applied Philosophy Lecture delivered on May 20, 2022, and I am grateful to the audience who braved a wave of COVID-19 to attend and participate in discussion. Early versions were trialed at other venues, and I thank those audiences for their questions and suggestions. Many thanks to the Society for inviting me to give the lecture and for enticing Pamela Hieronymi, Onora O'Neill, Philip Pettit, and Peter Railton, four brilliant philosophers whom I have long admired, to be my interlocutors. I am greatly honored. I owe Elizabeth Brake and Massimo Renzo special thanks for organizing everything and for rolling with the COVID punches.

## NOTES

- 1 In relation to the first dogma, see Chang, Ruth. 1997. "Introduction." In *Incommensurability, Incomparability and Practical Reason*, edited by R. Chang, 1–34. Cambridge, MA: Harvard University Press; Chang, Ruth. 2004a. "Putting Together Morality and Well-Being." In *Practical Conflicts*, edited by M. Betzler and P. Baumann, 118–58. Cambridge: Cambridge University Press; and Chang, Ruth. 2004b. "All Things Considered." *Philosophical Perspectives* 18: 1–22. In relation to the second dogma, see Chang, Ruth. 2002. "The Possibility of Parity." *Ethics* 112: 659–88; Chang, Ruth. 2005. "Parity, Interval Value, and Choice." *Ethics* 114: 331–50; Chang, Ruth. 2012. "Are Hard Choices Cases of Incomparability?" *Philosophical Issues* 22(1): 106–26; Chang, Ruth. 2016a. "Parity: An Intuitive Case." *Ratio* (special issue) 29: 395–411; and Chang, Ruth. 2016b. "Hard Choices." *The American Philosophical Association Journal of Philosophy* 92: 586–620. <https://doi.org/10.1017/apa.2017.7>. In relation to the third dogma, see Chang, Ruth. 2009. "Voluntarist Reasons and the Sources of Normativity." In *Reasons for Action*, edited by David Sobel and Steven Wall, 243–71. New York: Cambridge University Press; Chang, Ruth. 2013a. "Commitments, Reasons, and the Will." *Oxford Studies in Metaethics* 8: 74–113; Chang, Ruth. 2013b. "Grounding Practical Normativity: Going Hybrid." *Philosophical Studies* 164(1): 163–87. <http://link.springer.com/article/10.1007%2Fs11098-013-0092-z>; Chang, Ruth. 2015b. "Transformative Choices." *Res Philosophica* (special issue) 92(2): 237–82. <https://doi.org/10.11612/resphil.2015.92.2.14>; Chang, Ruth. 2020. "Do We Have Normative Powers?" *Aristotelian Society Supplement* 94: 275–300; and Chang, Ruth. 2021c. "What Is It to Be a Rational Agent?" In *The Routledge Companion to Practical Reason*, edited by R. Chang and K. Sylvan, 95–110. New York: Routledge.
- 2 Moore, G.E. 1903. *Principia Ethica*. Cambridge, UK: Cambridge University Press.
- 3 It might be worth noting that since appeal to a superproperty is only one way in which absolute claims might be understood, the first dogma is not equivalent to monism about normativity. The dogma (i) allows that there are *some* relativized normative truths, which may not be reducible to truths relativized to a superproperty, and (ii) even if all relativized truths reduced to absolute ones, it does not follow that they do so via a metaphysical superproperty; there are many relations that could hold between substantive normative considerations and a superproperty other than reduction even if the claims involving them reduce.
- 4 Again, this dogma should be read to encompass other sorts of possible absolute first-order normative truths, such as ones about what is rational, choiceworthy, best, supererogatory, suberogatory, permissible, optional, reasonable, virtuous, perfect, good enough, sufficient, protanto reason providing, and so on. Since its concern is only with first-order normative claims, it leaves open the possibility that higher-order normative claims are absolute (a possibility that helps to block regress worries that might arise if one rejects the first dogma, as I suggest we should, in favor of the view that all normative claims are relativized to ordinary substantive normative considerations). I have presented the dogma in its most plausible, 'weak', form, according to which only *some* normative truths are absolute. A stronger form would hold that *all* such normative truths are absolute, and that all relativized claims are reducible to these absolute claims. I have also sidestepped very tricky questions about relativization; note that the relativization of interest is to substantive normative considerations, not to time, places, and circumstances. Nor is the relativization of interest that of domains; Sidgwick thought that practical normativity divided into two irreducibly distinct domains, the moral and the prudential (see Sidgwick, Henry. 1874 (1981). *The Methods of Ethics*. London: Hackett Classics), and the first dogma

- should be understood as claiming that there can be absolute normative truths within a domain, whichever those domains might be.
- 5 A diagnosis specific to the second way of understanding simpliciter claims is that there seems to be no way to explain why all good things are good (or normative things normative) without appeal to some superproperty of being good (or of being normative). This thought is what motivated Moore to think that there could be absolute goodness claims. I suggest an alternative way of understanding what all good (or normative) things have in common below.
  - 6 I have tried to address the challenge that there are no such substantive considerations in Chang 2004a, b, op. cit., where I argue that there are often nameless 'unities' that explain how seemingly disparate substantive considerations are put together. While there are more unities than we might think, there are limits on how normative elements can be put together. One test I have proposed is the 'nominal-notable' test. If a nominal instance of one normative element can be compared with a notable instance of another, seemingly disparate, normative element, then there could well be a unity that puts them together. For a contrary position, see Copp, David. 1997. "The Ring of Gyges: Overridingness and the Unity of Reason." *Social Philosophy and Policy* 14(1): 86–106; Copp, David. 2021. "Normative Pluralism and Skepticism About 'Ought Simpliciter'." In *Routledge Handbook of Practical Reason*, edited by R. Chang and K. Sylvan, 416–37. New York: Routledge. But you do not have to accept expanding the boundaries of normative unities to reject the first dogma.
  - 7 Parfit, Derek. 2011. *On What Matters*, Vol. II. Oxford: Oxford University Press; Scanlon, Thomas. 2014. *Being Realistic About Reasons*. Oxford: Oxford University Press.
  - 8 Chang 1997 op. cit., p. 6.
  - 9 Thomson, Judith Jarvis. 1997. "The Right and the Good." *Journal of Philosophy* 94(6): 273–98; Thomson, Judith Jarvis. 2003. *Goodness and Advice*. Princeton, NJ: Princeton University Press; Thomson, Judith Jarvis. 2008. *Normativity*. Chicago, IL: Open Court.
  - 10 Another possible interpretation of simpliciter claims is in terms of *intrinsic* goodness. 'X is good, simpliciter' might be understood as the claim that, taking into account all the *intrinsic* goods, often understood as all the 'ethical' goods, X has intrinsic value (see Zimmerman, Michael J. 1999. "In Defense of the Concept of Intrinsic Value." *Canadian Journal of Philosophy* 29: 389–410). In these cases, 'X is good, simpliciter' expresses a complete thought, but not one which entails absolute normative truths as opposed to truths that are relativized to a type of ordinary – intrinsic – substantive normative considerations.
  - 11 If parity is possible, it will in general be a mistake to assume that Pareto superiority entails betterness. See Chang 2016a op. cit.
  - 12 See Pettit, Philip, 1991. "Consequentialism." In *A Companion to Ethics*, edited by Peter Singer, 230–40. Cambridge: Blackwell.
  - 13 None of this is to say that the first dogma is only assumed by consequentialists. The idea that normative claims are absolute is often assumed by deontologists, virtue theorists, perfectionists, etc. I think, for example, that it is at work in the classic statement of Kantian ethics of Onora O'Neill (see O'Neill, Onora. 2013. *Acting on Principle: An Essay on Kantian Ethics*. Cambridge: Cambridge University Press). As I will suggest, if we reject the first dogma, normative theorizing of whatever stripe should focus on neglected, difficult questions about the normative relations that might hold among ordinary substantive normative considerations.
  - 14 Since I have presented the dogma in its 'weak' form, that is, as the claim that only *some* normative truths are absolute, this worry might seem misplaced. Perhaps only those cases in which it makes sense to ask of two good things, which is better, are exactly those cases in which the question and its answer proceed in terms of an absolute normative consideration such as being good, simpliciter. But the worry is mitigated only insofar as the scope of being good, simpliciter matches up with the intelligibility of asking of two good things, which is better. And it is unclear whether there is any *a priori* reason to think that there will, indeed, be such a match.
  - 15 Recall that the first dogma is meant to hold within domains of normativity if there are such domains. Even within a domain, diversity undermines the idea that disparate substantive considerations can be put together exactly when theorists assume there is an absolute normative claim.
  - 16 See Rachels, Stuart. 2001. "A Set of Solutions to Parfit's Problems." *Nous* 35(2): 214–38; Temkin, Larry. 2012. *Rethinking the Good*. Oxford: Oxford University Press, p. 137. If we understand the Rachels–Temkin arguments as not presupposing the first dogma but instead as claiming that only certain qualified relations, such as 'betterness with respect to painfulness', are nontransitive, then it might be thought that those arguments would need to be supplemented with (i) detailed axiological arguments about the contributory components of painfulness and how they are normatively related in order to block the charge that the arguments equivocate across different conceptions of painfulness, and (ii) a clear declaration that it is these relativized relations that are nontransitive, not the more general relation of being better than, all things considered. I have

- argued that so-called 'continua' or 'spectra' arguments, whether they rely on the first dogma or not, can be avoided, once equivocation has been ruled out, by recognizing the qualitative change with respect to that substantive consideration that occurs along the spectrum so that items are no longer successively better than one another but on a par with one another. Parity breaks the putative chain of betterness to the repugnant conclusion. See Chang 2021a op. cit., *passim*.
- 17 Related puzzles over the nontransitivity of justification would also disappear. See Dorsey, Dale. 2016. *The Limits of Moral Authority*. Oxford: Oxford University Press, pp. 119–26.
  - 18 Although exclusionary reasons are second-order reasons, the claim that one has a reason that silences or cancels first-order reasons can be regarded as a first-order normative claim insofar as it is a claim about what is good or what we ought to do. In any case, one might think that second-order normative claims can also be relativized to ordinary substantive normative considerations.
  - 19 Some thinkers have suggested that reasons have two distinctive *kinds* of normative weight, namely *justifying* weight and *requiring* weight. See Gert, Joshua. 2003. "Requiring and Justifying: Two Dimensions of Normative Strength." *Erkenntnis* 59(1): 5–36; Gert, Joshua. 2016. "The Distinction Between Justifying and Requiring: Nothing to Fear." In *Weighing Reasons*, edited by E. Lord and B. Maguire, 157–72. Oxford: Oxford University Press; see also Tucker, Chris. 2022. "Parity, Moral Options, and the Weights of Reasons." *Nous*. <https://doi.org/10.1111/nous.12410>. A requiring reason makes omission of the act it requires impermissible, while a justifying reason makes an act that would otherwise be impermissible permissible. Putting aside other questions that could be raised about this distinction, it appears to trade on an assumption that truths about permissibility and impermissibility are absolute. For example, Gert argues that self-defense provides a justifying reason to do something – kill someone – that would otherwise be impermissible. Hence a special kind of 'justifying' normative weight. Here he is assuming that the claim 'it is impermissible to kill someone' is absolute. Once we relativize the claim to 'it is impermissible to kill a person with respect to the value of human life', there is no absolute impermissibility of killing that reasons of self-defense overturn since reasons of self-defense are also relevant to determining when you should not kill relative to the value of human life. Abandoning the first dogma renders the proliferation of such kinds of normative weight otiose.
  - 20 Dancy, Jonathan. 2004. *Ethics Without Principles*. Oxford: Oxford University Press.
  - 21 Kagan, Shelly. 1988. "The Additive Fallacy." *Ethics* 99(1): 5–31.
  - 22 Kamm, Frances. 1996. *Morality, Mortality: Rights, Duties, and Status, Vol II*. Oxford: Oxford University Press, p. 51; Kamm, Frances. 2007. *Intricate Ethics*. Oxford: Oxford University Press, p. 348.
  - 23 Chang 1997 op. cit., pp. 6–7; Chang 2004a, b op. cit., *passim*.
  - 24 Thompson 1997, 2003 op. cit.; Foot, Philippa. 1985. "Utilitarianism and the Virtues." *Mind* 94(2): 196–209; Geach, Peter. 1956. "Good and Evil." *Analysis* 17: 33–42.
  - 25 Chang 1997 op. cit., pp. 27–34.
  - 26 This idea should not be conflated with the idea that truth itself is relative. While the content of a normative truth is relativized to a substantive consideration, the truth itself may be absolute, i.e. not relativized to an agent's desires or a community's beliefs.
  - 27 Moore op. cit.; Wittgenstein, Ludwig. 1929. *Lectures on Ethics*. Cambridge: Cambridge University Press. Thanks to Duncan Richter for a fascinating conversation following the World Congress in Philosophy in China about Wittgenstein's take on the 3 dogmas: he probably accepted them.
  - 28 I have tried to make a start on both questions. In answer to the first, I argue that 'commitment', understood as the volitional activity of putting oneself behind something, helps to determine what matters in each set of circumstances and indeed which well-formed choice situation one should be confronting by grounding reasons and value (and thereby part of the story of how to avoid regress) (Chang 2021c op. cit.; Chang 2013b op. cit.), and in answer to the second, I argue that comparative truths and well-formed situations (in terms of which, I argue, trumping, canceling, outweighing, silencing, defeating, excluding, and various more complex apparently noncomparative relations can be understood) provide a way of focusing attention on the underlying logical form of the truths that determine how to live, regardless of which form of normative theory – utilitarianism, deontology, virtue theory, perfectionism, etc. – one favors. These truths, I argue, are comparative (see Chang, Ruth. 2015a. "Comparativism: The Grounds of Rational Choice." In *Weighing Reasons*, edited by B. McGuire and E. Lord, 213–40. Oxford: Oxford University Press). If this is correct, then the focus of normative theorizing should be on what comparative relations hold among normative elements.
  - 29 A basic relation is a member of a small set of relations that taken together exhaust the conceptual space of comparability between two items with respect to a covering consideration.

- 30 This dogma should be interpreted as allowing that the first dogma holds; thus we should allow that according to the second dogma, comparative relations are two-place, such as 'better than, simpliciter', rather than three-place, such as 'better than with respect to justice'. Each dogma is considered independently of the others.
- 31 There is a sense in which both dichotomists and trichotomists 'leave open' the range of basic relations of an ordering if we allow that items that fall in the gaps of a partial ordering are possibly rankable in some nontrichotomous way. But standard decision-theoretic approaches to normative orderings do not explore this possibility and thus are most plausibly interpreted as assuming trichotomy.
- 32 Chang 2009 op. cit.
- 33 Chang 2013a op. cit.
- 34 Chang 2016b op. cit.
- 35 Chang 2020 op. cit.
- 36 Chang, Ruth. 2021a. "How to Avoid the Repugnant Conclusion." In *Ethics and Existence: The Legacy of Derek Parfit*, edited by J. McMahan, T. Campbell, J. Goodrich, and K. Ramakrishnan, 389–429. Oxford: Oxford University Press.
- 37 Chang, Ruth. 2021b. "Are Hard Cases Vague Cases?" In *Value Incommensurability: Ethics, Risk, and Decision-making*, edited by H. Anderson and A. Herlitz, 50–70. New York: Routledge.
- 38 Chang 2021c op. cit.
- 39 As far as I am aware, there have been no positive arguments *as such* defending the claim that normativity is trichotomous as opposed to tetrachotomous in structure. Instead, there are negative arguments that attack positive arguments for parity or trichotomous incomparability (e.g. Klocksiem, Justin. 2010. "In Defense of Trichotomy." *Acta Analytica* 25(3): 317–27; Constantinescu, Cristian. 2012. "Value Incomparability and Indeterminacy." *Ethical Theory and Moral Practice* 15(1): 57–70; Elson, Luke. 2014. "Heaps and Chains: Is the Chaining Argument for Parity a Sorites?" *Ethics* 124(3): 557–71; Andersson, Henrik. 2015. "Parity and Comparability – A Concern Regarding Chang's Chaining Argument." *Ethical Theory and Moral Practice* 19(1): 245–53; Andreou, Chrisoula. 2015. "Parity, Comparability and Choice." *Journal of Philosophy* 112(1): 5–22) and positive (linguistic or metaphysical) arguments for complete determinate trichotomous rankings (e.g. Regan, Donald. 1997. "Value, Comparability and Choice." In *Incommensurability, Incomparability and Practical Reason*, edited by R. Chang, 129–50. Cambridge, MA: Harvard University Press; Broome, John. 1997. "Is Incommensurability Vagueness?" In *Incommensurability, Incomparability and Practical Reason*, edited by R. Chang, 67–89. Cambridge, MA: Harvard University Press; Williams, J. Robert. 2016. "Indeterminacy, Angst and Conflicting Values." *Ratio* 29: 412–33; Dorr, Cian, Jacob Nebel, and Jake Zuehl. 2022. "The Case for Comparability." *Nous*: 1–40. <https://doi.org/10.1111/nous.12407>). It is perhaps worth pointing out that I have offered a direct argument that our ordinary notion of comparability does not have trichotomy built into its meaning in the above discussion of the second dogma. Although there may be some pro tanto linguistic considerations for thinking that in some instances, comparability assumes trichotomy, this result is compatible with its being an open question as to whether comparability allows for tetrachotomy. Ultimately, the proof of parity is in the pudding; if parity can do distinctively important work, we should believe it exists. I have suggested that it can elsewhere.
- 40 Parfit, Derek. 2016. "Can We Avoid the Repugnant Conclusion?" *Theoria* 82(2): 110–27; see also Hsieh, Nien-he. 2005. "Equality, Clumpiness, and Incomparability." *Utilitas* 17(2): 180–204.
- 41 Elsewhere I have suggested three reasons to think that this strategy of maintaining trichotomy is mistaken: (i) that the resulting relation of being 'imprecisely equally good' has different logical features from the standard relation of being equally good and for that reason should be understood instead as an instance of a fourth basic relation, 'parity', (ii) that imprecise equality is in one way like betterness, in that the difference between items has magnitude, and in one way like standard equality in that the difference between items has no bias, and thus it is a mistake to privilege one similarity and conclude that it is a species of equality; instead it should be understood as a fourth relation, parity; (iii) that imprecise equality if a species of equality should hold of two identical objects that are nevertheless representable by the same imprecise cardinal measure (whatever that might be), but this fails to recognize the important difference between how two identical items might be imprecisely equal from how two rather different items might be imprecisely equal, and we can naturally mark that difference by understanding imprecise equality in the second type of case as parity. None of this is to say that 'imprecise' cardinal measurement, however one wishes to explain that idea, in conjunction with trichotomy, is not a perfectly acceptable and natural way in which to come to understand parity. But we should not conflate a trichotomous route to understanding parity with its reduction.

- 42 Aristotle. 2009. *The Nicomachean Ethics*, translated by W.D. Ross and Lesley Brown. Oxford: Oxford University Press.
- 43 Korsgaard, Christine. 1996. *The Sources of Normativity*. Cambridge: Cambridge University Press; Korsgaard, Christine. 2003. "Realism and Constructivism in 20th Century Moral Philosophy." *Journal of Philosophical Research* 28: 99–122.
- 44 These difficulties for the neo-Kantian view have been raised variously by Railton (Railton, Peter. 2004. "How to Engage Reason: The Problem of Regress." In *Reasons and Values: Themes from the Philosophy of Joseph Raz*, edited by J. Wallace, P. Pettit, S. Scheffler and M. Smith, 176–201. Oxford: Oxford University Press), Scanlon (Scanlon, Thomas. 2003. "Metaphysics and Morals." *Proceedings and Addresses of the American Philosophical Association* 77: 7–22), and perhaps most vividly by David Enoch (Enoch, David. 2006. "Agency, Shmagency: Why Normativity Will Not Come from What Is Constitutive of Agency." *Philosophical Review* 115: 169–98). Korsgaard, in answer to these challenges, has argued that it is a metaphysical fact about action that its good instances must be willed in conformity with the Categorical Imperative. (See Korsgaard, Christine. 2008. *The Constitution of Agency*. Oxford: Oxford University Press; Korsgaard, Christine. 2009. *Self-Constitution*. Oxford: Oxford University Press.) This move would, indeed, answer the challenge (and more), but it is controversial whether Korsgaard's ingenious arguments succeed.
- 45 Chang 2009, 2013a, b, 2016b, 2020, 2021c op. cit.
- 46 See Chang, Ruth. 2013c. "Practical Reasons: The Problem of Gridlock." In *Companion to Analytical Philosophy*, edited by B. Dainton and H. Robinson, 474–99. London: Bloomsbury Press.
- 47 Parfit 2011 op. cit.
- 48 Anscombe, Elizabeth. 1956. *Intention*. Oxford: Blackwell.
- 49 Schroeder, Mark. 2007. *Slaves of Passion*. Oxford: Oxford University Press.
- 50 Street, Sharon. 2009. "In Defense of Future Tuesday Indifference: Ideally Coherent Eccentrics and the Contingency of What Matters." *Philosophical Issues* 19(10): 273–98.
- 51 Smith, Michael. 2012. "Agents and Patients, or: What We Learn About Reasons for Action by Reflecting on Our Choices in Process-of-Thought Cases." *Proceedings of the Aristotelian Society* 112: 309–31; Smith, Michael. 2013. "A Constitutivist Theory of Reasons: Its Promise and Parts." *Law, Ethics and Philosophy* 1: 1–30; Smith, Michael. 2015. "The Magic of Constitutivism." *American Philosophical Quarterly* 52: 187–200; Sobel, David. 2017. *From Valuing to Value: A Defense of Subjectivism*. Oxford: Oxford University Press, chap. 9.
- 52 Mackie, John. 1977. *Ethics: Inventing Right and Wrong*. New York: Penguin Press; Korsgaard 1996, 2003 op. cit.
- 53 Another argument against externalism maintains that the externalist 'runs out' of reasons in exactly those cases in which additional reasons are required. I have called this the 'problem of explanatory shortfall' (Chang 2013b op. cit.). In a hard case between A and B, glossed as one in which A is better than B in some respects, B is better than A in other respects, and yet *it reasonably seems* that neither is at least as good as the other overall (perhaps with respect to a covering consideration), it could nevertheless be the case that A is better than B overall, even though one is passionate about B. In such cases, such as those between careers or life-long pursuits, the externalist explanation of why A is better than B must bottom out 'those are just the normative facts'. If we accept that there are reasons beyond given externalist ones, we have resources for explaining why A is better than B that appeal to the agent's will.
- 54 Chang 2009, 2013a, b, 2016b, 2020 op. cit.
- 55 Of course, the question of what grounds given reasons remains, but when put in a larger context in which will-based reasons play a central role in understanding normativity, the debate shifts. For example, is some of the motivation for thinking that given reasons are desire based – e.g. that an account of the ground of reasons must make *fundamental* room for the agent's perhaps peculiar motivational capacities – covered by the existence of will-based reasons?
- 56 I have suggested that there are normative constraints on will-based reasons: if, according to one's given reasons A is better (or worse) than B (simpliciter or with respect to a covering consideration) then will-based reasons cannot change this 'valence'. Put another way, when given reasons deliver the result that one item is better or worse than another, it follows that, 'all thing considered', that is, considering both given and will-based reasons, the same relation will hold. But if, according to one's given reasons, A and B stand in any other relation, what I call 'equipose', which covers their being equally good, on a par, incomparable, then in principle will-based reasons can determine what is true of A and B, 'all things considered', that is, taking into account the relevant given and will-based reasons. As a substantive matter, I believe that, although we can create reasons whenever given reasons are incommensurable but comparable, will-based reasons can only change the all-things-considered truth about A and B if A and B are on a par. This is because the kind of normative power

- involved in creating reasons does not, to my mind, plausibly sit well with situations in which, with respect to one's given reasons, it is intrinsically acceptable to flip a coin (the case of being equally good) or the options are outside the realm of reason (the case of being incomparable). This guarantees that we have no extensional difficulties beyond those dogging the usual accounts of given reasons. If, as I suggest, the right hybrid form of theory combines value-based and will-based reasons, we also have the resources to solve the extension problem raised for externalism, viz., the problem of explanatory shortfall according to which will-based reasons provide further reasons where value-based reasons run out.
- 57 Chang 2015b, 2016b op. cit.
- 58 Raz, Joseph. 1990. *Practical Reason and Norms*. Princeton, NJ: Princeton University Press; Wolf, Susan. 1990. *Freedom Within Reason*. Oxford: Oxford University Press.
- 59 It is perhaps worth noting that there is a loose sense in which even on the passivist view we can 'create' reasons. When I promise to wash your car tomorrow, I 'create' an obligation for myself to wash your car tomorrow, *ceteris paribus*. But as promises are typically understood, this is an anemic sort of creation; I simply trigger a pre-existing perhaps conditional reason to wash your car grounded in the value of the promising institution. I discuss the distinction between these 'weak' and 'robust' normative powers in Chang 2020 op. cit.
- 60 Callahan, Laura. 2021. "Epistemic Existentialism". *Episteme* 18(4): 539–54.
- 61 Arneson, Richard. 2002. "Egalitarianism." In *The Stanford Encyclopedia of Philosophy (Summer 2013 Edition)*, edited by E.N. Zalta. <https://plato.stanford.edu/entries/egalitarianism>. (accessed August 1, 2022).
- 62 A possibly equivalent way to describe expansive egalitarianism (or paritarianism) is to say that equality itself is what Rawls called a 'range property' (see Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Belknap Press, p. 508). Rawls was concerned to argue that the natural properties that provided the *basis* for equality were 'range properties', that is, they were properties with underlying differences in degree that were irrelevant to meeting a certain threshold for the range property to hold. Jeremy Waldron provides useful discussion (see Waldron, Jeremy. 2016. "Looking for a Range Property: Hobbes, Kant, and Rawls." Lecture 3 of the Gifford Lectures, University of Edinburgh. <https://www.youtube.com/watch?v=azKDbVGW11c>). For Rawls, the capacity to have a moral personality (i.e. a sense of one's good and a sense of justice) is the basis for equality among persons, but of course people can have this capacity to different degrees. The property of having the capacity for a moral personality, then, is a threshold property with respect to which different degrees of capacity for a moral personality might be said to be on a par. Paritarianism might be understood as the view that the property of being equal in some respect is itself a range property, where qualitatively different ways of being equal can all be on a par. Parity would then provide a more precise way of understanding such range properties. (accessed August 1, 2022).
- 63 Hart, H.L.A. 1961. *The Concept of Law*. Oxford: Oxford University Press; Dworkin, Ronald. 1986. *Law's Empire*. Cambridge, MA: Belknap Press. (accessed August 1, 2022).
- 64 Mayer, Colin. 2018. *Prosperity: Better Business Makes the Greater Good*. Oxford: Oxford University Press.
- 65 Klarman, Herbert, John Francis, and Gerald Rosenthal. 1968. "Cost Effectiveness Analysis Applied to the Treatment of Chronic Renal Disease." *Medical Care* 6(1): 48–54; Bush, J.W., S. Fanshel, and M. Chen. 1972. "Analysis of Tuberculin Testing Program Using a Health Status Index." *Socio-Economic Planning Sciences* 6(1): 49–69; for a history of the notion, see MacKillop, Eleanor, and Sally Sheard. 2018. "Quantifying Life: Understanding the History of Quality-Adjusted Life-Years (QALYs)." *Social Science and Medicine* 211: 359–66.
- 66 Herlitz, Anders. 2018. "Committing to Priorities: Incompleteness in Macro-Level Health Care Allocation and Its Implications." *The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine* 43(6): 724–45.
- 67 Railton, Peter. 2022. *Uheiro Lectures, "Ethics and AI"*. Oxford: University of Oxford.
- 68 Eckersley, Peter. 2019. "Impossibility and Uncertainty Theorems in AI Value Alignment, or Why Your AGI Should Not Have a Utility Function." arXiv:1901.00064 [cs.AI] <https://doi.org/10.48550/arXiv.1901.00064>; Christian, Brian. 2020. *The Alignment Problem: Machine Learning and Human Values*. New York: Norton; Goodman, Bryce. 2021. "Hard Choices and Hard Limits in Artificial Intelligence." AIES 2021: Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, 112–21. <https://doi.org/10.1145/3461702.3462539>; Chang, Ruth. 2021d. "How to Prevent AI from Taking Over the World." *New Statesman*. (accessed February 16, 2021). <https://www.newstatesman.com/science-tech/2021/02/how-prevent-ai-taking-over-world>.
- 69 Dobbe, Roel I.J., Thomas Krendl Gilbert, and Yonatan Mintz. 2020. "Hard Choices in Artificial Intelligence: Addressing Normative Uncertainty Through Sociotechnical Commitment." AIES 2020: Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, <https://doi.org/10.1145/3375627.3375861>, 242.