

## THE AMBITIONS OF CONSEQUENTIALISM

*Brian McElwee*

A CONSEQUENTIALIST ACCOUNT of some subject gives an assessment of it by reference to the production of good outcomes. This characterization leaves open (at least) four issues:

1. What is the subject being evaluated, e.g., acts, rules, motives, character traits?
2. Which components of an outcome contribute to making that outcome good, e.g., the promotion of well-being, the promotion of the glory of god, the universe being well ordered?
3. What distribution of those components makes for the best outcome, e.g., supposing well-being to be the sole relevant type of component, we can ask: Is the best outcome one in which overall well-being is maximized? Or may one outcome be better than another, even though it contains less well-being, so long as well-being is more equally distributed among subjects, or when the least well off are prioritized, or when all have enough?
4. How do we derive our evaluation of the subject from its relation to good outcomes, e.g., if the subject being assessed is acts, do we evaluate acts directly by reference to the extent to which they produce good outcomes? Or in some indirect way, such as their being compliant with a code of rules whose general acceptance would produce good outcomes?

In this paper, my central focus is on the first issue: What subject matter should consequentialists aim to address? Regarding the second and third issues, I intend what I say to be neutral across a wide variety of theories of what makes for good outcomes. Regarding the fourth issue, my focus will be on direct derivations of evaluations from good outcomes. However, indirect approaches such as rule consequentialism will figure briefly in our discussion.

## 1. CONSEQUENTIALISM AND RIGHT ACTION

The most well-known version of consequentialism is act utilitarianism. Its subject of evaluation is *acts*: it claims that the right act is the one that produces the best consequences. Its account of what makes outcomes good is *welfarist*: it claims that the best outcomes are those in which well-being or happiness is promoted. Its preferred distributive principle is *maximization*: the best outcome is the one in which well-being is maximized, no matter how that well-being is distributed across subjects. Finally, its evaluation of acts is *direct*: it assesses acts directly in terms of the goodness of their consequences (relative to the goodness of the consequences of the available alternative acts).

Let us begin then by considering consequentialist theories that, like act utilitarianism, are theories of right action, or of what we ought to do.<sup>1</sup> Such theories have significant initial plausibility, given that they embody the appealing thought that we should try to bring about what is good and try to eliminate what is bad. However, it is important to notice that talk of right action, or of what we ought to do, may be ambiguous. Consider the following two construals of the act consequentialist claim:

1. We are *morally obliged* to do whatever brings about the best (expected) consequences.<sup>2</sup> All other available actions are *morally wrong*.
2. What there is *most reason* for us to do is whatever brings about the best (expected) consequences. All reasons for action are grounded in production of the good.<sup>3</sup>

Formulation 1 faces very strong demandingness objections. It implies that, given the state of the world, where there is so much preventable suffering, we are morally obliged to devote almost all of our spare time, money, and energies to help-

- 1 A *fully fledged* act consequentialism perhaps is further committed to the view that what *makes* actions right is that they produce the best consequences. This stands in contrast, for example, to a divine-command theory that says that what makes actions right is that God commands them, but that God always commands those actions that produce the best consequences. See Heydt ("Utilitarianism before Bentham") for the historical importance of such versions of consequentialism.
- 2 Whether we characterize a consequentialist claim in terms of the best consequences or the best expected consequences may be seen as a fifth relevant issue. For the purposes of this paper, I lay this question aside.
- 3 It is quite open to a theorist to assert the first claim without the second—to say that reasons for action can have other sources besides promotion of the good, but that the fact that some action will best promote the good is always a trumping factor. I focus here on the stronger view, which makes both claims.

ing strangers in need.<sup>4</sup> If we do anything less than the most we can to help, we are acting morally wrongly. There are several aspects to such demandingness objections, each of which leaves this extremist form of act consequentialism looking seemingly untenable. First, it leaves us with very few morally permissible options; at every turn, we may only pursue one of those courses of action that will expectably bring about the most good. The options available will be generally unattractive, requiring me to forgo many of my most treasured projects, pastimes, and even relationships. Most of the time, a much better bet in maximally promoting the impartial good will rather involve working for effective charities, campaigning for trade justice, and so on. A second aspect of the extremist view's apparent implausibility is its condemnation of apparently morally admirable behavior as morally wrong. Someone who devotes a very substantial amount of her spare time and money to good causes, but who falls short of doing the most she could do is, by the lights of this view, failing in her moral obligations. A related unintuitive implication is that the view seems to leave no room for supererogatory action—action that is morally good, but that goes beyond what duty or obligation requires. Finally, the view suggests that we are only allowed to be partial toward our loved ones to a very limited degree—just to that degree that is mandated by an impartial calculus.

We should note that formulation 2 does not face the same objections. Standard demandingness objections apply specifically to accounts of moral obligation, not to mere rankings of the choiceworthiness of actions. To say that someone acts morally wrongly, that she has failed in her moral obligations, is not simply to say that she has done something other than the very best thing she could have done by the lights of morality, but to add a positive *criticism* of her acting as she does.

One common way of construing what is distinctive about judgments of moral obligation and moral wrongness is in terms of the distinctive sanctions of morality—those of the paradigmatic moral sentiments of blame and guilt. It is this feature that gives demanding theories of moral obligation their capacity to disturb—such theories imply that we must give up many of our treasured pastimes, projects, and relationships for the sake of aiding strangers, or else be deserving of distinctively painful feelings of blame and guilt.<sup>5</sup>

- 4 This conclusion is contingent, of course, on what account of the good is proposed. On any impartial account of the good that has a commonsensical conception of welfare, where intense suffering is counted as a significant bad, the conclusion seems unavoidable.
- 5 See Gibbard (*Wise Choices, Apt Feelings*, esp. ch. 3) and Skorupski (*The Domain of Reasons*, esp. ch. 12) for development and defense of the view that moral wrongness is centered around the sentiments of blame and guilt. Just how tight the connection is between moral wrongness and blameworthiness is a matter of dispute. In assuming in this paper that there

Formulation 2 has no such implications. It says simply that the action that has the best consequences is the one that has most to be said in its favor, without implying that if one does anything else, one merits criticism or moral sanction.<sup>6</sup> It might be pressed that an analogous demandingness objection applies to formulation 2, namely that it implies that we must live a very self-sacrificial lifestyle on pain of meriting charges of *irrationality*. But this charge is under-motivated. Formulation 2 does indeed imply, given a reasonable account of the good, that unless we do the most we can to help those in need, we are falling short of doing the very best we can by the lights of reason. But this is a charge we can happily live with. The life of the moral saint may be one that has the most to be said for it overall, the one that gets the strongest endorsement from the standpoint of reason. But one does not merit positive charges of irrationality, in the ordinary, rhetorically loaded sense of that term, simply for acting otherwise.<sup>7</sup> One's behavior may be perfectly *understandable* (and thus not *irrational*) if grounded in good reasons, even when those reasons are less strong than reasons to do something that will be difficult or costly to one's own well-being.

Our first conclusion then is that if consequentialism is to be understood as

---

is indeed a significant connection between the two, I am *not* assuming the very controversial view that there can never be a gap between the wrong and the blameworthy. For instance, we might confidently judge some action (say, insulting a colleague) to be morally wrong, but think the agent is not blameworthy because of some extenuating circumstance (she had just received extremely upsetting news). In another type of case, we might judge that some action is "objectively" wrong, but the agent is not blameworthy because she is in no epistemic position to know the features of the action that make it a bad idea (for example, she is in no position to know that putting the plug in this socket will result in someone getting a nasty electric shock). In both these sorts of cases, even though moral wrongness and actual blameworthiness come apart, it still seems natural to think there is a close connection between the judgment that the act is wrong and *hypothetical* blameworthiness. In the first case, part of what is meant by saying the action is wrong is that it is the sort of thing that one would merit feelings of blame for doing *absent* extenuating circumstances. In the second case, in saying that the action is "objectively" wrong, we are committing to the idea that one would merit feelings of blame if one was fully informed of the relevant facts and yet performed that action.

- 6 It is perhaps more common in everyday talk to speak simply of what we should do, rather than of what there is most reason to do. I deploy the latter formulation simply to avoid the implications of meriting blame or serious criticism that can come with "should" and more frequently with "ought" in certain contexts. It is noteworthy that when we evaluate possible courses of action, we very frequently use very close synonyms for "reason" and "most reason," e.g., "There's something to be said for staying a bit longer, but I think there's most to be said for leaving now." Furthermore, in the case of ordinary normative talk about beliefs and feelings, which I discuss below, we tend to use the language of "reason" and "most reason" very frequently.

7 For further discussion, see McElwee, "Impartial Reasons, Moral Demands."

a direct theory of right action, or of what we ought to do, then this is best construed as a claim about what we have most reason to do, not as a theory of our moral obligations.<sup>8</sup>

## 2. CONSEQUENTIALISM AS A GENERAL THEORY OF NORMATIVITY

Consequentialism may be best known as a theory assessing actions. But why focus solely on actions? Other things besides actions require ethical evaluation. And other things besides actions can have good or bad consequences. Consequentialists have thus been led to make distinctively consequentialist assessments of many other things besides acts—in particular, motives, character traits, moral rules, moral codes, and sets of institutions. As Jeremy Bentham says, “It is with disposition [of character] as with everything else: it will be good or bad according to its effects: according to the effects it has in augmenting or diminishing the happiness of the community.”<sup>9</sup>

The idea that everything should be assessed in consequentialist terms has come to be known as global consequentialism, a term coined by Philip Pettit and Michael Smith:

Global consequentialism identifies the right  $x$ , for any  $x$  in the category of evaluands—be the evaluands acts, motives, rules, or whatever—as the best  $x$ , where the best  $x$ , in turn is that which maximises value.<sup>10</sup>

Similarly, Shelly Kagan writes that the “most plausible version of consequentialism will be direct with regard to everything.”<sup>11</sup>

Derek Parfit makes explicit that consequentialism is to cover all possible evaluands:

Consequentialism covers, not just acts and outcomes, but also desires, dispositions, beliefs, emotions, the colour of our eyes, the climate, and everything else. More exactly,  $C$  covers anything that could make outcomes

8 A useful comparison here is with the “scalar” consequentialism defended in Norcross, “The Scalar Approach to Utilitarianism,” and Crisp, *Reasons and the Good*. Such a view endorses a consequentialist ranking of actions from best to worst (or from what there is most reason to do to what there is least reason to do). But what is distinctive about the scalar view is not its consequentialist ranking, but its rejection of any supplementary account of moral obligation. I argue against such a rejection in McElwee, “Consequentialism and Permissibility” and “Should We De-Moralize Ethical Theory?”

9 Bentham, *An Introduction to the Principles of Morals and Legislation*, 246.

10 Pettit and Smith, “Global Consequentialism,” 121.

11 Kagan, “Evaluative Focal Points,” 151.

better or worse. According to C, the best possible climate is the one that would make outcomes best.<sup>12</sup>

What attractions might there be in adopting such a global consequentialism? Julia Driver notes that one prominent rationale for the view comes in response to the virtue ethical challenge to consequentialism.<sup>13</sup> Consequentialists, along with other modern moral theorists, are accused of giving exclusive attention to the evaluation of action, particularly couched in terms of the morally obligatory, at the expense of attention to motives and character. One consequentialist response to this is to assert that consequentialism has a ready-made way of evaluating motives and character traits; this can be done in just the same way as actions are to be evaluated, directly in terms of the consequences they (tend to) produce. Another attraction Driver claims for global consequentialism is that it gives an attractive account of what she calls “normative ambivalence,” where “a stable evaluation, or a unitary evaluation, is hard to achieve because we are really thinking about two different things: [for example] the agent’s action and the character the agent is expressing through the action.”<sup>14</sup> As Robert Adams notes, the consequentially optimal set of motives for an agent to have may lead that agent on occasion to act in a consequentially suboptimal way.<sup>15</sup> In such cases, we may feel a tension in our normative assessments, at the same time wanting to endorse the agent’s behaving as she does and yet wanting to criticize it. But this tension, Driver suggests, can be relieved simply by making a twofold claim, offering a direct consequentialist account of the action (it is wrong because it has overall bad or suboptimal consequences) and another direct consequentialist account of the set of motives in play (it is right because it has better overall consequences than any other set of motives, even if in this case it leads to wrong actions). Global consequentialism thus allows us to say all we want to say.

The focus of Pettit and Smith’s discussion is to argue that this global consequentialism is preferable to “local” forms of consequentialism. Local forms of consequentialism give a direct consequentialist account of a privileged evaluand, and then give an assessment of other evaluands by reference to some relation they stand in to the privileged one. Rule consequentialism is an example of this local consequentialist pattern: its privileged evaluand is sets of moral rules. The

12 Parfit, *Reasons and Persons*, 25.

13 Driver, “Global Utilitarianism.”

14 Driver, “Global Utilitarianism,” 169.

15 Adams, “Motive Utilitarianism.”

right set of moral rules is that whose general acceptance produces the most good. A right action is one that complies with this right set of rules.<sup>16</sup>

Pettit and Smith argue that global consequentialism is to be preferred to *all* forms of local consequentialism. After all, why should we privilege any particular evaluand? Local consequentialisms will unduly ignore those consequences of each evaluand that are not mediated by the privileged evaluand. We should, for example, reject a local consequentialism that defines the right act as one that has the best consequences, but then defines the right motives or rules as those that lead to the promotion of right acts, on the following basis:

Someone's possession of certain motives, or his or her having internalised certain rules, may have consequences that are not mediated by any act to which those motives or rules give rise. Your clear benevolence towards me, and mine towards you, can provide each of us with a sense of warmth and reassurance independently of any acts that it occasions. And the mere knowledge that you have internalised a rule of promise-keeping provides me, well in advance of any contract we enter into, with a rich sense of the arrangements we may form.<sup>17</sup>

However, in order to establish that global consequentialism is the most plausible version of consequentialism, it is not enough to establish that it is more plausible than local consequentialisms that give a direct consequentialist treatment of a privileged evaluand and then supplement that with accounts of other evaluands that are derived from the privileged one. One would also need to show that global consequentialism is to be preferred to theories that are directly consequentialist about actions (or some other particular evaluand), and then complement that claim with a *non*-consequentialist account of other evaluands—i.e., an account that is neither directly consequentialist, nor one that assesses these other evaluands by reference to some relation they stand in to the privileged evaluand.

In fact, global consequentialism comes off badly when compared to some such “hybrid” views. I will put the point first in terms of reasons, before going on to address global consequentialists on their own terms, in section 4. I have suggested that consequentialism has some attraction as a theory of what we ought to do, specifically when interpreted as the claim that what there is *most reason* to do is what brings about the best consequences. We can apply the global consequentialist's challenge—“Why privilege actions?”—to this formulation. If

<sup>16</sup> For detailed discussion of different possible formulations of rule consequentialism, see Hooker, *Ideal Code, Real World*.

<sup>17</sup> Pettit and Smith, “Global Consequentialism,” 122.

consequentialism is a plausible theory of reasons for action, might it not equally be a plausible theory of reasons for other reason-responsive states—of reasons for belief, and of reasons to feel?

Let us take first the claim as applied to beliefs:

There is most reason to believe whatever belief will bring about the best consequences. There is reason to believe *B* just to the extent that believing *B* will bring about some good.

This consequentialist account of reasons to believe looks very unpromising.<sup>18</sup> Consider a pair of detectives investigating a murder. In working out whether there is reason to believe Jones committed the crime, they should (and detectives typically do) reflect on the evidence available to them. Is there evidence that Jones was at the scene of the crime at the time the crime was committed? Or evidence that he was somewhere else? Is there evidence that Jones had some motive to commit the crime? Or evidence that he had no such motive?

In working out whether there is reason to believe that Jones committed the crime, they should not (and detectives typically do not) start thinking about the goodness of the consequences of their having various beliefs about whether Jones committed the crime. Suppose one detective says, “We’ve already got Jones locked up in a cell. If I were to believe that Jones committed the crime, I’d be able to sleep better tonight, believing that the killer is behind bars. So there is good reason to believe Jones committed the murder.” This would seem patently absurd—this is not the way our ordinary talk of reasons to believe operates.<sup>19</sup> Rather, reasons to believe are based on evidence. A direct consequentialist account of reasons to believe, evaluating beliefs in light of the same goods used to evaluate actions, seems implausible.<sup>20</sup>

Might we not be more tempted in some other cases to say that the conse-

18 At least if it simply applies the same account of the good being used to assess reasons to act. Perhaps a more plausible account of reasons to believe that may yet be described as consequentialist is one in which there is a *sui generis* good in light of which we should evaluate beliefs, distinct from the good in light of which we should evaluate actions (e.g. well-being). For example, it might be suggested that the belief there is most reason to believe is the one that maximizes overall true belief, or significant true beliefs, or the ratio of true beliefs to false ones. This sort of view would be quite different from a unified consequentialism, with a single account of the good, of the kind that global consequentialist writers clearly have in mind, and I do not aim to make any assessment of it in this paper. See Dunn, “Epistemic Consequentialism.”

19 Note that the absurdity does not lie in this being a suboptimal “decision procedure” for establishing what there is reason to believe. Rather, it lies in a misidentification of what the reason-making factors are.

20 Chappell (“Fittingness”) similarly observes that the case of beliefs (and that of feelings)

quences of believing should bear on what we ought to believe? An opponent of the view under discussion might respond by saying, "Of course the consequences of having beliefs matters. Believing (unwarrantedly) that you are the best man for the job can help you get the job. So you ought to try to cultivate such a belief."

Insofar as we think reasons for *action* are provided by the consequences of our actions, that final claim about what we ought to *do* may be absolutely correct. Our beliefs are not generally under the immediate control of our will, but there are courses of action we can take in order to cultivate certain sorts of beliefs and to suppress others. When it is possible for us, at sufficiently little cost, to cultivate beliefs that will produce good consequences, we plausibly have reason to do so; there are frequently reasons to *bring it about* that I have certain beliefs. But these reasons are reasons for *action*.

The natural thing to say about such cases, I think, is that there is no reason to believe that I am the best man for the job (if all my evidence suggests otherwise) but that there may be practical reason to try to bring it about that I believe I am, if doing so is really going to bring about some good consequence. Some may prefer to talk instead about *two types of reasons to believe*: reasons of evidential warrant and consequence-based reasons. Perhaps not that much hangs on whether we call the latter as well as the former reasons for belief. What is more important is that there seems to be no intelligent composite question of what I have most reason to believe overall, which somehow combines reasons of the two kinds. Instead, there is the practical question of what beliefs I ought to try to cultivate, and there is the distinct question of what beliefs are epistemically warranted.

In any case, what is relevant for present purposes is the relatively weak claim that *not all* reasons for belief are grounded in the consequences of those beliefs. A *purely* consequentialist account of reasons to believe seems clearly implausible.

Why is this important? Whoever thought that consequentialism was supposed to provide an account of reasons to believe? Well, first we seem to have established that some evaluands lend themselves more readily to a consequentialist treatment than others—*contra* the global consequentialist's claim. Notwithstanding the fact that beliefs can have good and bad consequences, such features clearly do not exhaust the considerations bearing on their normative status. The import becomes much more significant, however, when we turn to the third category of reason-responsive states—namely, feelings.

Consider the claims:

---

looks especially jarring for global consequentialists in their attempt to extend consequentialism beyond its proper domain of action.

There is most reason to feel whatever feeling will bring about the best consequences. There is reason to feel *F* just to the extent that feeling *F* will bring about some good.

A consequentialist account of reasons to feel looks just as unattractive as a direct consequentialist account of reasons to believe. Natural, appealing claims about reasons to feel, making no reference to the consequences of having the feeling, abound:

There is reason to feel sad when you are bereaved.

There is reason to feel grateful when someone does you a good turn.

There is reason to feel disappointed when you lose a cup semifinal.

There is reason to feel relief when you are cured of a dangerous illness.

We do not need to know that feeling a certain way will have good consequences in order to know what there is reason to feel. Reasons to feel appear to be determined by the fittingness of the feeling to the object of the feeling, rather than by the good or bad consequences of having the feeling.<sup>21</sup>

Again, an opponent might respond at this point by saying, "Of course the consequences of having feelings matters. Some feelings are very harmful. Feelings of envy, for instance, can eat away at you and cause misery to you and your loved ones. We ought to try to eliminate such feelings." And again, if we think reasons for *action* are provided by the consequences of our actions, the claim about what we ought to *do* may be quite correct. As in the case of belief, there are courses of action we can take in order to cultivate certain types of emotional responses and to suppress others. When we are able to costlessly cultivate feelings that will produce good consequences, we have some reason to do so. There are *practical* reasons to bring it about that I feel some way or another.

This sort of distinction has figured in discussions of fitting attitude analyses of value, under the heading of the "wrong kind of reasons." Just because it will have good consequences if I feel admiration for some cruel and powerful tyrant, this does not mean that the tyrant is admirable. The most natural-sounding thing to say, I think, is that there is no reason to admire the tyrant (he is not admirable) but that there is practical reason to try to bring it about that I admire him, if doing so is, for example, going to spare me a painful death at the hands of

21 Again, it might be suggested, in a "consequentializing" spirit, that there is a *sui generis* good in light of which we should evaluate feelings, distinct from the good appropriate for assessing acts. For example, we have most reason to feel whatever feeling best maximizes overall fittingness between one's feelings and their objects. I do not intend to evaluate this sort of view in this paper.

his henchmen.<sup>22</sup> Instead of talking about *two types of reasons to feel* (reasons of fittingness and consequence-based reasons), it seems more natural to talk about (fitting) reasons to feel, and practical reasons to cultivate feelings. But again, it is more important to recognize that there is no composite question of what I have most reason to feel overall, combining reasons of the two kinds. Instead, there is simply the practical question of what feelings I ought to try to cultivate, and a distinct question of what feelings are fitting, or merited, or apt, or appropriate.

As before, what is most relevant here is the weaker claim that *not all* reasons for feelings are grounded in the consequences of those feelings. A *purely* consequentialist account of reasons to feel seems clearly implausible. Our second conclusion, then, is that consequentialism is *not* a plausible theory of reasons in general.

### 3. REASONS TO FEEL AND ETHICAL EVALUATION

The correct treatment of reasons to feel has wider importance than is generally appreciated. Many of our ethical (and more generally evaluative) questions are, in significant part, matters of what there is reason to feel.

#### 3.1. Moral Obligation

Take the case of moral obligation, already addressed briefly in section 1. Consequentialists frequently claim that we are morally obliged to do whatever brings about the best (expected) consequences. We noted that such claims face strong demandingness objections, on any substantive account of the morally obligatory, which understands this as implying more than that an action is morally best. Why does it seem so objectionable to say that we are *morally obliged* to devote our lives to helping distant strangers, that we would be acting *morally wrongly* if we did otherwise? Its objectionableness is plausibly explained by the sentimental core of charges of moral wrongness and of violating moral obligations. When a moral theory claims merely that a life of extreme altruism would be *morally best*, we do not typically reply, “That’s too demanding to be morally best.” But when a moral theory says that the life of extreme altruism is *morally required* or *obligatory*, we confidently judge that it is too demanding. What is added by claims about what is morally obligatory is the imputation of blame or guilt for failure to act in the recommended way. Again, this is why the extreme act consequentialist claim is so unsettling—it says we must give up our relatively comfortable lifestyles or else be such as to merit the sanctions of the moral sentiments.

Consider once more the agent who does a very substantial amount to

22 Rabinowicz and Rønnow-Rasmussen, “The Strike of the Demon.”

help needy strangers—devoting perhaps a third of his spare income and three nights a week of his spare time to helping the poor. It may be true of him that he could be doing even better, but it seems absurd to accuse him of acting in a way that is morally wrong. The reason why it sticks in the throat to say that he is acting morally wrongly, that he has failed in his moral obligations, is that he seems clearly not to merit feelings of blame for the level of altruism embodied in his behavior. His altruistic efforts in fact seem to merit admiration. If this line of thinking is correct, then our best account of our moral obligations will be shaped by our best judgments about when there is reason to feel blame toward an agent. It will not be a maximizing account like the extremist version of act consequentialism.

An alternative consequentialist account of moral obligation might endorse the connection between moral obligation and blameworthiness, but then go on to offer a distinctively consequentialist account of norms for blaming. One way to do this would be to give a direct consequentialist account of when we have most reason to perform distinctive blame *actions*, such as criticizing, remonstrating, shunning, or formally punishing. (Such an account will be of a piece with a more general direct consequentialist account of reasons for action.) But this is very jarring as an account of blameworthiness. Even if we accept the (controversial) view that we should perform such blame actions just when they will produce the best consequences, our judgments of blameworthiness seem tied not to the expediency of such blame actions, but instead to when there is reason to *feel* blame *feelings* toward the agent.<sup>23</sup> And, as argued in section 2, reasons to feel are more plausibly treated according to fittingness considerations, rather than in terms of the consequences of having the feelings in question.<sup>24</sup>

### 3.2. *Virtue*

A second example of a central ethical concept that must plausibly take account of reasons to feel is virtue. Julia Driver, following Bentham and Mill, argues that virtues are character traits that are systematically instrumental in promoting

23 And so there may be (consequence-based) reason to shun or punish someone who is not blameworthy, and there may be (consequence-based) reason to forgo shunning or punishing someone who is blameworthy.

24 It is beyond the scope of this paper to give an account of when it is fitting to feel blame toward an agent for acting as she did. But such an account will generally need to give significant attention to the value of the consequences of the act being assessed. All else being equal, one is frequently blameworthy for performing actions with bad consequences, and not generally blameworthy for performing actions with good consequences.

good consequences.<sup>25</sup> But this view fits awkwardly with our intuitive judgments about virtue.

Virtues do generally produce good consequences in the actual world, so the clearest-cut examples where this instrumentalist treatment seems wrong may be rather unusual. But any correct account of what makes for virtue must cover unusual counterfactual cases. Suppose a powerful demon made it the case that malice systematically produced good consequences. Would this really suffice to make malice—*wishing suffering upon others*—virtuous? This rings false. A plausible explanation why is that our judgments of virtue are in large part governed by judgments about reasons to feel—in this case, about what traits there is reason to admire. Those traits that we acclaim as virtues are not necessarily those that have the best consequences, but those that we have reason to morally admire. We should not regard malice as a virtue even in these strange circumstances, because malice is not admirable—there is no good reason to feel admiration for malicious people. Rather, the circumstances are merely ones where, unusually, we have strong *practical* reason to cultivate the vice of malice.

More homely examples make a similar point, though perhaps less starkly. Consider the disposition to obey the law from fear of punishment. If this trait is likely to be more reliable in getting people to obey (good) laws, then it is a disposition we have good practical reason to cultivate, but it rings hollow to call it a virtue because it is not an especially admirable trait to possess. There may be more noble motivations to obey the law, which we do have reason to admire and that we would thus more readily describe as virtuous. But if the former are significantly more reliable than the latter in promoting the good, then we have good reason to cultivate the former in ourselves and others, given the importance of the goods at stake.

On the present proposal, there being good practical reason to cultivate a trait in prevailing circumstances is not sufficient to make the trait a virtue; rather, we only acclaim a trait as a virtue if there is reason to feel admiration toward the agent who has the trait. This account still leaves open what substantive conception of virtue to adopt. But a strong candidate view that fits well with the proposal is the Aristotelian one that virtuous traits are ones that involve correct responsiveness to value—traits that involve “loving the good,” as Thomas Hurka puts it, or “being for the good,” as Robert Adams calls it.<sup>26</sup> We morally admire, and thus call virtuous, those traits that embody such correct responsiveness to value, rather than those that simply lead to good consequences. And such cor-

25 Driver, *Uneasy Virtue* and “The Virtues and Human Nature.” See Skorupski (“Externalism and Self-Governance”) for critical discussion along the lines pursued here.

26 Hurka, *Virtue, Vice and Value*; Adams, *A Theory of Virtue*.

rect responsiveness to value itself partly consists in feeling what there is (fitting) reason to feel: taking pleasure in, for example, the happiness of others and being pained by their suffering.

This very natural view about what underlies common moral assessments goes some way toward undercutting Driver's suggestion that global consequentialism best explains the phenomenon of "normative ambivalence." The sort of case Driver has in mind is one where some character trait, or set of motives, that we would generally endorse leads an agent to perform an action other than the one that we would wish them to perform. So, for example, the compassionate person fails to take the opportunity to kill an evil dictator because she cannot bring herself to ruthlessly strangle him with a shoelace.<sup>27</sup> Driver is surely correct to say that such cases are best treated by giving separate evaluation of the act and of the operative motives or character traits. But even if we endorse a direct consequentialist evaluation of the act, this does not mean that a direct consequentialist evaluation of the motive or trait is most plausible. A deep-seated dislike of violently inflicting pain and suffering is *fitting*, not just instrumentally good. Talk of doing the right thing for the wrong reason or from a bad motive, and of doing the wrong thing for the right reason and from a good motive is a staple of common moral judgment. But talk of good motives is most naturally construed in terms of motives that aim at the good, not in terms of their being instrumental in producing the good.

That some character traits involve feelings and actions that constitute intrinsically fitting responses to value is something that consequentialists really ought to accept but often overlook in offering a purely instrumentalist treatment of virtue. Take the utilitarian axiology that happiness is good and suffering bad. Surely part of what is involved in making this claim is commitment to the idea that it is fitting to desire, to approve, to take pleasure in the production of happiness, and likewise fitting to abhor, to disapprove of, to be pained by suffering. We need not await some further judgment about the consequences of having such attitudes to know that there is something correct about them, that they enjoy some positive normative status. Such responses can themselves be evaluated instrumentally—they can be good as means to securing happiness. But this does not exhaust their normative import. Consequentialists have generally eschewed talk of fitting attitudes, perhaps because it seems to open the door to other sorts of values—"intuitionist" values of the kind that Mill was combating in the nineteenth century. But a consequentialism about reasons for action can fit happily with talk of fitting feelings.<sup>28</sup>

27 Driver, "Global Utilitarianism," 172–73.

28 See Hurka (*Virtue, Vice and Value*) for a form of consequentialism that makes room for ap-

Our third conclusion is that offering plausible accounts of reasons to feel is extremely important because many ethical judgments are partly constituted by judgments about what there is reason to feel.

#### 4. GLOBAL CONSEQUENTIALISM

Let us return now to global consequentialism. As we saw earlier, Pettit and Smith characterize the view in terms of *rightness*:

Global consequentialism identifies the right  $x$ , for any  $x$  in the category of evaluands—be the evaluands acts, motives, rules, or whatever—as the best  $x$ , where the best  $x$ , in turn is that which maximises value.<sup>29</sup>

So, in summary, the right  $x$  is the  $x$  that maximizes value.

It is useful to look at how talk of the “right  $x$ ” figures in ordinary discussion for different instances of  $x$ . Consider again those cases where  $x$  is a reason-responsive state, but not an act:

What’s the right thing to believe about this?

What’s the right way to feel about this?

When phrases like these crop up in everyday talk, they pretty clearly refer to what there is reason (fitting reason) to believe or to feel. It would ordinarily be decidedly odd to start talking about the consequences of having the belief and of having the feeling in response to such questions couched in terms of rightness. The right thing to *do* may in unusual circumstances be to bring about an unfitting belief (e.g., where the evidentially warranted belief is very distressing) or an unfitting feeling (e.g., where having the fitting feeling will lead to my being tortured). But the most natural construal of the question of the “right belief” and the “right feeling” is in terms of, respectively, reasons to believe (in the sense of what is fitting to believe) and reasons to feel (in the sense of what is fitting to feel).

Global consequentialists most often discuss motives, character traits, and codes of moral rules as the  $x$  to be given a consequentialist treatment. But we have already seen that motives and character traits plausibly involve feelings that may be fitting or unfitting, so talking of the “right motive” and the “right charac-

---

appropriate response and characterizes virtue in these terms. See Chappell (“Fittingness”) for excellent discussion of how consequentialists can benefit by making fitting attitudes central to their theory.

29 Pettit and Smith, “Global Consequentialism,” 121.

ter trait” as if these statuses were settled in a purely consequentialist way is again likely to be jarring.

#### 4.1. *Motives*

Consider a scenario where the right thing for Bob to do is to help Andrea. We might ask what is the right motive for Bob to have and to act upon. Compare these two motives:

1. Bob helps Andrea out of compassion.
2. Bob helps Andrea out of a desire to humiliate her lover.

There are possible circumstances in which it would have the best outcome for Bob to act out of the latter motive—perhaps this will have a much higher probability of Bob *successfully* helping Andrea if he acts from this motive, and for once, the malicious motive is unlikely to have significant further effects. But does this really settle the question of whether it is the *right* motive? A relevant normative feature of the two motives seems to have been ignored. On the view outlined above, motives can be intrinsically fitting, involving correct responsiveness to value. In the present example, the first motive involves being pained by something bad, Andrea’s suffering, and desiring something good, the relief of her suffering. The second motive involves taking pleasure in and desiring something bad, the humiliation of Andrea’s lover. Given this, I think we would more naturally describe the first motive as the “right” one, notwithstanding its inferior consequences. “Rightness” talk, like reasons talk, seems to go more readily with fittingness considerations than with consequence-based ones. At the very least, fittingness considerations seem a second relevant normative feature of motives alongside the consequences of having the motive.

#### 4.2. *Character Traits*

As we saw in the discussion of virtue, above, it seems that similar things should be said about character traits. Talk of the “right character traits” could be construed as referring to the character traits we have reason to cultivate (because, on the consequentialist view, their cultivation will best promote the good). But it is at least as naturally interpreted as referring to those traits that involve fitting motives, fitting feelings, fitting responses to value. We might thus be tempted to say that the “right” character trait for Bob to have is compassion, rather than a determined one-upmanship, even when the latter is just as effective, or even more effective, in leading him to act such as to promote the good. The instrumentalist about the evaluation of character traits may respond that compassion is to be preferred to these latter traits because it is in general more likely to lead to the

promotion of the good. But again this seems to leave out something extremely important from our assessment.

#### 4.3. *Codes of Moral Rules*

Similar observations apply to codes of moral rules, when we ask which is the “right” code of moral rules. We should begin by asking what a code of moral rules is. It is, at least in part, a collection of judgments of the form, “One is morally obliged to do  $x$  in circumstances  $C_1$ ,” “One is morally obliged to do  $y$  in circumstances  $C_2$ ,” and so on. But our best judgments about what is morally obligatory, I have suggested, involve judgments about when there is reason to *feel* blame toward agents who fail to comply. So we might naturally construe the phrase “the right code of moral rules” as shorthand for talking about the right way to feel about someone who fails to do  $x$  in circumstances  $C_1$ , the right way to feel about someone who fails to do  $y$  in circumstances  $C_2$ , and so on. And as noted above, these seem to be judgments that most convincingly admit of non-consequentialist treatments.

If we ask the *practical* question about which code of moral rules we should we try to inculcate in ourselves and others, then perhaps in some circumstances we have good reason, grounded in the good consequences of so doing, to inculcate a code that involves treating as morally obligatory something that is not morally obligatory; we should inculcate patterns of feeling that include occasional feelings of blame toward those who do not merit such feelings, and perhaps even *beliefs* that there is reason to feel blame toward someone whom there is no reason to feel blame toward. But saying that this code is the “right” one, just in virtue of its inculcation having the best consequences in certain circumstances, is at best misleading. What is essential is to separate out the practical question, which admits of a plausible consequentialist answer, and the various questions about what there is reason to feel, which do not admit of a plausible consequentialist answer.

So in the global consequentialist’s favorite cases—motives, character traits, moral codes—a direct consequentialist treatment of the “right  $x$ ” seems most plausible insofar as the specific evaluative question being addressed is reducible to a question about what there is reason to do. But even in those cases, the most natural construal of the question about the “right  $x$ ” is not the practical one, but one that involves questions about reasons to feel, about fitting feelings, which do not admit of a plausible consequentialist treatment.<sup>30</sup>

30 What about those evaluands that seem to involve no reason-responsive state whatsoever? Driver notes the sheer oddity of talking of the right eye color or the right climate: “This is odd because these objects are not agents, and we tend to intuitively restrict moral evalua-

Our fourth conclusion then is that the global consequentialist claim is at best misleading. It claims that, for  $x$  in general, the right  $x$  is the one that results in the best outcomes. But this is most plausible when the “right  $x$ ” is construed as the “ $x$  there is reason to act to bring about.”<sup>31</sup> So again, consequentialism appears most plausible simply as a theory of reasons for action.<sup>32</sup>

# 5. THE PRESSURE FROM SENTIMENTALISM

My aim in this paper has not been to offer positive support for consequentialism. Rather, I have tried to establish that a form of consequentialism that limits itself to being a theory of reasons for action, and that complements this with non-consequentialist accounts of other normative questions, is more plausible than one that gives a direct consequentialist account of everything.<sup>33</sup> That is not to say that any form of consequentialism is ultimately convincing. Indeed, one might think that a very plausible competitor view emerges naturally from our discussion—one in which practical reasons have two sources, the good and the sentiments.

I suggested above that in order to know what reasons to feel there are, we need not know the consequences of having those feelings. It is sufficient for knowing that there is (at least some) reason to feel gratitude toward Tom that I know that Tom has done me a good turn. I do not need to investigate whether feeling grateful to Tom will have some good effect. Might we not then be tempt-

---

tion to features relevant to agency. . . . Moral agents are sensitive to reasons; climates are not” (“Global Utilitarianism,” 173). One way of making sense of such judgments of the “right climate” or the “right eye color” is to read them as claims about which is the climate or eye color that it is right to choose when there is reason to choose between climates or between eye colors. Again, this interpretation of the global consequentialist’s claims simply makes them instances of direct act consequentialism, claims specifically about reasons for *action*. For further discussion of such cases, and their relation to the “ought implies can” principle, see Streumer, “Can Consequentialism Cover Everything?”; Brown, “Blameless Wrongdoing and Agglomeration”; and Streumer, “Semi-Global Consequentialism and Blameless Wrongdoing.”

- 31 Perhaps, more precisely, the claim is most plausible when the “right  $x$ ” is construed as the “ $x$  there is most reason to bring about *when possible*,” as there may be cases where the “right  $x$ ” seems one that we have (fitting) reason to will, wish for, or desire—it is the “optimal  $x$ ”—but would be impossible to bring about.
- 32 The arguments offered here buttress those offered in Chappell (“Fittingness”) for a similar conclusion.
- 33 I have offered no explicit argument that such a view is to be preferred to rule consequentialism. Though see Pettit and Smith (“Global Consequentialism”) and Kagan (“Evaluative Focal Points”) for arguments against forms of rule consequentialism that have an ultimately consequentialist justification.

ed to say just the same thing about *actions* that are expressions of reason-supported feelings? Why not say that the fact that Tom has done me a good turn is sufficient to establish that I have (some) reason to *thank* Tom? I need not inquire into whether thanking Tom will have some good consequence to know that there is at least some reason to thank him, even if that reason were to be outweighed by some bad consequences of thanking him. Intuitively, some actions—those that constitute expressions of feelings—can seem fitting in just the same way that feelings can.

We may conclude that just as there is reason to feel blame or resentment toward someone if they cheat you, so there is reason to *remonstrate* with them or to *protest*. Just as there is reason to admire your performance if it shows great skill, so there is reason to *applaud* you.

This idea has been articulated by John Skorupski, who defends what he calls the *bridge principle*:<sup>34</sup>

Whatever facts give  $x$  reason to feel  $F$  give  $x$  reason to do the  $F$ -prompted action, in virtue of being a reason to feel  $F$ .<sup>35</sup>

There is significant appeal to saying that some reasons to perform actions that constitute expressions of feelings are grounded in the fact that there is reason to feel those feelings, while other reasons for action may be based in the good consequences of the action.

One option for the consequentialist is to concede that there are indeed reasons for action grounded in the sentiments, and not in good consequences, but to insist that nonetheless there is always most reason to do that that brings about the most good. On this view, if one knows that some action will bring about the most good, then one can safely conclude that that is what one should do.

The relative pros and cons of these views will need to be addressed in future work.<sup>36</sup> My aim here has simply been to establish what consequentialism is most

34 On Skorupski's view, reasons for action come from *three* distinct sources. Some reasons for action are consequence-based reasons, grounded in the good; others are grounded sentimentally, via the bridge principle; and some are grounded in rights, by the demand principle, which claims that an agent has reason to do that which some person may permissibly demand of him. See Skorupski, *The Domain of Reasons*, pt. III, for a full discussion, and "The Triplism of Practical Reason" for a summary of the view.

35 Skorupski, "The Triplism of Practical Reason," 129.

36 A second important way in which a non-consequentialist account of reasons to feel may put pressure on a purely consequentialist account of reasons for action concerns sentiments of blame. Reasons to feel blame interact directly with reasons to act: a judgment that there is reason to feel blame toward some agent  $A$  appears to presuppose that  $A$  had stronger reason to do other than he did.

plausibly a theory of. A consequentialism more limited in its ambitions, which makes room for non-consequentialist answers to some normative questions, looks much more promising than one that attempts to provide a direct consequentialist answer to every normative question.<sup>37</sup>

University of Southampton  
b.m.mcelwee@soton.ac.uk

#### REFERENCES

- Adams, Robert Merrihew. "Motive Utilitarianism." *Journal of Philosophy* 73, no. 14 (August 1976): 467–81.
- . *A Theory of Virtue: Excellence in Being for the Good*. Oxford: Oxford University Press, 2006.
- Bentham, Jeremy. *An Introduction to the Principles of Morals and Legislation*. Oxford: Blackwell, 1948.
- Brown, Campbell. "Blameless Wrongdoing and Agglomeration: A Response to Streumer." *Utilitas* 17, no. 2 (July 2005): 222–25.
- Chappell, Richard Yetter. "Fittingness: The Sole Normative Primitive." *Philosophical Quarterly* 62, no. 249 (October 2012): 684–704.
- Crisp, Roger. *Reasons and the Good*. Oxford: Oxford University Press, 2006.
- Driver, Julia. "Global Utilitarianism." In *The Cambridge Companion to Utilitarianism*, edited by Ben Eggleston and Dale E. Miller, 166–76. Cambridge: Cambridge University Press, 2014.
- . *Uneasy Virtue*. Cambridge: Cambridge University Press, 2001.
- . "The Virtues and Human Nature." In *How Should One Live? Essays on the Virtues*, edited by Roger Crisp, 111–30. Oxford: Oxford University Press, 1996.
- Dunn, Jeffrey. "Epistemic Consequentialism." *Internet Encyclopedia of Philosophy*. <https://www.iep.utm.edu/epis-con>.
- Gibbard, Allan. *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Oxford: Oxford University Press, 1990.
- Heydt, Colin. "Utilitarianism before Bentham." In *The Cambridge Companion*

37 Early versions of this paper were presented at the University of St. Andrews, and at Senate House, University of London at a symposium in honor of John Skorupski. I am very grateful to the participants at each of these for their feedback. I would also like to thank David Plunkett, Ben Sachs, Justin Snedegar, and Richard Yetter Chappell for very helpful comments on drafts of the paper.

- to *Utilitarianism*, edited by Ben Eggleston and Dale E. Miller, 16–37. Cambridge: Cambridge University Press, 2014.
- Hooker, Brad. *Ideal Code, Real World*, Oxford: Oxford University Press, 2001.
- Hooker, Brad, Elinor Mason, and Dale E. Miller, eds. *Morality, Rules, and Consequences: A Critical Reader*. Edinburgh: Rowman and Littlefield, 2000.
- Hurka, Thomas. *Virtue, Vice and Value*. Oxford: Oxford University Press, 2001.
- Kagan, Shelly. "Evaluative Focal Points." In Hooker, Mason, and Miller, *Morality, Rules, and Consequences*, 134–55.
- McElwee, Brian. "Consequentialism and Permissibility." *Utilitas* 22, no. 2 (June 2010): 171–83.
- . "Impartial Reasons, Moral Demands." *Ethical Theory and Moral Practice* 14, no. 4 (August 2011): 457–66.
- . "Should We De-Moralize Ethical Theory?" *Ratio* 23, no. 3 (September 2010): 308–21.
- Norcross, Alastair. "The Scalar Approach to Utilitarianism." In *The Blackwell Guide to Mill's Utilitarianism*, edited by Henry R. West, 217–32. Oxford: Wiley-Blackwell, 2006.
- Parfit, Derek. *Reasons and Persons*. Oxford: Oxford University Press, 1984.
- Pettit, Philip, and Michael Smith. "Global Consequentialism." In Hooker, Mason, and Miller, *Morality, Rules, and Consequences*, 121–33.
- Rabinowicz, Wlodek, and Toni Rønnow-Rasmussen. "The Strike of the Demon: On Fitting Pro-Attitudes and Value." *Ethics* 114, no. 3 (April 2004): 391–423.
- Skorupski, John. *The Domain of Reasons*. Oxford: Oxford University Press, 2010.
- . "Externalism and Self-Governance." *Utilitas* 16, no. 1 (March 2004): 12–21.
- . "The Triplism of Practical Reason." *Ratio* 25, no. 2 (June 2012): 127–47.
- Streumer, Bart. "Can Consequentialism Cover Everything?" *Utilitas* 15, no. 2 (July 2003): 237–47.
- . "Semi-Global Consequentialism and Blameless Wrongdoing: Reply to Brown." *Utilitas* 17, no. 2 (July 2005): 226–30.