Consequentialism

Douglas W. Portmore

The defining feature of consequentialism is that it ranks outcomes (e.g., those associated with performing certain acts, complying with certain codes of rules, or inculcating certain sets of motives) and then takes the normative statuses of actions to be some increasing function of how those outcomes rank. Little else can be said unequivocally about consequentialism, as consequentialists disagree about most everything else. Consequentialists disagree on whether we should assess the normative statuses of actions directly in terms of how their outcomes rank (act-consequentialism) or indirectly in terms of whether, say, they comply with the code of rules with the highest-ranked associated outcome (rule-consequentialism). They disagree on whether the relevant function is a maximizing one (maximizing consequentialism) or a satisficing one (satisficing consequentialism). And they disagree on whether there is just one ranking of outcomes that is the same for all agents (agent-neutral consequentialism) or a potentially different ranking for each agent (agent-relative consequentialism).

As most see it, consequentialism is a theory about the permissibility of actions, but some hold instead that it is a theory about only the comparative moral value of actions (scalar consequentialism). And whereas some hold that consequentialism is committed to ranking outcomes in terms of their impersonal value, others deny this. Even those who agree that outcomes are to be ranked in terms of their impersonal value disagree about whether outcomes are to be ranked in terms of their actual value (objective consequentialism) or their expected value (subjective consequentialism).

Given all this disagreement, there can be no illuminating set of necessary and sufficient conditions that captures all the various uses of the term 'consequentialism'. Instead, as Walter Sinnott-Armstrong (2003) points out, "the term 'consequentialism' seems to be used as a family resemblance term to refer to any descendant of classic utilitarianism that remains close enough to its ancestor in the important respects." And since philosophers disagree on which are the important respects, there is no agreement on how to define 'consequentialism'. But even if there is no agreement on how to define 'consequentialism', there is, as Sinnott-Armstrong suggests, an archetype: viz., maximizing act-utilitarianism. In this essay, I will start with this archetype, abstract from it the most traditional version of consequentialism, and then explain how various objections to this traditional

version have led philosophers to adopt more nuanced versions of consequentialism.

Maximizing Act-Utilitarianism and Traditional Act-Consequentialism

Maximizing act-utilitarianism is a species of *traditional act-consequentialism*, which holds:

> **TAC:** (1.0) A subject's performing x is morally permissible if and only if, and ultimately because, there is no available alternative, y, whose outcome ranks higher than x's outcome, and (2.0) y's outcome ranks higher than x's outcome if and only if y's outcome is impersonally better than x's outcome.

Or, more simply put, TAC holds:

> A subject's performing x is morally permissible if and only if, and ultimately because, their doing so would maximize the impersonal good.

Two aspects of TAC need clarification. First, note that a subject's performing x would maximize the impersonal good (hereafter, simply 'the good') if and only if there is no available alternative act whose outcome is better than x's outcome. For one, this means that there can be more than one act that would maximize the good, as there can be more than one act that is tied for first place in terms of the production of the good. For another, this means that TAC is not committed to there being a complete ranking of outcomes. Indeed, it may be that some outcomes are incommensurable with others such that they are not better than, worse than, or just as good as the others. Thus, on TAC, an act can be permissible even though there are available alternatives whose outcomes are incommensurable with its outcome. So long as there is no available alternative whose outcome is *better* than its outcome, it will be permissible.

Second, note that, by including the words "and ultimately because," I am acknowledging that someone could fail to be a traditional act-consequentialist (a TACist) even though she accepts the following bi-conditional.

> **BC:** A subject's performing x is morally permissible if and only if their doing so would maximize the good.

Take, for instance, a divine command theorist who accepts not only that a subject's performing x is morally permissible if and only if, and ultimately because, their doing so is not forbidden by God, but also that their performing x is

forbidden by God if and only if their doing so would fail to maximize the good. This divine command theorist would accept BC, but she would not, I presume, count as a TACist, for she would deny that, when acts are permissible, they are so *ultimately because* they maximize the good. To deny this is to deny TAC.

Now, maximizing act-utilitarianism is a species of TAC, for it is committed to both TAC and the following further claim about what is good. That is, *maximizing act-utilitarianism* is committed to the following two claims.

> **MAU** (i) A subject's performing *x* is morally permissible if and only if, and ultimately because, their doing so would maximize the good, and (ii) their performing *x* would maximize the good if and only if their doing so would maximize aggregate utility.

The aggregate utility that would be produced by an act is the sum of all the utility it would produce minus the sum of all the disutility that it would produce, where utility is a measure of whatever it is that enhances a subject's welfare, and disutility is a measure of whatever it is that diminishes a subject's welfare. The classical utilitarians (e.g., Mill 1861 and Bentham 1789) held that pleasure is the only thing that contributes to a subject's welfare and that pain is the only thing that detracts from a subject's welfare. This view about welfare is known as *hedonism*. Hence, the view according to which both hedonism and maximizing act-utilitarianism are true is known as *hedonistic act-utilitarianism* (**HAU**). HAU is perhaps the simplest version of TAC. For this reason, I will often use it to illustrate some of the putative problems with TAC.

Objections to Traditional Act-Consequentialism

TAC amounts to essentially three claims: (I) it is never permissible to fail to maximize the good, (II) it is always permissible to maximize the good, and (III) if a subject's performing *x* is morally permissible, it is ultimately because their doing so would maximize the good. Given these three claims, there are at least three general sorts of objections to TAC, each corresponding to the rejection of one of these three. I consider them below.

First objection: It is sometimes permissible to fail
to maximize the good.

Agent-Sacrificing Options and the Self-Other Asymmetry:

According to TAC, maximizing the good is always obligatory. According to commonsense morality, by contrast, it is often optional. On commonsense morality, there is often the moral option either to act to make things better overall

but worse for oneself (or others) or to act to make things better for oneself (or others) but worse overall. These options, which are known as *agent-centered options*, provide agents with the freedom to give their own interests more or less weight than they have from the impersonal perspective. They come in two varieties: *agent-sacrificing options* and *agent-favoring options*.

Agent-sacrificing options permit agents to perform suboptimal acts so long as these acts are suboptimal only because they involve producing less good for the agent. The underlying thought is that agents are permitted to make things worse for themselves so long as they do not thereby make things worse for others overall.[1] To illustrate, suppose that Abe is stranded on a deserted island. Given his isolation, what Abe does will affect only himself. Now, suppose that he is presently sitting next to an ant hill and that the ants are starting to bite him. Assume that letting the ants continue to bite him is not what is best for him. What would be best for him is his getting up and collecting coconuts. Of course, since he is the only one affected by his actions, his getting up and collecting coconuts is not only what is best for him but also what is best overall. So, on a traditional act-consequentialist theory such as HAU, it is wrong for Abe to let the ants continue to bite him.

This may seem counterintuitive. Although it is clearly stupid, foolish, and irrational for Abe to let the ants continue to bite him, there does not seem to be anything immoral about his doing so. After all, in doing so, he does not hurt anyone but himself. Of course, it might, in certain circumstances, be immoral for Abe to let the ants bite someone else, but that is someone else. On commonsense morality, there seems to be an asymmetry between the self and others. Whereas there is a *pro tanto* obligation to ensure that others enjoy pleasure and avoid pain, there is no such obligation with regard to ourselves.

Given this asymmetry, which is known as the *self-other asymmetry*, we often have the option of sacrificing our own greater utility for sake of promoting others' lesser utility. I can, for instance, permissibly give you my last two aspirins to relieve your slightly less severe headache even if that entails my forgoing the opportunity to relieve my slightly more severe headache. But a traditional act-consequentialist theory, such as HAU, entails that it would be wrong for me to do so. For if my taking the aspirins is what would maximize aggregate utility, then that is what I'm required to do. On HAU, by contrast, there is no self-other asymmetry; there is just as much of an obligation to promote one's own utility as to promote anyone else's. Thus, HAU seems objectionable in that it is unable to accommodate agent-sacrificing options and the self-other asymmetry.

To meet this objection, some have proposed revising HAU. Ted Sider (1993), for instance, has proposed that we instead adopt *self-other utilitarianism:*

> **SOU** (1.0) A subject's performing *x* is morally permissible if and only if, and ultimately because, there is no available alternative, *y*, whose outcome

ranks higher than x's outcome, and (2.1) y's outcome ranks higher than x's outcome if and only if both (a) y's outcome contains more utility for others (i.e., those other than the given subject) than x's outcome and (b) y's outcome contains more overall utility than x's outcome.

In other words, SOU holds that an act is morally permissible so long as there is no available alternative that produces both more utility for others and more overall utility. SOU thereby allows agents to give their own utility anywhere from no weight at all to the same weight that everyone else's utility has from the impersonal perspective.

SOU accounts for both agent-sacrificing options and the self-other asymmetry. For instance, it accounts for our intuition that Abe does no wrong in allowing the ants to bite him given that there is no available alternative whose outcome contains more utility for others. And SOU accounts for our intuition that it would be permissible for me to give my last two aspirins to someone else so that they can relieve their less severe headache. This is permissible, because there is no available alternative whose outcome contains *both* more overall utility and more utility for others. There is only an available alternative whose outcome contains more overall utility: the one in which I take the aspirins. But this alternative produces less utility for others.

Agent-Favoring Options:

Although SOU accounts for both agent-sacrificing options and the self-other asymmetry, it fails to account for agent-favoring options. Agent-favoring options permit agents to favor their own interests and, thus, to promote their own utility even when doing so entails doing less to promote the impersonal good. SOU's failure to accommodate agent-favoring options gives rise to what is known as the *demandingness objection*. Theories such as TAC, HAU, and SOU are too demanding in that they imply that agents are morally required to sacrifice their projects, interests, and special relationships whenever doing so would maximize the good. According to both HAU and SOU, I am morally required to sacrifice my interest in living in material comfort, to neglect my relationship with my daughter, and to abandon my project of writing a book if I could thereby produce more aggregate utility. Suppose, for instance, that what would maximize utility is my taking all my money and moving to some poor developing nation in Africa, founding a school for the education and empowerment of young women. Assume that this is, of all my options, the one that would produce the most utility. Yet, to perform this option, I must give up living in material comfort, abandon my project of writing a book, and desert my daughter, leaving her to become a ward of the state. Theories such as HAU and SOU imply that I am required to make such sacrifices for the sake of maximizing utility. But many philosophers (such as: Stroud 1998, Hurley 2006, and Portmore 2011) think that this is to demand more from me than can be reasonably

demanded of me.

Proponents of utilitarianism might try to respond to this objection by arguing that the problem lies not with HAU (or SOU), but with the sorry state of the world that we live in. That is, they might argue that it is only because we live in a world in which there is so much suffering and hardship that agents like myself are required to make such tremendous sacrifices. But, in fact, this is false. For theories such as HAU and SOU are too demanding even in utopian worlds. To illustrate, imagine that we live in a utopian world in which everyone leads very happy and fulfilled lives. And imagine that my blood contains the makings of some new drug that would make everyone but myself slightly happier. Suppose, though, that to synthesize this drug I would need to sacrifice my life so that scientists could drain out all my blood. Assume that they need all my blood at once to synthesize the drug. And imagine, for the moment, that we can accurately quantify the resultant changes in people's utilities in terms of *utiles*—a measurement of utility equivalent to the utility of a person's experiencing the mildest of pleasures for the briefest of moments. Assume that if this drug were synthesized and evenly distributed among the remaining population of eight billion and one people, each person would get an additional utile. And assume that in laying down my life I would thereby be sacrificing exactly eight billion utiles.

In this case, I could provide a net gain of one utile of aggregate utility by sacrificing my life. And so, on HAU and SOU, I am morally required to do so. But to demand that I make such a tremendous sacrifice for the sake of such a miniscule gain in the total aggregate utility (i.e., one utile) is, it seems, to demand more from me than can be reasonably demanded of me. For it seems that I have at least as much reason, all things considered, to refrain from sacrificing my life in this instance. Perhaps, I have some reason to sacrifice my life for the sake of producing one extra utile of utility, but given that a utile is a rather small measure of utility this will not be a very strong reason. Moreover, it seems that I have a strong agent-relative reason to safeguard my own utility by refusing to sacrifice my life. This reason is an agent-relative reason in that it is a reason that is specific to me, and it goes beyond the agent-neutral reason we each have to promote the total aggregate utility. The thought, then, is that although we each have an agent-neutral moral reason to promote the total aggregate utility, we also each have an additional agent-relative nonmoral reason to promote our own utility. And, given this additional reason, we sometimes have sufficient reason, all things considered, to do something that will promote our own utility at the expense of failing to maximize the total aggregate utility. Indeed, many utilitarians (e.g., Singer 1999, Sidgwick 1907) concede as much.

At this point, such utilitarians argue that although utilitarian views such as HAU and SOU demand that we sacrifice more than we have decisive reason, all things considered, to sacrifice, this is no objection to their view. For they claim that it is just a fact about morality that it demands more from us than we have decisive reason, all things considered, to give. And, so, it is no objection to these

views that they exhibit the unreasonable demandingness that is inherent in morality itself. As David Sobel puts it, "What morality asks, we could say, is too much to be the thing the agent has most reason to do all things considered, but not too much to count as what morality asks" (2007, 14).

Critics of utilitarianism argue, though, that in salvaging utilitarianism from the demandingness objection by claiming that moral requirements are not necessarily anything that we have decisive reason, all things considered, to obey, the utilitarian wins this battle only to lose the war. That is, they succeed in defending the high bar that they set for what morality demands of us only by lowering the bar for what reason demands of us by way of our compliance with morality, thereby implausibly marginalizing the role that morality plays in our practical reason and deliberation (Hurley 2006, 705). And other critics argue that this type of defense is wholly misguided, for morality *is*, by necessity, supreme in its rational authority (Portmore 2011).

There is, however, another way of responding to the demandingness objection. We could revise SOU so that it accommodates agent-favoring options. We can do so by adopting *Schefflerian utilitarianism:*

> **SU** (1.0) A subject's performing x is morally permissible if and only if, and ultimately because, there is no available alternative, y, whose outcome ranks higher than x's outcome, and (2.2) y's outcome ranks higher than x's outcome if and only if both (a) y's outcome contains more utility for others (i.e., those other than the given subject) than x's outcome and (b*) y's outcome contains more egoistically-adjusted utility than x's outcome, where egoistically-adjusted utility includes everyone's utility but adjusts the overall total by giving the subject's utility, say, ten times the weight of anyone else's.

In other words, SU holds that an act is morally permissible so long as there is no available alternative that produces both more utility for others and more egoistically-adjusted utility. SU thereby allows agents to give their own utility anywhere from no weight at all to ten times the weight that their utility has from the impersonal perspective.

Since SU allows agents to give their own utility less weight than it has from an impersonal perspective, SU accommodates agent-sacrificing options just as SOU does. But, unlike SOU, SU can also accommodate agent-favoring options and, as we will see below, supererogatory acts as well. To illustrate, consider the case of Edouard, who has the following four mutually exclusive and jointly exhaustive alternatives: a_1, a_2, a_3, and a_4. (To ensure that they are jointly exhaustive, let a_1 be the act of performing something other than a_2, a_3, or a_4.) The consequences of each of these alternatives are laid out in Table 1.

Table 1[2]

Let S = the subject, $U_s(x)$ = the utility that would accrue to the subject if they were to perform x, let $U_{-s}(x)$ = the utility that would accrue to others if the subject were to perform x, let $U_{+s}(x)$ = $U_{-s}(x)$ + [10 × $U_s(x)$], and let $U(x)$ = the overall utility that would be produced if S were to perform x.

x	$U_s(x)$	$U_{-s}(x)$	$U(x)$	$U_{+s}(x)$	*deontic status*
a_1	7	**10**	17	**80**	merely permissible
a_2	2	**4**	6	**24**	impermissible
a_3	1	**15**	16	**25**	supererogatory[3]
a_4	−1	**20**	19	**10**	supererogatory

Whereas Edouard is, on HAU, obligated to perform a_4, he is, on SU, permitted to perform a_1 instead, which is better for him. Thus, SU permits him to perform a_1 even though his performing a_4 would be both better for others and better in terms of the impersonal good. Edouard has, then, the agent-favoring option of performing a_1 as opposed to a_4. Furthermore, SU accommodates agent-sacrificing options. For Edouard can permissibly choose to perform a_3 as opposed to a_1, thereby providing others with a net benefit of five utiles at a cost of six utiles to himself.

Supererogatory Acts:

Interestingly, SU can also accommodate *supererogatory acts*, acts that go above and beyond the call of duty. This is a plus, as HAU is usually thought to be unable to accommodate supererogatory acts or, at least, unable to accommodate the range of acts that we take to be supererogatory.

An act is supererogatory if and only if it meets the following three conditions: (i) it is morally optional, (ii) it is morally commendatory, and (iii) it goes beyond the call of duty.[4] Clearly, HAU can meet the first two conditions. Some acts will be morally optional on HAU, for sometimes there will be more than one available alternative that would maximize utility. And some of these morally optional alternatives will be morally commendatory. Suppose, for instance, that Smith has the choice of benefiting himself or Jones and that either way utility will be maximized. If Smith chooses to benefit Jones as opposed to himself, this certainly seems to be morally commendatory. Moreover, it is morally optional since it is one of two acts that would maximize utility. So, here is an act that meets the first two conditions for being supererogatory.

What about the third? Does it go beyond the call of duty? Some argue that it does (Harwood 2003, Vessel 2010). They argue that an agent goes beyond the call of duty so long as she does more for others than she is required to do. But we might wonder whether this correctly specifies the relevant sense of 'going

beyond the call of duty'. If we hold that going beyond the call of duty amounts only to doing more for others than is required, then we must deny the possibility of supererogation with respect to self-regarding duties. Yet it certainly seems possible to go beyond what such duties require (Kawall 2003, Portmore 2008). For instance, we might think that there is a duty to develop our talents and that this is an imperfect duty that doesn't require that we take advantage of every opportunity to develop our talents, but requires only that we take advantage of a sufficient number of such opportunities. But if this is right, then it is surely possible to go above and beyond our duty to develop our talents, and yet doing so may be of no benefit to anyone besides ourselves.

Consider also that a supererogatory act must involve doing more of whatever there is more moral reason to do. After all, we would not think that perspiring more than is required is supererogatory, because we do not think that there is any moral reason to perspire more than is required. So, we should think that doing more for others than is required is supererogatory only if there is more moral reason to do more for others than is required. Yet, on HAU, Carlos has no more moral reason to enhance Sergio's utility by some amount than to enhance his own utility by that same amount. All that matters on HAU is that utility is maximized, and utility would be maximized either way.

Even if we allow that an agent goes beyond the call of duty so long as they do more for others than is required and, thus, allow that HAU can accommodate certain supererogatory acts, such as Carlos's choosing to benefit Sergio rather than himself, HAU still cannot accommodate the range of acts that we take to be supererogatory. On HAU, no *optimific act* could ever be supererogatory.[5] Yet, intuitively, many such acts are supererogatory. For instance, Edouard's performing a_4 seems supererogatory. Yet HAU implies that it is obligatory. To illustrate, suppose that Edouard has exactly $3,000 in disposable income each month and, thus, that he can reasonably afford to perform a_1, which involves giving $1,000 a month to charity. By contrast, his performing a_4 involves his giving $3,000 a month to charity, thereby sacrificing everything beyond what is needed for his continued subsistence. Now, his making such a sacrifice would, we will suppose, be optimific in that his donating all of his disposable income to charity would do more good than his spending any of it on himself. Even so, on commonsense morality, this level of sacrifice would be considered supererogatory, not obligatory, as HAU implies.

Worse yet, HAU implies that many intuitively supererogatory acts are actually morally wrong (McConnell 1980). Suppose, for instance, that Carlos has just been given the last piece of cherry pie. Since, unlike Sergio, cherry pie is Carlos's favorite, Carlos would enjoy eating the last piece of pie more than Sergio would. Suppose, then, that Carlos must choose between giving Sergio the last piece of pie, thereby enhancing Sergio's utility by five utiles, or keeping the last piece for himself, thereby enhancing his own utility by six utiles. In this case, HAU implies

that it would be wrong for Carlos to choose to benefit Sergio instead of himself. Yet, arguably, his giving Sergio the last piece of pie is supererogatory. To avoid such counterintuitive implications and to accommodate a wider range of supererogatory acts, the act-consequentialist must adopt a view such as SU instead.

So, as we've seen above, one general sort of objection to TAC is that it implausibly holds that it is never permissible to fail to maximize the good. We just considered two instances of this objection: (i) that it is permissible to fail to maximize the good when doing so will make only oneself worse off and (ii) that it is permissible to fail to maximize the good when the alternative would be unreasonably demanding. In the next section, we'll consider a third instance of this general objection: (iii) that it is permissible to fail to maximize the good when one doesn't know what would maximize the good.

Traditional Act-Consequentialism and Not Knowing What Will Maximize the Good:

We rarely know what will maximize the good. Although it is sometimes possible to anticipate the immediate effects of our actions, it is clearly impossible to anticipate all the effects that our actions will have over an indefinite future. And these unforeseen future effects may be truly massive, affecting, perhaps, the identities of people over countless generations (Lenman 2000). Such massive unforeseen effects would likely dwarf whatever immediate effects we do foresee. The upshot of all this is that we are literally clueless as to which of our actions would maximize the good (Lenman 2000). And, given our cluelessness, TAC can be of no practical use in deciding what to do.

At this point, the TACist can point out that this is really no objection to the theory, for TAC provides only a *criterion of rightness* and not a *decision-making procedure* (Bales 1971). Whereas a decision-making procedure must be something that actual agents with limited knowledge and imperfect calculating abilities can use to guide their decisions about what to do, a criterion of rightness does not. A criterion of rightness need only specify the fundamental features of actions that determine whether they are right or wrong, and it may turn out that actual agents are often unable to discern whether their actions have these features. So, if TAC is meant to provide only a criterion of rightness, which it is, then it is no objection to it that it fails as an adequate decision-making procedure.

At this point, the critics of TAC may wish to distinguish between *objective rightness* and *subjective rightness*. Subjectively right (i.e., required) actions are those that we can legitimately expect *actual* agents with limited knowledge and imperfect calculating abilities to perform if acting conscientiously and, thus, are the ones that we can rightfully blame them for failing to perform. Objectively right actions, by contrast, are those that we could legitimately expect only *idealized* agents with full knowledge and perfect calculating abilities to perform if acting conscientiously.

Having made this distinction, the critics of TAC can argue that the primary function of an ethical theory is to provide an account of rightness that can be used to guide our practical deliberations and to inform our assessments of blame and praise—that is, to provide an account of subjective rightness. And this, they argue, is something that TAC fails to do. To see why, consider the following example:

> *Mine Shafts:* A hundred miners are trapped underground in one of two mine shafts. Floodwaters are rising, and the miners are in danger of drowning. Carmine can close one of three floodgates. Depending both on which floodgate she closes and on which shaft the miners are in, the results will be one of the six possibilities given in Table 2.

Table 2

		The miners are in…	
		Shaft A	Shaft B
	Gate 1	Carmine saves 100 lives	Carmine saves zero lives
Carmine closes…	Gate 2	Carmine saves zero lives	Carmine saves 100 lives
	Gate 3	Carmine saves 90 lives	Carmine saves 90 lives

Carmine does not know which shaft the miners are in, but she knows all the above as well as the fact that, given her evidence, there is a 50 percent subjective chance that the miners are in Shaft A as well as a 50 percent subjective chance that the miners are in Shaft B. As a matter of fact, though, the miners are in Shaft A. Assume that Carmine is not in any way at fault for her imperfect epistemic position with regard to the location of the miners. And assume that everything else is equal and, thus, that the unforeseen remote effects of her actions would be the same regardless of which gate she closes and that Carmine knows this.[6]

Carmine does not know whether it is closing Gate 1 or Gate 2 that would maximize the good, but she knows that there is no chance that closing Gate 3 would maximize the good. In fact, we are told that the miners are in Shaft A, which means that it is closing Gate 1 that would maximize the good. Thus, according to TAC, Carmine is required to close Gate 1. But if we are talking about what Carmine is required to do in the subjective sense—the sense that is supposed to guide her practical deliberations and inform our assessments of blame and praise for her actions, then it is counterintuitive to suppose that Carmine is required to close Gate 1. Indeed, it seems that, in the subjective sense, Carmine should instead close Gate 3. Here, then, is a case in which it seems (subjectively) morally permissible for an agent to fail to maximize the good. And so here we have a genuine objection to TAC, assuming that the primary function of an ethical theory is to provide a criterion of subjective rightness. The objection would be that claim (I)—the claim that it is never permissible to fail to maximize the good—is false, for Carmine should close Gate 3 even though her doing so would fail to maximize the good.

To see why Carmine should, in the subjective sense, close Gate 3 and not

Gate 1, recall that Carmine has no idea whether it is closing Gate 1 or Gate 2 that would maximize the good. Thus, we should not expect Carmine, if morally conscientious, to close either one. Doing so would be too risky, carrying a substantial subjective risk of saving no one. In closing Gate 3, by contrast, she guarantees that 90 miners will be saved. And given the associated risk, we would blame Carmine if she closes Gate 1 even though this is what TAC implies that she was required to do. So, if we are looking for an account of rightness that can be used to guide our practical deliberations and inform our assessments of blame and praise, TAC is not it.

There are at least two ways to respond to this objection. One is to deny that the primary function of an ethical theory is to guide our practical deliberations and inform our assessments of blame and praise. Perhaps, the primary function of ethical theory is instead to figure out what our objective ends should be (e.g., maximizing the good, acting virtuously, or abiding by the Categorical Imperative). This would leave it to decision theory or some other branch of practical philosophy to provide an account of how agents with limited knowledge should make practical decisions given both their normative uncertainty about what their objective ends should be and their nonnormative uncertainty about how best to achieve them.[7]

Another way to respond to the above objection is to accept that the primary function of an ethical theory is to provide an account of subjective rightness and try to revise TAC so that it can serve this function. Some (e.g., Jackson 1991) have suggested that the way to do this is to adopt *subjective consequentialism:*

> **SC** (1.0) A subject's performing x is (subjectively) morally permissible if and only if, and ultimately because, there is no available alternative, y, whose outcome ranks higher than x's outcome, and (2.3) y's outcome ranks higher than x's outcome if and only if y's outcome is better than x's outcome in terms of expected value.

A subject maximizes expected value if and only if they maximize $\sum_i Pr(O_i/A_j) \times V(O_i)$, where Pr is the subject's probability function at the time, V is consequentialism's value function, O_i are the possible outcomes, and A_j are the possible actions (Jackson 1991, 463–4). But even if SC gets the intuitively correct answer in *Mine Shafts*, it suffers from two significant problems. First, given agents' limited knowledge and imperfect calculating abilities, they are hardly ever in the position to reliably figure out which of the countless alternatives available to them would have the highest expected value over an indefinite future (Feldman 2006). After all, the only reason that SC can be used in *Mine Shafts* to guide Carmine's practical deliberations and to inform our assessments of blame and praise is because it was artificially stipulated that the unforeseen remote effects of

Carmine's actions would be the same regardless of which gate she closes and that she knows this. So, although SC might be useful in a few artificial cases like *Mine Shafts*, it is not much of an improvement over TAC in providing action guidance.

Second, SC seems suspect in that it allows for an agent's nonnormative uncertainty, but not for her normative uncertainty, to affect the subjective rightness of her actions. Whereas SC allows that an agent's uncertainty about which act will maximize the good affects the subjective rightness of her actions, it does not allow for her uncertainty as to whether she ought to be seeking to maximize the good to affect the subjective rightness of her actions. This is potentially problematic, for arguably an agent's nonculpable uncertainty with regard to the relevant normative facts (e.g., whether her objective end should be to maximize the good or something else) is just as important in determining what we can legitimately expect her to do and to blame her for not doing as her nonculpable uncertainty with regard to the relevant nonnormative facts is (e.g., whether this or that act will best achieve the objective ends that she ought to be pursuing). And, because of this, SC seems to fall prey to a counterexample that is analogous to *Mine Shafts*, only in this example we swap normative uncertainty for nonnormative uncertainty.

I will present the counterexample shortly, but first let me state my assumptions. First, I will assume that although impermissibility does not come in degrees, the moral badness of impermissible acts does. That is, some impermissible acts are morally worse than others. For instance, on TAC, any act that fails to maximize value is impermissible. But it is plausible to suppose that some acts that fail to maximize value are worse than others, as some fall further from the mark of maximal value than others do. Likewise, on Kantianism, any act that violates the Categorical Imperative is impermissible. But it is plausible to suppose that some violations of the Categorical Imperative are morally worse than others. For instance, it seems that violating the Categorical Imperative by killing someone with an otherwise bright future is much worse than violating the Categorical Imperative by telling a relatively harmless lie. Second, I will assume that we can measure the degree to which acts are morally good/bad in terms of their *deontic value*, where the deontic value of a supererogatory act is always positive, the deontic value of an impermissible act is always negative, and the deontic value of a non-supererogatory optional act is always zero. (Let 'DV(x)' stand for 'the deontic value of x'.) Third, and more controversially, I will assume that we can make inter-theoretic comparisons of deontic value.[8] Fourth and last, I will assume that TAC and Kantianism are accounts of objective rightness. Now, here is the counterexample to SC:

> *Normative Uncertainty:* Norma must choose among the following three mutually exclusive and jointly exhaustive options: a_1, a_2, and a_3. (To ensure that they are jointly exhaustive, let a_3 be the act of doing something other than either a_1 or a_2.) If she does a_1, she will maximize value (not deontic value, but nonmoral value)[9] but violate the

Categorical Imperative. If she does a_2, she will abide by the Categorical Imperative but fail to maximize value. And if she does a_3, she will both fail to maximize value and violate the Categorical Imperative. Depending both on which theory of objective rightness is true and on which act she performs, the deontic value of her action will be one of the six possibilities given in Table 3.

Table 3

		The true theory is . . .	
		TAC	Kantianism
	a_1	$DV(a_1) = 0$	$DV(a_1) = -100$
Norma performs...	a_2	$DV(a_2) = -100$	$DV(a_2) = 0$
	a_3	$DV(a_3) = -10$	$DV(a_3) = -10$

Norma does not know whether TAC or Kantianism is true, but she knows all the above as well as the fact that, given her evidence, there is a 50 percent subjective chance that TAC is true as well as a 50 percent subjective chance that Kantianism is true. Assume that Norma is not in any way at fault for her imperfect epistemic position with regard to which ethical theory is true. And note that since Norma knows that her performing a_1 is what would maximize value, it follows that Norma's performing a_1 is also what would maximize expected value.

Norma does not know whether it is performing a_1 or a_2 that is morally permissible, but she knows that there is no chance that performing a_3 is morally permissible. Of course, since Norma knows, as stipulated, that her performing a_1 would maximize value, it follows that her performing a_1 would maximize expected value (not expected deontic value, but expected nonmoral value). So, according to SC, Norma is required to perform a_1. But if we are talking about what Norma is required to do in the subjective sense—the sense that is supposed to guide her practical deliberations and inform our assessments of blame and praise for her actions, then it is counterintuitive to suppose that Norma is required to perform a_1. After all, given her ignorance of whether it is TAC or Kantianism that is true, we should expect Norma, if morally conscientious, to perform a_3. For performing either a_1 or a_2 is just too risky. They both carry the substantial subjective risk of performing an act with a deontic value of -100, whereas if she performs a_3, she can be certain that the deontic value of her act will be no worse than -10. Given the associated risk, we would blame Norma for performing a_1 even though this is what SC implies that she is required to do. So, if we are looking for an account of rightness that can be used to guide our practical deliberations and inform our assessments of blame and praise, then SC is not it.

Second objection: It is sometimes impermissible to maximize the good.

Agent-Centered Constraints:

As we saw above, TAC seems too demanding. Interestingly, it seems to be too permissive as well. On TAC, there are no types of acts that are off-limits, not even rape, murder, torture, or genocide. Of course, the TACist can hold that, say, murder is especially bad and, indeed, so bad as to make one murder worse than any number of deaths by natural causes. The TACist can, then, hold that it is impermissible to murder one person even to produce a vaccine that would save countless lives. Nevertheless, the TACist must still concede that it would be permissible to commit one murder to prevent two other comparable murders. Since no matter how bad murder is, two murders will, other things being equal, be worse than one murder, the TACist must concede that it will be permissible to commit murder to minimize the overall commissions of murder. Thus, on TAC, any type of act, no matter how intrinsically bad, will be permitted in those circumstances in which committing one instance of that act-type will minimize overall commissions of that act-type.

This illustrates a more general problem with TAC: it is unable to accommodate *agent-centered constraints*. An agent-centered constraint is a constraint on maximizing the good that it would be wrong to infringe upon even in some circumstances in which doing so would minimize comparable infringements of that constraint.[10] On commonsense morality, there are two types of agent-centered constraints: *agent-centered restrictions* and *special obligations*. Agent-centered restrictions *prohibit* agents from performing certain types of acts (e.g., murder) even in some circumstances in which doing so would minimize comparable commissions of that act-type. Special obligations, by contrast, *require* agents to do certain things (e.g., to keep their promises) even in some circumstances in which their failing to do so would minimize comparable failings of that type. These special obligations often arise as the result of performing certain past acts (e.g., making a promise) but also come with occupying certain roles (e.g., certain professional and familial roles). Since TAC is unable to accommodate these agent-centered constraints, it can seem too permissive. This is the *permissiveness objection*.

If we give credence to the intuitions that ground the permissiveness objection, we might think that TAC must be abandoned for the sake of some version of consequentialism that can accommodate agent-centered constraints. One such version can be obtained by modifying SU, the result being:

> **Mod-SU** (1.0) A subject's performing x is morally permissible if and only if, and ultimately because, there is no available alternative, y, whose outcome ranks higher than x's outcome, and (2.4) y's outcome

ranks higher than *x*'s outcome if and only if there is no available alternative that would produce both more constraint-adjusted utility and more comprehensively-adjusted utility than *x* would, where the constraint-adjusted utility of an outcome is just the sum of the utility for others, adjusted by multiplying any disutility (or loss of utility) resulting from the subject's infringement of an agent-centered constraint by 500, and where the comprehensively-adjusted utility of an outcome is just its constraint-adjusted utility added to the product of the subject's utility times ten.

In other words, Mod-SU holds that an act is morally permissible so long as there is no available alternative that produces both more constraint-adjusted utility and more comprehensively-adjusted utility. Like SU, Mod-SU allows agents to give their own utility anywhere from no weight at all to ten times the weight that everyone else's utility has from the impersonal perspective. But, unlike SU, Mod-SU requires that agents give the disutility caused by their infringements of constraints 500 times the weight that this disutility has from the impersonal perspective.

Mod-SU could easily accommodate a non-absolute constraint against, say, committing murder.[11] Agents would, other things being equal, be prohibited from committing murder even to prevent up to 500 others from each committing some comparable murder. And Mod-SU could easily accommodate a non-absolute constraint against failing to fulfill the special obligation that parents have to nurture their own children. Parents would, other things being equal, be prohibited from failing to nurture their own children even to prevent up to 500 others from each comparably failing to nurture their children. Clearly, then, Mod-SU can accommodate both agent-centered restrictions and special obligations.

Infinite Goodness and the Strong Pareto Principle:

TAC can also seem too permissive in cases in which all the alternatives open to an agent would produce an infinite amount of goodness. Suppose, for instance, that you have the choice either to do *x* or to refrain from doing *x*. If you do *x*, W_x will be the actual world. If you refrain from doing *x*, W_{-x} will be the actual world.[12] These two possible worlds are depicted in Table 4. Assume that the two worlds have the same locations of value (e.g., the same people or times) and that they have an infinite number of locations.[13] Thus, the ellipses indicate that the 1's should continue indefinitely in each direction in both W_x and W_{-x}.

Table 4

W_x	$\ldots, 1, 1, 1, 1, -8, 1, 1, 1, 1, \ldots$
$W_{\sim x}$	$\ldots, 1, 1, 1, 1, -2, 1, 1, 1, 1, \ldots$

Intuitively, it seems wrong for you to do x. Yet if we think that one world can be better than another if and only if the total sum of goodness contained by the one is greater than the total sum of goodness contained by the other, then TAC implies that its permissible to do x. After all, the total sum of goodness contained by both W_x and $W_{\sim x}$ is infinity, and no total sum is greater than infinity. Thus, on the assumption that one world can be better than another if and only if the total sum of goodness contained by the one is greater than the total sum of goodness contained by the other, TAC counterintuitively implies that it is permissible for you do x, for $W_{\sim x}$ would, then, be no better than W_x. So, here is another case in which it seems impermissible to maximize the good, for, in this case, it seems impermissible to maximize the good by doing x.

One way to defend TAC against this objection is to reject the assumption that one world can be better than another if and only if the total sum of goodness contained by the one is greater than the total sum of goodness contained by the other and to hold instead the *strong pareto principle:* if every location has at least as much goodness in W_i as it does in W_j, and at least one location has more goodness in W_i than it does in W_j, then W_i is better than W_j (Lauwers and Vallentyne 2004). On this principle, $W_{\sim x}$ is better than W_x. Thus, when combined with the strong pareto principle, TAC gets the intuitively correct result, implying that it is impermissible for you to do x.

Another response to the objection is to accept the assumption that one world can be better than another if and only if the total sum of goodness contained by the one is greater than the total sum of goodness contained by the other and so accept that we need to replace TAC with something along the lines of the following view, which I call *rational-preference consequentialism:*

> **RPC** (1.0) A subject's performing x is morally permissible if and only if, and ultimately because, there is no available alternative, y, whose outcome ranks higher than x's outcome, and (2.5) y's outcome ranks higher than x's outcome if and only if the subject has decisive reason, all things considered, to prefer y's outcome to x's outcome.

RPC ranks alternative outcomes not in terms of their impersonal goodness, but in terms of the subject's (agent-neutral and agent-relative) reasons to prefer each to the others. So, even if $W_{\sim x}$ is not, strictly speaking, better than W_x, you ought to prefer $W_{\sim x}$ to W_x. For the fact both that every location has at least as much goodness in $W_{\sim x}$ as it does in W_x and that at least one location has more goodness in $W_{\sim x}$ than it does in W_x would seem to constitute a decisive reason for you to

prefer W_{-x} to W_x. Thus, RPC implies that it is impermissible for you to maximize the good by doing x.

Third objection: It is not the case that if the subject's performing x is morally permissible, it is ultimately because their performing x would maximize the good.

One might accept that the subject's performing x is morally permissible if and only if their performing x would maximize the good but deny that what makes their performing x morally permissible is that their doing so would maximize the good. Philippa Foot (1985), for instance, holds that it is the other way around: that what makes a subject's performing x an act that would maximize the good is that it is a morally permissible act, for she argues that we have no grasp of the notion of a good state of affairs that is independent of what is required by the virtue of benevolence. And, on her view, benevolence is itself constrained by the other virtues; that is, benevolence comes into play only in choosing among actions that are not ruled out by the other virtues (such as, justice). Thus, it makes no sense to say that the state of affairs in which you commit one murder is better than the one in which five others each commit murder, for your committing murder, even to prevent five others from doing the same, is ruled out by justice. On Foot's view, then, we can make sense of one state of affairs being better than another only if we see benevolence as requiring us to bring about the one as opposed to the other and, thus, see no other virtue as prohibiting us from doing so. And, therefore, we can have no grasp of one state of affairs being better than another that is independent of our understanding of whether it would be permissible for us to bring about the one as opposed to the other.

Others, though, have questioned this view. For instance, Samuel Scheffler asks: "[D]o we really cease to understand what is meant by 'a better state of affairs' if the question is raised whether infringing a right or telling a lie or treating a particular individual unfairly might perhaps produce a better state of affairs than failing to do so?" (1985, 412). But even if we side with Scheffler on this point, we might think that some of the versions of consequentialism discussed above, such as Mod-SU and RPC, are particularly prone to this sort of Footian worry. On these consequentialist views, the outcomes of actions are not given an agent-neutral ranking in terms of their impersonal goodness, but are instead given an agent-relative ranking in terms of the agent's reasons for preferring each outcome to its alternatives. And it may seem that the correct agent-relative ranking was settled based on our intuitions about what it is right and wrong to do. The worry, then, is that insofar "as an agent-relative ordering of all states of affairs from better to worse can be constructed only once we have figured out which actions are right, it cannot serve as a guide to rightness" (van Roojen 2004, 169). Thus, the Footian worry is that these more nuanced versions

of "consequentialism" avoid counterintuitive implications only at the cost of abandoning one of consequentialism's core commitments: specifically, the commitment to the idea that the deontic statuses of actions are determined by how their outcomes rank, not vice versa.[14]

Other Responses to These Objections

I have suggested that one move that the TACist can make in response to an objection is to modify TAC so that it avoids the objection. Specifically, I have suggested modifying TAC by modifying the way it ranks outcomes—that is, by modifying (2.0) in the formulation of TAC given at the beginning of the section "Act-Utilitarianism and Traditional Act-Consequentialism." This strategy is called *consequentializing*. The idea is to start with some plausible non-consequentialist theory that avoids the objection in question and take whatever features that it considers to be relevant to assessing the deontic statuses of actions and incorporate them as features that are relevant to how outcomes rank, thereby generating a consequentialist theory that avoids the objection.[15]

There are, of course, other ways for the TACist to respond to an objection. One is to accept the general conciliatory approach of thinking that TAC must be modified to avoid the objection but to modify TAC by modifying its account of how its ranking of outcomes determines the deontic statuses of actions—that is, by modifying (1.0) as opposed to, or in addition to, modifying (2.0) in the above formulation of TAC. Another way for the TACist to respond to an objection is not at all conciliatory and involves trying to discredit the intuitions on which the objection is based. I explain both strategies below.

Modifying (1.0) of TAC

If one takes the intuitions behind the demandingness objection seriously, one might reject TAC and adopt *satisficing consequentialism* in its place:

> **SAT** (1.1) A subject's performing x is morally permissible if and only if, and ultimately because, x's outcome ranks high enough relative to the best available alternative(s), and (2.6) x's outcome ranks high enough relative to the best available alternative(s) if and only if the total amount of (impersonal) goodness contained in x's outcome plus n ($n > 0$) is greater than or equal to the total amount of (impersonal) goodness contained within each of the best available alternatives.[16]

SAT can accommodate agent-centered options and supererogatory acts. The problem, though, is that it permits agents to sacrifice their own good for the sake of making things worse for others (Bradley 2006). To illustrate, consider the case

depicted in Table 5.

Let S = the subject, U(x) = the overall utility that would be produced if the subject does x, and let $n = 80$. Assume that a_1 is the act of your keeping to yourself while you work to complete some important project for your job, that a_2 is the act of your instead dissuading six others from donating money to Oxfam, that a_3 is the act of your instead dissuading seven others from donating money to Oxfam, and that a_4 is the act of your instead dissuading eight others from donating money to Oxfam. Assume that your failure to complete the important project will cost you ten utiles. And assume that for each person that you dissuade from donating money to Oxfam there will be ten fewer utiles of aggregate utility.

Table 5

x	$U(x)$	*deontic status*
a_1	**+190**	supererogatory and best
a_2	**+120**	supererogatory
a_3	**+110**	merely permissible
a_4	**+100**	impermissible

SAT implies that it is permissible for you to neglect your work to dissuade seven others from donating their money to Oxfam, thereby making things worse not only for you, but also for others. What is worse, SAT implies that it is actually supererogatory for you to neglect your work to dissuade six others from donating their money to Oxfam, thereby making things worse not only for you, but also for others. The idea that making a personal sacrifice for the sake of making things worse for others can be permissible and even supererogatory is utterly absurd.

Perhaps, then, the best solution to the demandingness objection is, not to adopt SAT, but to adopt *scalar consequentialism*:

> **SCAL** (1.2) A subject's performing y is morally better than their performing x if and only if, and ultimately because, y's outcome ranks higher than x's outcome, and (2.0) y's outcome ranks higher than x's outcome if and only if y's outcome is (impersonally) better than x's outcome.[17]

Since SCAL does not provide a criterion for assessing the permissibility of actions but provides only a criterion for assessing the comparative deontic value of alternative acts, SCAL avoids issuing any requirements (or permissions) at all. And by avoiding issuing any demands at all, it certainly avoids the demandingness objection. But it does so at the price of not even providing a criterion of rightness. Thus, SCAL succeeds in avoiding the demandingness objection only at the price of

failing to fulfill the function of an ethical theory, which is, it seems, to provide a criterion of rightness.

If we think that this is indeed too high a price to pay, we could instead avoid the demandingness objection, and the permissiveness objection as well, by adopting *rule-consequentialism:*

> **RC** (1.3) A subject's performing *x* is morally permissible if and only if, and ultimately because, *x* is permitted by the code of rules with the highest-ranked associated outcome, and (2.7) the outcome associated with one code of rules ranks higher than the outcome associated with another code of rules if and only if the one is (impersonally) better than the other.[18]

Not very long ago, the consensus was that RC is untenable, the thought being that it either collapses into practical equivalence with TAC or suffers from incoherence. Brad Hooker (2000) has been largely responsible for shattering this consensus. He has shown that if we formulate RC so that it ranks codes of rules in terms of the amount of good that would result from their gaining widespread *acceptance* as opposed to the amount of good that would result from their being universally *complied with*, RC will not collapse into practical equivalence with TAC.

The thought was that RC will collapse into practical equivalence with TAC, because a rule that never permits failing to maximize the good (such as "Take care of your loved ones except when doing otherwise would maximize the good") will always win out over one that sometimes permits failing to maximize the good (such as "Take care of your loved ones"), for the former type of rule would produce more good if universally complied with. So, when formulated in terms of compliance, RC collapses into practical equivalence with TAC.

However, this does not happen when we formulate RC in terms of acceptance. For there are costs associated with getting each new generation to accept a given code of rules, and, thus, the ideal code of rules cannot be too complicated or too demanding or the costs of getting the rules inculcated will exceed the benefits of having them widely accepted. Consider, for instance, that the costs of getting a rule such as "Take care of your loved ones except when doing otherwise would maximize the good" to become widely accepted would be incredibly high given people's natural inclination to favor their loved ones. Thus, when formulated in terms of acceptance, RC avoids the collapse objection by including rules that sometimes permit failing to maximize the good.

However, the worry now is that if RC includes rules that sometimes permit failing to maximize the good, then RC must be guilty of incoherence. After all, it may seem that RC is ultimately committed to the maximization of the good, and it certainly seems incoherent to be committed to the maximization of the good while holding that it is sometimes permissible to fail to maximize the good. However, Hooker (2000) argues that RC is not committed to the maximization of

the good. The argument for RC is not that it is the theory that derives from some overarching commitment to the maximization of the good, but that it does a better job of meeting various desiderata (e.g., internal consistency, coherence with our considered moral convictions, etc.) than any other alternative moral theory does.

The dismissive reply

In the above, I have suggested several ways that we might modify TAC to avoid various objections. Most of these objections stem from certain commonsense intuitions about what it is right and wrong to do in particular cases. Some consequentialists have argued, though, that instead of modifying TAC in response to such objections, we should instead dismiss the intuitions on which these objections are based. The thought is that if TAC conflicts with our commonsense morality and the intuitions associated with it, then so much the worse for our commonsense morality. Recently, this sort of blusterous response has been given more substance via an appeal to the work of the psychologist Joshua D. Greene. Greene has conducted experiments using functional magnetic resonance imaging that purport to show that emotional neural processes tend to generate what he takes to be characteristically deontological judgments, whereas cognitive neural processes tend to generate what he takes to be characteristically consequentialist judgments. And Peter Singer (2005) and Greene (2003, 2008) have both argued that these empirical findings impugn the credibility of these characteristically deontological intuitions without impugning the credibility of these characteristically consequentialist intuitions. The empirical findings purport to show that the emotional processes that give rise to these characteristically deontological intuitions are tracking whether or not the relevant harm is up-close and personal. And they argue that since this is clearly a morally irrelevant factor, it follows that the emotional processes and the characteristically deontological intuitions that they give rise to should be deemed unreliable. Singer explains:

> If…our intuitive responses are due to differences in the emotional pull of situations that involve bringing about someone's death in a close-up, personal way, and bringing about the same person's death in a way that is at a distance, and less personal, why should we believe that there is anything that justifies these responses? …For most of our evolutionary history, …violence could only be inflicted in an up-close and personal way—by hitting, pushing, strangling, or using a stick or stone as a club. To deal with such situations, we have developed immediate, emotionally based responses to questions involving close, personal interactions with others. The thought of pushing the stranger off the footbridge elicits these emotionally based responses. Throwing a switch that diverts a train that will hit someone bears no resemblance to anything likely to have happened in circumstances in which we and our ancestors lived…. But what is the moral salience of the fact that I have killed someone in a way that was possible a million years ago, rather than in a way that became possible only two

Thus, Singer and Greene argue that there is good reason to question the reliability of our deontological intuitions: they are tracking what appear to be morally irrelevant factors. But there seems to be no equivalent reason for thinking that the consequentialist factors that our cognitive processes are tracking are likewise morally irrelevant. Indeed, they argue, unlike in the case of our characteristically deontological intuitions, we cannot give an evolutionary debunking of our characteristically consequentialist intuitions. And if all that is right, then perhaps we should reject the deontological intuitions on which the objections to TAC rest rather than reject TAC itself.[20]

A Different Kind of Objection to Traditional Act-Consequentialism

Thus far, we've considered objections to TAC's three claims: (I) it is never permissible to fail to maximize the good, (II) it is always permissible to maximize the good, and (III) if a subject's performing x is morally permissible, it is ultimately because their doing so would maximize the good. But some objections don't take aim at these claims, but instead take aim at some overarching implication of the view. I'll now consider three such overarching objections.

First overarching objection: TAC implausibly implies that it can be indeterminate whether a subject has acted permissibly.

Imagine that a sniper shoots and kills Gunnar as he is attempting to enter an elementary school with a gun. Did the sniper maximize the good in shooting Gunnar? Well, it depends on whether shooting Gunnar was necessary to protect others, which in turn depends on whether Gunnar would have hurt anyone had he not been shot. But it's entirely possible that Gunnar had the kind of libertarian freedom that implies that there is just no fact of the matter as to what he would have done had he not been shot. And, in that case, it is indeterminate whether the sniper maximized the good.

Now, on TAC, whether the sniper acted permissibly in shooting Gunnar just depends on whether he maximized the good in doing so. So, if TAC implies that it's indeterminate whether he maximized the good, then it also implies that it's indeterminate whether he acted permissibly. And if it's indeterminate whether the sniper acted permissibly, then it is indeterminate whether should blame him for what he did. What's more, it's indeterminate whether the sniper should act in the same way should he ever again face a relevantly similar situation. But if TAC can't tell the sniper what he should do, nor tell us whether we should blame him for what he does, then we might wonder what's the point of such a moral theory. What's worse, it seems that whether the sniper acted permissibly just depends on factors such as what

the objective probability was that Gunnar would have hurt people had he not been shot and on how much hurt he would have then caused. But, instead, TAC implies that, regardless of such factors, it's just indeterminate whether the sniper acted permissibly. And that seems implausible.

Second overarching objection: TAC implausibly implies that permissibility doesn't distribute over conjunction.

Imagine that I would produce the most good if I were to raise both my arms, that a distant second best would be my raising neither arm, and that worst of all would be my raising only one arm. And let's assume that I have all the following distinct options: raising both my left arm and my right arm, raising neither my left arm nor my right arm, raising my left arm, raising only my left arm, raising my right arm, and raising only my right arm. Lastly, assume that, because I'm a bad person who's inclined to make things go as badly as possible, I would not raise my right arm if I were to raise my left arm. And, likewise, I would not raise my left arm if I were to raise my right arm. Note that the assumption here is not that I *could not* raise my left arm were I to raise my right arm, but rather that I *would not* raise my left arm were I to raise my right arm. Now see Table 6 for a list of my options and their resultant outcomes.

Table 6

Option	Produces	Ranking
Raise my left arm	The least good	Worst
Raise my right arm	The least good	Worst
Raise only my left arm	The least good	Worst
Raise only my right arm	The least good	Worst
Raise both my arms	The most good	Best
Raise neither of my arms	Neither the most nor the least good	2nd Best

The problem here is that, intuitively, permissibility distributes over conjunction. If I'm permitted to do x and y, then I'm permitted to do x (as well as permitted to do y). Yet, TAC denies this. For it implies both that it's permissible for me to raise both my left arm and my right arm but that it's impermissible for me to raise my left arm. It implies that I'm permitted to raise both my left arm and my right arm, because there is no alternative that would produce more good. But it implies that I'm not permitted to raise my left arm, for there is an alternative (namely, my raising neither arm) that would produce more good.[21]

Third overarching objection: TAC is implausibly collectively inefficacious.

TAC seems to run into trouble when it comes to what we would expect a collection of agents to produce if they were all to act morally. To illustrate, let's consider a case where there are just two individuals: Darcy and Dougal. Each has their own button in front of them, and each has exactly two options: push or don't push. And let's assume that neither can affect what the other will do but each knows what the other will do. And see Table 7 for the four possible outcomes. Now, imagine that both Darcy and Dougal are bad people who each wants both what's best for themself and what's worse for others. Consequently, both end up not pushing their buttons, and this results in an outcome with −10 units of goodness, which is an outcome that's far worse than the best outcome that they could collectively achieve: the one where they both push and thereby produce an outcome with 8 units of goodness. Nevertheless, according to TAC, Darcy and Dougal both act permissibly. For they both did what would, given what the other was going to do, maximize utility.

Table 7

	Dougal pushes	Dougal doesn't push
Darcy pushes	Dougal: −1 Darcy: −1 <u>Others: 10</u> Total: 8	Dougal: 0 Darcy: −1 <u>Others: −10</u> Total: −11
Darcy doesn't push	Dougal: −1 Darcy: 0 <u>Others: −10</u> Total: −11	Dougal: 0 Darcy: 0 <u>Others: −10</u> Total: −10

Nevertheless, many believe that the correct moral theory must be collectively efficacious such that the agents who abide by it, whoever and however numerous they may be, are guaranteed to produce the morally best world that they could collectively achieve. After all, morality is supposed to be a collective enterprise. So, we should be able to make things as morally good as possible if we all act morally. But TAC implies otherwise. For, according to TAC, Darcy and Dougal have done everything that TAC directs them to do even though they fall well short of bringing about the morally best world that they could collectively achieve.[22]

Conclusion

Historically, the debate over consequentialism has focused on TAC with consequentialists responding to objections either by abandoning TAC for some other more nuanced version of consequentialism or by attempting to discredit the intuitions on which the objections are based. Consequentialism has thus proved amazingly resilient to the onslaught of criticisms that have been leveled against it. Ultimately, though, whether consequentialism will rule the day depends on what the best version of it is and how this, the best version, compares to the best versions of its non-consequentialist rivals. Only time will tell.[23]

Bibliography

Arneson, R. (2005), "Sophisticated Rule Consequentialism: Some Simple Objections."
  Philosophical Issues 15, 235–51.

Bales, R. E. (1971), "Act-Utilitarianism: Account of Right-Making Characteristics or
  Decision-Making Procedure?" *American Philosophical Quarterly* 8, 257–65.

Bentham, J. (1789), *An Introduction to the Principles of Morals and Legislation*, J. Burns and H.
  L. A. Hart (eds.). London: Athlone Press, 1970.

Berker, S. (2009), "The Normative Insignificance of Neuroscience." *Philosophy &
  Public Affairs* 37, 293–329.

Bradley, B. (2006), "Against Satisficing Consequentialism." *Utilitas* 18, 97–108.

Bykvist, K. (2003), "Normative Supervenience and Consequentialism." *Utilitas* 15,
  27–49.

Dreier, J. (1993), "Structures of Normative Theories." *The Monist* 76, 22–40.

Feldman, F. (2006), "Actual Utility, the Objection from Impracticality, and the Move to
  Expected Utility." *Philosophical Studies* 129, 49–79.

Foot, P. (1985), "Utilitarianism and the Virtues." *Mind* 94, 196–209.

Greene, J. D. (2008), "The Secret Joke of Kant's Soul," in W. Sinnnott-Armstrong (ed.),
  *Moral Psychology, Vol. 3: The Neuroscience of Morality: Emotion, Brain Disorders, and
  Development*. Cambridge, MA: MIT Press, 35–79.

——. (2003), "From Neural "Is" to Moral "Ought": What Are the Moral Implications of
  Neuroscientific Moral Psychology?" *Nature Reviews Neuroscience* 4, 847–50.

Harwood, S. (2003), "Eleven Objections to Utilitarianism," in L. Pojman (ed.), *Moral
  Philosophy: A Reader,* 3rd Edition. Indianapolis: Hackett Publishing, 179–92.

Hooker, B. (2000), *Ideal Code, Real World.* Oxford: Clarendon Press.

Hurley, P. (2006), "Does Consequentialism Make Too Many Demands, or None at All?"
  Ethics 116, 680–706.

Jackson, F. (2014), "Procrastinate Revisited.*" Pacific Philosophical Quarterly* 95, 634–47.

——. (1991), "Decision-Theoretic Consequentialism and the Nearest and Dearest
  Objection." *Ethics* 101, 461–82.

Kagan, S. and Vallentyne, P. (1997), "Infinite Value and Finitely Additive Value Theory." *The
  Journal of Philosophy* 94, 5–26.

Kagan, S. (1991), *The Limits of Morality*. Oxford: Oxford University Press.

Kawall, J. (2003), "Self-Regarding Supererogatory Actions." *Journal of Social
  Philosophy* 34, 487–98.

Lauwers, L. and Vallentyne, P. (2004), "Infinite Utilitarianism: More Is Always
  Better." *Economics and Philosophy* 20, 307–30.

Lawlor, R. (2009a), *Shades of Goodness: Gradability, Demandingness and the Structure of Moral
  Theories.* New York: Palgrave Macmillan.

  ——. (2009b), "The Rejection of Scalar Consequentialism." *Utilitas* 21, 100–16.

Lenman, J. (2000), "Consequentialism and cluelessness." *Philosophy & Public Affairs*,
  29, 342–370.

Louise, J. (2004), "Relativity of Value and the Consequentialist Umbrella." *Philosophical
  Quarterly* 54, 518–36.

McConnell, T. C. (1980), "Utilitarianism and Supererogatory Acts." *Ratio* 22, 36–8.

McNamara, P. (2011). "Supererogation, Inside and Out: Toward an Adequate Scheme for Common-sense Morality." *Oxford Studies in Normative Ethics*, 1, 202–35.

——. (1996), "Making Room for Going beyond the Call." *Mind* 105, 415–50.

Mill, J. S. (1861), *Utilitarianism*, in Roger Crisp (ed.), Oxford: Oxford University Press, 1998.

Norcross, A. (2006), "Reasons without Demands: Rethinking Rightness," in J. Dreier (ed.), *Contemporary Debates in Moral Theory*. Oxford: Blackwell, 5–20.

Parfit, D. (2011), *On What Matters, Volume One*. Oxford: Oxford University Press.

Portmore, D. W. (2019), *Opting for the Best*. New York: Oxford University Press.

——. (2011), *Commonsense Consequentialism: Wherein Morality Meets Rationality*. New York: Oxford University Press.

——. (2009a), "Consequentializing." *Philosophy Compass* 4, 329–47.

——. (2009b), "Rule-Consequentialism and Irrelevant Others." *Utilitas* 21, 368–76.

——. (2008), "Are Moral Reasons Morally Overriding?" *Ethical Theory and Moral Practice* 11, 369–88.

Regan, D. (1980), *Utilitarianism and Co-operation*. New York: Oxford University Press.

Scheffler, S. (1985), "Agent-Centered Restrictions, Rationality, and the Virtues." *Mind* 94, 409–19.

Sepielli, A. (2010), "Normative uncertainty and intertheoretic comparisons," unpublished manuscript.

——. (2009), "What to Do When You Don't Know What to Do." *Oxford Studies in Metaethics* 4, 5–28.

Sider, T. (1993), "Asymmetry and Self-Sacrifice." *Philosophical Studies* 70, 117–32.

Sidgwick, H. (1907), *The Methods of Ethics, Seventh Edition*. London: Macmillan. First Edition 1874.

Singer, P. (2005), "Ethics and Intuitions." *The Journal of Ethics* 9, 331–52.

——. (1999), "A response," in D. Jamieson (ed.), *Singer and His Critics*. Oxford: Blackwell, 269–335.

Sinnott-Armstrong, W. (2003), "Consequentialism." *The Stanford Encyclopedia of Philosophy (Summer 2003 Edition)*, Edward N. Zalta (ed.), http://plato.stanford.edu/archives/sum2003/entries/consequentialism.

Smith, M. (2006), "The Right, the Good and Uncertainty," in T. Horgan and M. Timmons (eds.), *Metaethics after Moore*. Oxford: Oxford University Press, 133–48.

Sobel, D. (2007), "The Impotence of the Demandingness Objection." *Philosophers' Imprint* 7, 1–17.

Stroud, S. (1998), "Moral Overridingness and Moral Theory." *Pacific Philosophical Quarterly* 79, 170–89.

Thomson, J. J. (1990), *The Realm of Rights*. Cambridge, MA: Harvard University Press.

Timmons, M. (2008), "Toward a sentimentalist deontology," in W. Sinnott-Arm- strong (ed.), *Moral Psychology, Volume 3: The Neuroscience of Morality: Emotion, Brain Disorders, and Development*. Cambridge, MA: MIT Press, 93–104.

van Roojen, M. (2004), "The Plausibility of Satisficing and the Role of Good in Ordi- nary Thought," in M. Byron (ed.), *Satisficing and Maximizing*. Cambridge: Cam- bridge University Press, 1–13.

Vessel, J-P. (2010), "Supererogation for Utilitarianism." *American Philosophical Quarterly*,

47, 299–319.

Zimmerman, M. J. (1996), *The Concept of Moral Obligation*. Cambridge: Cambridge University Press.

Notes

1 Some might hold that there is sometimes, on commonsense morality, an option to act to make things worse both for oneself and for others. Whether this is plausible or not depends on whether there is, on commonsense morality, a *pro tanto* moral reason to promote the impersonal good, as Shelly Kagan (1991) has argued. (A *pro tanto* moral reason is one that has genuine weight such that it generates an obligation in the absence of countervailing reasons.)

2 The third and fifth columns appear in boldface type, for these are the two columns that, on SU, are most salient to determining moral permissibility.

3 I'm assuming that, although it wasn't explicitly stated above, SU includes the following claim: a subject's performing x is supererogatory if and only if, and ultimately because, there is a permissible alternative, y, whose outcome contains less utility for others (i.e., those other than the given subject) than x's outcome.

4 I would actually prefer to do away with condition (ii)–see Portmore (2011). That is, I prefer to let the term "supererogatory" denote the pre-theoretic notion of going above and beyond the call of duty (i.e., that of doing more than one has to), and an act can go above and beyond the call of duty without being commendatory, as where the agent performs such an act with a bad motive—see McNamara (2011). If we were to let the term "supererogatory" refer to the pre-theoretical notion of exceeding the minimum that morality demands, as I would prefer, we would be following McNamara (1996). Nevertheless, I include condition (ii) here, for many philosophers take it to be a necessary condition for an act's being supererogatory.

5 An act is optimific if and only if its outcome outranks that of every other available alternative.

6 I borrow this example, with some minor revisions, from Derek Parfit (2011).

7 Michael Smith (2006) defends such a view.

8 See Sepielli (2009, 2010) for a defense of this controversial assumption as well as some suggestions about how this might be done.

9 Whereas only acts can be *morally* good and bad and, thus, have deontic value, states of affairs (such as the state of affairs in which people experience more pleasure) can have nonmoral value—that is, value, *simpliciter*. The deontic value of an act is a measure of how much moral reason there is to perform that act.

10 If there is a constraint against performing a certain act-type, then any commission of an act of that type constitutes an *infringement* of that constraint. But not all infringements of a constraint are wrong. All and only those infringements that are wrong constitute *violations* of that constraint. I follow Thomson (1990, 122) and Zimmerman (1996, 181) in making this terminological distinction.

   Constraints needn't be absolute. It may be permissible to infringe upon a constraint if enough good is at stake, and it may be permissible to infringe upon a constraint if doing so would prevent a sufficient number of comparable infringements of that constraint.

11 If we want Mod-SU to include an absolute constraint against committing murder, we need only replace "500" with "∞" in Mod-SU.

12 I will assume that for each act alternative there is some determinate fact as to what the world would be like were she to perform that alternative. This assumption is some-times called *counterfactual determinism*—see, e.g., Bykvist (2003, 30). Although this assumption is controversial, nothing that I will say here hangs on it. I make the assumption only for the sake of simplifying the presentation. If counterfactual determinism is false, then instead of correlating each alternative with a unique possible world, we will need to correlate each alternative with a probability distribution over the set of possible worlds that might be actualized if S were to perform that alternative. I consider the possibility of such indeterminacy below.

13 This example is borrowed with minor modifications from Kagan and Vallentyne (1997).

14 For a response to this objection, see Portmore (2009a, 2011).

15 See, for instance, Dreier (1993), Portmore (2009a, 2011), and Louise (2004).

16 This is modeled on Ben Bradley's formulation of comparative level satisficing consequentialism—see Bradley (2006, 101).

17 For a defense of this view, see Norcross (2006). For a critique of this view, see Lawlor (2009a, 2009b) and Portmore (2011).

18 What outcome is associated with a given code of rules will depend on how rule-

consequentialism is formulated. It may be the outcome that would result from universal compliance with it or from widespread acceptance of it or from something else. For a defense of rule-consequentialism, see Hooker (2000). For criticisms of rule-consequentialism, see Arneson (2005) and Portmore (2009b).

19 Those unfamiliar with the cases involving switching the train (or trolley) from one track to another and pushing the fat man off the footbridge should see http://www.youtube.com/watch?v=Fs0E69krO_Q.

20 For responses to these sorts of arguments by Greene and Singer, see Berker (2009) and Timmons (2008).

21 For more on this issue, see Jackson (2014) and Portmore (2019).

22 For more on this issue, see Regan (1980) and Portmore (2019).

23 For helpful comments on an earlier draft, I thank Christian Miller.