Gert contra Consequentialism

Walter Sinnott-Armstrong

Bernard Gert is ambivalent about consequentialism. It might not seem so, since he criticizes consequentialism and utilitarianism repeatedly and viciously. Here are some of my favorite excesses: “All forms of consequentialism, including negative consequentialism, are so vague as to be almost totally useless as moral guides” (253). “I think utilitarianism not only an incorrect position but an extremely dangerous one” (373). Gert is less restrained in casual conversation, as I can testify.

In response, I will argue that, when consequentialism is properly understood, it is not useless or dangerous, as Gert claims, because its project is different than his.

Is Gert a Closet Consequentialist?

He doth protest too much, methinks. Despite his vehemence, Gert’s own moral theory and system have important consequentialist elements that allow it to be interpreted as an indirect form of consequentialism. Or so I will argue.

At the very foundation of Gert’s moral theory lies his theory of rationality and reasons. One section heading reads, “No beliefs unrelated to avoiding harms or gaining benefits are basic reasons” (74; cf. 61). Since the consequences of an act include both the harms it avoids and the benefits it gains, Gert holds that all basic reasons for an act must be beliefs about the consequences of that act or about some feature of the act that is related to consequences.

This last disjunct is needed so that a belief that an act violates one of Gert’s second five moral rules will always count as a reason not to do that act. Such beliefs are reasons only because “moral rules are good rules” (75), since moral rules pick out act kinds whose instances usually increase the risk of some bad consequence. A belief that a particular act is of such a kind is a reason to believe that the particular act will increase the risk of some bad consequence. This reason can be undermined; but, if all we know about an act is that it breaks a promise, for example, then this is a reason to believe that it risks some harm to somebody. That does not mean that each individual violation actually increases the objective probability of some bad consequence (pace 127); but it does mean that, in the absence of any underminer, we will always have some epistemic reason to believe that such an act will increase the risk of some bad conse-
sions of rule consequentialism. Instead of asking, "What would happen if everyone did that?" Gert asks, "What would happen if everyone knew that they were allowed to do that?" This is an important improvement. As Gert points out (121), it would be disastrous if everyone stood on their heads all day, but nothing disastrous would happen if everyone knew that they were allowed to stand on their heads all day, since very few would choose or be able to do it. That explains why it is not morally wrong for one person to stand on her head all day.

Another advance is that Gert tells us what that is, i.e., which kind of act is done or known to be allowed. Austin blithely refers to "the class" of acts with no limits on which classes might be relevant. That lack of limits leads to well-known problems. When one tries to formulate rules precisely (cf. 214). If "the class" is acts that maximize utility, "acts of the class are generally done" when everyone does acts that maximize utility, and then the probable effect on the general happiness" would be very good. Thus, every single act that maximizes utility passes Austin's test, so Austin's formulation reduces to act consequentialism. Gert avoids that reduction by specifying a canonical description of the act type in terms of morally relevant features (225 ff).

Admittedly, Gert restricts morally relevant features to concepts that all rational people share, so that justified violations, like moral rules, will be known or knowable by all agents who are bound by the moral system (150-151, 226). Some consequentialists might question why this much knowledge and publicity is necessary for a theory of moral wrongness as opposed to moral responsibility. Nonetheless, each rule consequentialist needs to decide which kinds of rule and procedure for justifying violations determine the moral wrongness of acts. The rationale for making moral wrongness depend on one kind of rule or procedure rather than another need not itself be consequentialist in order for the theory to be consequentialist. Gert does not give a consequentialist rationale for focusing on universal public rules and procedures. Nonetheless, once Gert decides to seek universal public rules and procedures to determine moral wrongness, and once the canonical description of an act is specified, then Gert judges competing candidates for the role of being publicly allowed solely by their consequences. Here again Gert displays the exclusive focus on consequences that makes him a consequentialist.

Of course, Gert's theory is not hedonistic or quantitative or additive. Gert never tells us exactly how to determine whether the consequences of universally publicly allowing all acts of a kind are better than the consequences of not allowing. He says that he would put more weight on preventing harms than on gaining benefits. He probably also would consider other factors in addition to the total harms and benefits of the options, including distribution. But that is still compatible with consequentialism, as Gert recognizes (213). Consequentialism about moral wrongness or wrongness claims only that whether an act morally ought to be done or is morally wrong is determined solely by the consequences of anything, such as the particular act or universally publicly allowing all acts of the same kind. This broad claim is compatible with almost any way of comparing values. Thus, even if Gert does not quantify or add harms or benefits, and even if he considers distribution as well as objective goods, none of that keeps him from being a consequentialist in a broad sense that is common in moral theory.

I conclude that Gert's own moral system can be interpreted as a sophisticated form of negative objective universal public rule consequentialism.
Gert contra Consequentialism

Sophisticated Act Consequentialism

Even so, Gert’s system still contrasts sharply with act consequentialism. That is the view that Gert criticizes so strongly.

It is easy to poke fun at simplistic versions of act utilitarianism, but refuting the crudest versions of a class of views will not refute the whole class. To assess act consequentialism in general, we need to compare Gert’s sophisticated rule consequentialism with the best sophisticated version of act consequentialism.

I will sketch one such version. I will not endorse, much less argue for, this system here. I cannot even spell out all of its details. All I hope to do is to make it clear enough and plausible enough to set up a fair fight with Gert.

Act consequentialism is an instance of direct consequentialism. Direct consequentialists determine which acts a person ought to do by the consequences of that person doing those acts. They can also determine which public rules our society ought to have by the consequences of our society having those public rules. They can also determine which motives (and states of character) a person ought to have by the consequences of that person having those motives (and states of character). What ought to be at each level is determined by the consequences of the very thing being assessed. That is what makes this view direct.

In contrast, rule consequentialism about acts is indirect in that whether an act ought to be done is determined not by the consequences of the act itself but instead by the consequences of some group obeying or accepting a general rule that the act violates or conforms to. Such indirect consequentialism strikes many people as implausible to the extent that it is hard to see why one thing should be evaluated by the consequences of something else. Direct consequentialism also enables us to explain our ambivalence in complex conflicts by allowing different evaluations at different levels, such as when an act morally ought to be done although it violates a good moral rule. The plausibility and power of direct consequentialism then extend to its part, act consequentialism, which is the claim that an agent ought to do an act if and only if that act has better consequences than any alternative act.

This claim itself says nothing about which consequences are best, so it needs to be supplemented by separate claims about value. Classical utilitarians held hedonistic or subjective theories of value, but most today reject hedonism, so our consequentialist theory will probably seem more plausible if it includes some values that are objective in the sense that they are independent of any conscious mental state. For example, one can have life and freedom even while one is sleeping so deeply that one has no conscious mental states at all. To make the contrast with Gert stark, I will assume Gert’s own theory that the only basic goods are pleasure, freedom, ability, and life (or consciousness), and the only basic evils are pain and losses of goods. Personally I would add other items, such as false beliefs (217n.1), and take away some items, such as life (which has no value independent of pleasure and some ability to control consciousness, pace 93); but these minor changes do not affect my points here, so I will ignore them.

This list of goods and evils does not itself tell us how to determine which is better among the total sets of consequences that result or would result from different actions. It does not even tell us which consequences count, or for whom. Act consequentialism can be held about all reasons or just about morality. There are consequentialist systems of prudence, which consider only consequences to the agent, and even aesthetics, which does not consider economic effects.

Our concern here is a moral system. Consequently, I will assume that the system is universalistic in the sense that harms and benefits to everyone count and that it is egalitarian in the sense that the harms and benefits to each count equally (in some sense). The theory can still depart from classical utilitarianism by allowing comparisons of total consequences to be affected by the distribution and mix of values within the total consequence (213). There might be some algorithm, but I do not know of any; so I will discuss a version of act consequentialism on which one total consequence is objectively better than another in a morally relevant way if the first would be preferred by all rational impartial persons. We can again follow Gert in allowing disagreements among rational impartial persons. If some rational impartial persons prefer one total consequence, and others prefer another, the total consequences are incomparable in the sense that neither is better and they are also not equal. (Compare 77–78.) This shows how some moral disagreements and conflicts can be irresolvable.

This version of act consequentialism, thus, shares many of the advantages of Gert’s own system. Insofar as Gert’s theory gains plausibility by allowing reasonable moral disagreements, as he claims, that plausibility will be inherited by this version of act consequentialism. Insofar as Gert’s theory of harms and benefits is plausible, that plausibility will be inherited by this version of act consequentialism. Insofar as Gert’s emphasis on universal public rules is plausible, act consequentialists can talk about which universal public rules morally ought to be used, taught, and enforced. This will be just one special application of their basic system. Sophisticated act consequentialism can thereby incorporate much of Gert’s own moral system. Moreover, it has the additional virtues of directness and flexibility. So I hope it seems plausible, at least initially. Anyway, this is the version of act consequentialism that I will discuss henceforth.

Gert’s Criticisms

The crucial question is whether any such act consequentialism can adequately respond to Gert’s main objections. I will argue that it can.

Too Much Impartiality

Gert’s first main criticism is that act consequentialism requires or demands too much. Gert makes this charge specifically with regard to impartiality:

it is humanly impossible for anyone to act impartially with respect to preventing or relieving the pain or suffering of all moral agents, much less to promoting their pleasure or happiness. It is absurd to claim that morality requires one to act impartially toward all moral agents with respect to the consequences of one’s action on their happiness when it is humanly impossible to do so. (136, my emphasis; cf. 123, 152)

This might seem confused. It is not literally impossible to be impartial toward all moral agents. In each case, one can use an arbitrary and hence impartial method (say, flip a coin) to choose whom to help (cf. 133). This is not far from what people do when they help those needy people whom they happen to meet or hear about. Gert might respond that we cannot be impartial all the time. However, one could donate a dollar per year to a charity that reduces suffering as much as possible with no regard to who is suffering. Then, for each portion of the year, one donates a portion of a dollar and hence does something impartial.
Gert can still respond that not all of our acts can be impartial in this way, because human nature gives us all such a strong tendency to favor ourselves and our families and friends in at least some acts. This makes universal impartiality humanly impossible.

Some act consequentialists, such as Godwin, might respond that all this shows is that human nature leads us to do what is morally wrong. After all, human nature also leads people to seek revenge even when doing so is irrational, and so on. But Gert could respond that it is humanly possible to lead a whole life without ever seeking revenge, whereas it is humanly impossible to live a whole life without ever being partial to oneself or anyone else.

The argument then seems to depend on the principle that "ought" implies "can," or at least that one must be able to avoid doing what is morally wrong. This principle could be challenged. Moreover, even if one cannot avoid ever acting partially, one can avoid each particular act of partiality. That seems to be enough to save act consequentialism from Gert's argument.

Nonetheless, I will focus on another response that is available to act consequentialists. Act consequentialists can simply distinguish what is morally required from what morally ought to be done. The principle of act consequentialism, as I and most others formulate it, determines what agents morally ought to do. Still, some acts that morally ought to be done are not morally required. An act is morally required only when failure to do that act is morally wrong. Moral wrongness in this view implies some kind of punishability. As Mill wrote,

We do not call anything [morally] wrong unless we mean to imply that a person ought to be punished in some way or other for doing it—if not by law, by the opinion of his fellow creatures; if not by opinion, by the reproaches of his own conscience. There are other things, on the contrary, which we wish that people should do, which we like or admire them for doing, but yet admit that they are not bound to do; it is not a case of moral obligation; we do not blame them, that is, we do not think they are the proper objects of punishment.

Mill's distinction between what is morally wrong to omit and what morally ought to be done, of course, parallels Gert’s distinction between moral rules and ideals. My point is that a similar distinction can be drawn within act consequentialism.

The resulting theory of moral wrongness will be indirect insofar as the moral wrongness of an act depends on the consequences of a different act (of punishment), but that much indirectness is plausible if claims of moral wrongness are used primarily for judging other people or for imagining how other people should judge us, since then the main issue is how one person should react to acts by another person. Moreover, the basic principle of what morally ought to be done can remain direct act consequentialism.

One might object that the notion of punishments and punishability are vague. Granted. But they are no clearer in Gert's theory. Gert does say that punishment must inflict evil, and he does chide Mill for including "reproaches of his own conscience" (16), but beyond this Gert never specifies what counts as punishment for him. Imprisonment and flogging are clearly punishment, but what about fines and civil damages? What if you privately or publicly rebuke your neighbor? What if a miscreant makes you refuse to hire or promote someone? What if you just refuse to invite them to a party? Are these punishments? Gert can draw a line, of course. The problem is to justify drawing the line in one place rather than another. In discussion, Gert said that punishment must be inflicted "by authorized persons in accordance with some established procedure." But what is needed for authorization? Gert sees parents as authorized to punish their children, but who authorizes them? Gert's main model seems to be law. However, what is morally wrong should not depend on what is legally punishable, because one role of morality is to determine what legal systems may and may not punish. Moreover, many moral wrongs should not be punished legally or by parents. Even if no parent or legal authority should get involved, it still might be morally wrong for me to lie to a friend. Is that friend authorized to punish me somehow? Is their mutual friend somehow authorized? Since it is not clear who counts as "authorized" or when a procedure is "established," it is still not clear what counts as punishment in Gert's system.

My own view is that we should distinguish different degrees of obligation and moral obligation according to different kinds of negative reactions that are warranted by violations, so there is no need to draw a hard and fast line between punishment and nonpunishment. Mill could agree, but Gert needs a stable line between punishment and nonpunishment in order to draw his distinction between moral rules and moral ideals. It is not clear where or how he will draw or justify that line. Anyway, however Gert eventually defines punishment, act consequentialists can adopt or adapt the same definition. Thus, if Gert can contrast moral rules with moral ideals in terms of punishability versus encouragement, then act consequentialists should be able to use the same tools to distinguish what one is morally required to do from what one morally ought to do.

Armed with this distinction, act consequentialists can agree with Gert that impartiality with regard to consequences is not always morally required (136, 152). Specifically, such impartiality is not morally required when partiality is not morally wrong or morally ought not to be punished. It morally ought not to be punished when punishment would not have the best consequences. When is that? Sidgwick argues plausibly that we should grade on a curve: "we think that moral progress will on the whole be promoted by our praising acts that are above the level of ordinary practice, and confining our censure—at least if precise and particular—to acts that fall clearly below this standard." Sidgwick admits that this line “must be inevitably vague,” but the basic idea is clear. When so many people are so partial, there are too many cases to censure them all, so it is best to pick out the worst offenders for censure. If we censure those who show only moderate partiality in cases where almost everyone else would also, this will do little good and will probably discourage the worst offenders from raising themselves up to the norm, since even then they will not escape censure. For such reasons, we should reserve censure and punishment for cases where people show too much partiality or show it in situations where impartiality is especially important, such as in courtrooms or classrooms or in other situations where Gert counts a kind of act as an unjustified violation of a moral rule. It is only in those cases, then, that partiality is morally wrong and impartiality is morally required.

Act consequentialists can still explain why very rich people might be morally required to contribute a great deal to charity, much more in absolute terms and in percentages than the middle class. There is more to be gained by inducing a single very rich person to contribute. Moreover, that person has less to lose, since a multibillionaire would still be very rich if he gave away a billion dollars. Consequently, if private or public censure would increase the millionaire’s charitable contributions and would not have too many bad side effects, then censure morally ought to be applied. The government could pass steeply graduated taxes enforced by legal punishments, but moral censure still might be useful in
addition to taxes. The billionaire could not complain that he is censured without moral wrongdoing, if he did fail to contribute enough, since whether that failure is morally wrong is the issue at stake. At least a billionaire who gives little to charity morally ought to censure himself by feeling guilty for not giving more. Those emotions of guilt would be justified by their consequences in such circumstances. Insofar as such censure and guilt feelings count as punishment, their being justified implies some degree of moral requirement for the very rich that did not exist for the middle class. Act consequentialism thereby retains much of its radical and reformative character, even while it limits what is morally required for most people.

Many details remain to be worked out, but enough has been said to show how act consequentialists can avoid the charge that they require too much (or too little) impartiality. Act consequentialists still hold that everyone morally ought to be impartial in a way that would greatly affect most people's lives, but that weaker claim does not seem so implausible, at least to act consequentialists.

**Too Much Information**

A separate but related criticism is that act consequentialism requires too much information. Gert writes,

If the relevant consequences are taken to be actual consequences, then act utilitarianism leads naturally to skepticism. No one can possibly know all the future consequences of his proposed actions, let alone all of the actual future consequences of all of the possible alternative actions, so there is no way that anyone can ever know that any given act is morally good or right. (129n9)

If one rejects such moral skepticism, act consequentialism will seem inadequate, at least "if the relevant consequences are taken to be actual consequences."

One popular response is for act consequentialists to turn to foreseeable consequences. Gert recommends this move, and it is available within the general framework of act consequentialism.

But there are reasons to resist this temptation. The turn to foreseeable consequences raises a question, "Foreseeable by whom?" Gert talks about what is foreseeable by the agent whose acts are judged, but that produces problems. If you do not know and have no way of knowing that your act of keeping a trivial promise will cause many deaths, but an observer does know this and also knows that this act will have no benefits for anyone, then that observer should not say (to you or to another) that you morally ought to do this act. If possible, the observer should tell you that you morally ought not to do it. This "ought" is moral because it is based on harm to others, rather than harm to self, economic benefit, law, religion, or anything else apart from morality; and you are able to respond to advice from the observer. Of course, if you do kill those people because of a reasonable mistake, then you have an excuse, you are not responsible, and you should not be blamed or judged to be a morally bad person, because you had no way of knowing that you were killing. Such moral judgments about responsibility and persons, then, might depend on whether consequences are foreseeable. Also, of course, agents have to decide on the basis of consequences that they can foresee. But judgments about the moral rightness or wrongness of acts, when made by observers, do not seem restricted to what was foreseeable by the agent.

Moreover, Gert's decisive question asks about the consequences of publicly allowing all acts of a kind. The relevant kind of act is determined by what is foreseeable to the agent of that act, so different agents can do the same kind of act only if they are able to foresee the same consequences of their particular acts. But different agents still might foresee and be able to foresee very different consequences of publicly allowing all acts of that kind. Thus, whether it is morally wrong for an agent to do an act of that kind cannot be determined by whether that agent can foresee that publicly allowing all such acts would have worse consequences than not publicly allowing all such acts. If it did so depend, then acts of the very same kind would be morally wrong for one agent but right for another. Active euthanasia, for example, would be morally wrong for some doctors but not for others, even when the doctors are able to foresee the same consequences of the acts. That is a result that Gert wants to avoid. So his decisive question, at least, cannot refer to effects that are foreseeable by the agent.

Gert might respond that what matters is what is foreseeable by the observer making the moral judgment. But that leads to other problems. There could be several observers. Suppose a first observer knows that your act will cause many deaths, but a second observer cannot foresee these or any bad consequences of your act. Accordingly, the first observer says that your act is morally wrong, and the second observer says that your act is not morally wrong. Both of these moral judgments would be correct if what is morally wrong depended on the consequences that are foreseeable by the person making the moral judgment. But these judgments cannot both be correct (even though both might be justified, given different empirical situations of the two observers). Consequently, act consequentialists should not make what is morally wrong depend on what is foreseeable by the person making the moral judgment. This leaves no satisfying answer to the question "Foreseeable by whom?"

Other problems arise when we ask when a consequence is foreseeable. Does it have to be foreseeable given what the person actually believes or what the person should believe or should we subtract false beliefs or just believe the person should know are false? Different theories give different answers, and no answer fits all of our intuitions.

It is hard to see any way to avoid such problems except by making what is morally wrong depend on actual consequences (or some equivalent, such as an ideal observer). This conclusion might, however, seem problematic when applied to retrospective moral judgments. Suppose again that you have no way of knowing that your act will cause many deaths, and nobody warns you, so you do it, and it does cause many deaths. According to act consequentialists who count actual consequences, your act was morally wrong. Gert claims that no normal speaker would say this. I am not so sure whether this is so or why.

Gert argues that such retrospective judgments would be like judgments of nonhuman animals, but we do not make moral judgments about nonhuman animals (21-22). However, act consequentialists can accept that we do not classify a judgment as moral unless it is about humans. That fact about common usage might merely reflect the inability of nonhuman animals to follow advice given in moral judgments. After all, we also do not advise animals to brush their teeth. In contrast, even when a normal adult human has no way of knowing that an act will cause many deaths, such a human is still capable of responding to advice if an observer says that the act will cause deaths and that the human morally ought not to do it. Even after the act is done, a human is still the kind of agent who is normally responsible and subject to moral judgment, and who was capable of responding to advice if any had been given in advance. So, even if you could not have foreseen the deaths, you are not like an animal in any way that need preclude retrospective moral judgments of your acts.
Moreover, after you cause the deaths that you could not have foreseen, you and your victims wish that someone had warned you, since your act turned out to be the wrong thing to do. Its wrongness was not prudential or economic or aesthetic or religious or anything other than moral. What makes it wrong is harm to others without adequately compensating benefits, as in standard cases of moral wrongness. Such considerations might lead some to call your act morally wrong. Of course, they should immediately add that you were not responsible, blameworthy, punishable, or bad as a person; but the moral wrongness of your act is a separate matter.

The reason why speakers are so reluctant to call such acts morally wrong might be just that we usually make such retrospective judgments in order to blame agents, so an audience might be misled into thinking that the agent is blameworthy or a bad person. This explanation can be supported by the oddness of saying that your act was morally right when you killed so many people. For different reasons, then, we usually avoid expressing either moral judgment out loud. But it still might be correct that your act was morally wrong, even if nobody likes to say it.

The other problem is that actual consequences go on forever, so they seem unknowable, and this version of act consequentialism seems to lead quickly to skepticism about moral knowledge. (Recall 129n9.) This implication would hold if knowledge required certainty. One cannot ever be certain about the long-term consequences of any particular act, so, if moral wrongness depends on those consequences, one cannot ever be certain that any particular act is morally wrong. But that implication does not seem absurd, if one does not forget that one still can know that the agent is punishable on the basis of what was intended or foreseeable.

Moreover, agents and observers can still have very good reasons to believe that the overall consequences of an act will be good or bad. It is extremely unlikely that the overall consequences would be good if I killed my neighbors right now. It is possible that these deaths will lead to world peace, but that is unlikely. There is at least as much reason to think that they would lead to world war. Thus, although I cannot be certain, I can be justified in believing that my killing my neighbors now would be morally wrong, according to act consequentialists who count actual consequences. In other cases, where it is harder to justify beliefs about overall consequences, it is harder to be justified in believing that an act is morally right or wrong. But that just shows that life is tough. No news there. And no problem for actualistic act consequentialism, since any plausible moral theory should admit that we often cannot tell what is morally right or wrong.

Even if it still seems implausible to count actual consequences, it does not matter for my purposes whether act consequentialists count actual consequences or some kind of foreseeable consequences. So here I will not take either position. Since Gert counts foreseeable consequences and usually assumes that act consequentialists do the same, I will write as if act consequentialists count foreseeable consequences. Any remaining differences with Gert must then be due to other factors. Readers who favor actual consequentialism must still reframe my points in terms of actual consequences. Whichever version of act consequentialism seems best, the question here is whether that version is as plausible as Gert’s moral system.

To compare these systems, we need to determine whether Gert’s moral system avoids the epistemic limits of consequentialism. He claims that “everyone who is judged by it knows what morality prohibits, requires, encourages, and allows” (4; cf. 7, 10, etc.). This is too strong as it stands. One problem is that Gert’s second morally relevant feature requires us to determine the consequences of each particular act in order to determine what kind of act it is (228). The fifth, sixth, and eighth morally relevant features also refer implicitly or explicitly to consequences (231-33). Gert tells us that he refers to foreseeable consequences of acts and their alternatives and public allowings. However, an agent might not actually foresee or know consequences that she is able to foresee or know. Then the agent will not actually know what kind an act is or whether that act is morally permissible.

Gert might respond that at least the moral status of the act is knowable because its consequences are foreseeable. However, even if he can solve the above problems about foreseeability, this weaker claim does not fit with his limitation to rationally required beliefs and concepts in his moral theory. Many beliefs and concepts that are not actually shared by all rational people, nonetheless, could be held by all rational people. This includes much science, if what could be known is broad enough. It is not clear why those additional beliefs and concepts should not be used in moral theory if agents are judged not by what they do foresee but by what they could foresee.

Even if this tension can be resolved, Gert faces worse epistemically problems regarding his morally decisive question, which asks about the effects of a kind of violation being publicly allowed (236). This speculative counterfactual question about the consequences of a practice that we do not follow in a world that we do not inhabit will be harder to answer than the act consequentialist’s question (especially if the act consequentialist counts only foreseeable consequences of particular acts). Gert seems to assume that it is easy to extrapolate from what would happen in our society to what would happen if we started to think differently about what we are allowed to do. This is not at all easy, because so many other social circumstances might change along with this change in our public allowings. For example, if we all viewed ourselves as allowed to break laws when doing so causes risks and harms, then this might undermine “the order and stability that is essential for any society to function well”, as Gert claims (208). Or it might lead to very few and minor illegal acts. There need be no increase in murder, rape, theft, or any harmful or dangerous acts, since those acts are still ruled out. Some people would drive through stop signs on clearly deserted streets, but other drivers would know to watch more carefully. This new attitude to law might even lead to reforms of the legal system, including better enforcement to increase risks of punishment and better laws in areas where breaking the law harms nobody. All of this is speculation, but it can be countered only by more speculation or ideology (238). So justified beliefs about moral rightness and wrongness are at least as hard to obtain on Gert’s version of act consequentialism as on act consequentialism.

Gert will probably respond that his moral system has the advantage of concrete rules, whereas act consequentialism cannot be applied without detailed examination of each situation. This might seem to make act consequentialism useless as a guide to conduct in those cases where we cannot form justified beliefs about consequences. Moreover,
should not appeal directly and self-consciously to consequences alone when they decide what to do in most everyday situations. Act consequentialist conditions of moral rightness might be useless and even dangerous if misused as a direct practical guide. But that is not their intended use, and these dangers of misuse do not show that such conditions of moral rightness are useless (pace 253). They can still be useful in determining which practical guide is best. In choosing a guide for everyday use, act consequentialists will have to consider the fallibility of humans. Gert criticizes act consequentialism for "failing" to recognize that the fallibility of persons is an essential presupposition of morality (265-6; cf. 245). But the most that they do is refuse to build fallibility directly into their conditions of moral rightness. Their practical guide still can and should be tailored to reflect common mistakes and also to help people work with the information that is available to them in various situations. By providing such a practical guide that is separate from its conditions of moral rightness, act consequentialism thereby avoids requiring too much information.

Act consequentialists can even incorporate publicity and universality, since they ask which rules should be publicly announced and taught as rules for all people to use self-consciously, that is, as public guides. The results might be very close to the content Gert's moral system.

But there will still be differences. Suppose two people get married. Each wants both to feel free to have sexual relations with others, so their marriage vows include no promise of sexual exclusivity. Later such does have sex with others and tells the spouse. This does not bother either of them, because they know that they really love each other, regardless of how each gets physical pleasure on the side. They do not have children who could be harmed by their activity; and they keep it secret from their neighbors, because they know that their neighbors and most of their society disapprove of adultery, and they do not want to upset anybody unnecessarily. They do not care whether other marriages are like theirs, but they like theirs the way it is. Nonetheless, Gert would say that their act of adultery violates his moral rule against cheating if the institution of marriage in their society requires sexual exclusivity, as in modern Western societies, according to Gert (196-97). His point cannot be that this couple violates any explicit or implicit promise, nor that they deceive others by representing themselves as having promised sexual exclusivity, since he insists that his rule against cheating is independent of his rules governing promises and deception (192, 208-9). Gert's claim is that this couple's acts of adultery are morally wrong just because publicly allowing adultery like theirs would undermine the existing institution of marriage, that institution is generally beneficial, and their only reason to violate it is to get sexual pleasure.

Old hippies and libertarians might respond that here Gert takes publicity too far. To these critics, it seems unfair to let the dominant public institution determine the moral status of their private sexual acts. Act consequentialists can agree with these critics insofar as, even if a public prohibition on adultery has better consequences generally, this unusual couple's sexual acts are not morally wrong, when those acts have the best consequences among their alternatives. Whether or not you agree, the point here is just that act consequentialism remains at least plausible and distinct from Gert's system.

What enables act consequentialists to adopt such a position is that they can separate conditions of moral rightness from their universal public guide. Gert confines these stages into one when he claims that the universal public guide determines what is morally right and wrong. This conflation prevents Gert from drawing distinctions that others can. In this respect, act consequentialists can
provide everything Gert does, and more.

Indeed, act consequentialists can provide many practical guides. One question asks which practical guide has the best consequences when everyone knows that everyone knows that everyone follows it. Less universal but no less public practical guides might, however, be best for certain societies or certain groups within society, such as doctors. A slightly different practical guide might have the best consequences if followed by a given individual. Opponents might scorn this proliferation of practical guides or any nonpublic or nonuniversal morality. (215), but act consequentialists can respond that their theory thereby recognizes more of the levels and complexities of our moral lives (pace 6).

Even if this flexibility is not seen as an advantage, at least the two approaches seem to be at a standoff. Gert seeks direct practical guidance or rules for all to follow publicly and self-consciously. Act consequentialists instead seek to specify conditions of moral rightness. Gert cannot fairly complain that they do not give him what he wants, if they are not trying to do that. Gert’s project is not the only worthwhile project in moral theory. If act consequentialists are up to something else, that cannot be used against them.

Too Little Common Sense

Some of you are probably feeling impatient. Haven’t I left out the most obvious and strongest criticism of act consequentialism? Yes, but no longer.

The most common criticism of act consequentialism is that it has absurd implications in particular cases. It is supposed to license harming one person to benefit others, which is supposed to violate the victim’s rights. Many examples have been given, including proverbial sheriffs hanging innocent people to stop riots, doctors cutting up one healthy person to save five dying patients, and so on.

Gert agrees that such acts are morally wrong, but he argues that sometimes it is permissible to kill an innocent person in order to prevent more harm to more people (124-25). So Gert’s main counterexample to utilitarianism is different: “It is a universally accepted criticism of some forms of utilitarianism that it would allow the infliction of pain on one person in order to promote a great amount of pleasure for many others if a sufficiently large number will receive the pleasure” (124; cf. 15, 231, 373). One common example is Romans forcing slaves to fight lions as entertainment for citizens in the coliseum stands.

It is not clear that act consequentialists must allow any such act. Even hedonistic act utilitarians can argue plausibly that there will always be some better form of entertainment with less cost in the long run. Sophisticated act consequentialists can also count the freedom and life lost by the slave; and they can assign relative values so that these losses always override any entertainment value. Act consequentialists can also count distribution in a way that magnifies losses to those previously deprived, and they can discount pleasures at pain to others. Such moves enable sophisticated act consequentialists to minimize divergence between their view and common morality.

Gert will insist that act consequentialism still diverges from common morality in some examples. One of his favorites involves cheating on a medical school final exam (cf. 203). To ensure that cheating has the best consequences and the student can foresee this, Gert needs to add details: the student must be able to know that more harm will result if he does not cheat (because there is no other way to avoid failing, and failing will hurt his career and his parents, who are doing and dying). He must also be able to know that no less harm will come if he does cheat (because he won’t be caught or feel too guilty, he won’t become an incompetent doctor or a bad person, and other students won’t be harmed in later competition or because these grades are curved). The student must also be able to know that more good will come if he does cheat (since he will pass and be happy, his parents will be very proud of him, and he will get his degree, start his career, help many patients, and live happily). If enough details are spelled out properly, cheating will have the best consequences, and the student will be able to know this, so act consequentialism will imply that the student morally ought to cheat.

Gert claims that this implication of act consequentialism conflicts with common morality and also with his and most moral intuitions. He backs up this moral intuition with an argument from public morality: rational impartial people would never advocate publicly allowing cheating in circumstances like these, because publicly allowing this kind of act would make impossible all restricted tests or would lead to too many abuses, as students make mistakes about the consequences of cheating and its alternatives.

For act consequentialists could respond by adjusting their theory of value so that cheating does not have the best consequences in this case, or they could argue that it is not really countermoral to hold that cheating is morally permissible in such extreme circumstances. However, I would like to explore a third possible response, namely, that, even though act consequentialism is countermoral in such examples, that does not refute it.

The crucial question for this response asks why it is so bad for a theory to be countermoral. At first, Gert answers this way: “it would not count as justifying morality unless the code of conduct being justified was virtually identical to the moral system that is now implicitly used in deciding how to act morally and in making moral judgments” (7). This test has to be hedged by the term “virtually”, because no moral progress is possible without some change, and also because Gert’s own theory is not completely identical with the moral system in current use. Once that hedge is inserted, however, it is not clear why act consequentialism fails this test. As Sidgwick argued in some detail, act consequentialism agrees with common morality in a wide range of cases. There is even more convergence with sophisticated versions of act consequentialism. Partly because secrets are hard to keep, better consequences will almost always be produced by following common morality or Gert’s rules, except when violations are justified by Gert’s procedure. So these systems are “virtually” identical, unlike Nietzsche’s system (17-18).

These systems still differ somewhat, so we need to ask whether a theory is refuted if it diverges at all from common morality. Gert himself goes on to say, “Any divergence counts against the adequacy of the generated system and ultimately against the moral theory that generated it. This is precisely what happened to utilitarianism” (7). Here Gert claims that every divergence counts, but how much? Act consequentialism has virtues that might override some minor divergence from common morality, so its critics also need to show that its countermoral implications are numerous and important enough to refute it.

Gert and other critics do claim this, but why? Maybe because the systems diverge in so many cases. One book lists over a hundred counterexamples to utilitarianism (although many do not apply to sophisticated act consequentialism, and many that do apply are questionable). In principle, act consequentialism diverges from common morality and from Gert’s system in an infinite number of possible cases. The raw number is not decisive. Act consequentialists can respond that divergence occurs only in cases that are recognizably abnormal
both statistically and in the sense that they are emergencies, a feature that Gert admits to be morally relevant (235).

More importantly, the rare cases are unrealistic. Cheating seems to have the best foreseeable consequences in Gert's case only because the student knows more than any real student could ever really know or justifiably believe. As the number of convictions shows, students do not know as much as they think they know about whether they will be caught. As Kaskokinov found out in Dostoevsky's Crime and Punishment, it is also hard to tell how guilty one will feel, or how such an act will affect one's character and future actions. Similarly, when critics talk blithely about cutting up one person to save five, they forget that real transplants require many staff members, any one of whom could leak information that would ruin the doctor and the hospital. Such counterexamples depend on secrecy, but secrecy is fragile in real life. These critics also forget that doctors never know whether transplants will work, that often another treatment might work without costing lives, that such acts affect the agents' characters and future actions, that it would be a massively unlikely coincidence if five patients all happened to be compatible with a complete stranger who happens to walk into the hospital, and so on. Such details are essential and justify calling such counterexamples unrealistic.

These examples are still logically possible. If act consequentialism is supposed to hold necessarily, such cases refute it if the acts in such examples are morally wrong. So why does it matter that these examples are unrealistic?

This lack of realism matters because the arguments against act consequentialism rest on moral intuitions that these acts are morally wrong. Our moral intuitions evolved to help us deal with normal and realistic cases. They were then shaped by social indoctrination and teaching that was also aimed at the most common kinds of cases. Only philosophers worry about unrealistic situations. Given their origin, it is not at all clear why moral intuitions should be trusted in rare and unrealistic cases. If moral intuitions should not be trusted in such cases, then they cannot refute act consequentialism.23

These cases still might leave a dirty taste in one's mouth. No teacher likes to say that it is morally permissible to cheat. No doctor likes to say that it is morally permissible to cut up one person to save five, even in unrealistic circumstances. However, act consequentialists can easily explain why we are and should be reluctant to publicly announce and even think that such acts would be morally permissible. If we announce these permissions to others, or if we think them to ourselves, then we or our audience will be more likely to make mistakes and cheat or cut up when such acts are not morally permissible, which is all or almost all of the cases that we will really encounter. Our speech acts and thoughts affect our actions, so speech acts and thoughts can be evaluated by their consequences. But those consequences of speech and thought still need not reveal the truth of those claims and thoughts or the moral status of the acts that those claims and thoughts are about. So this explanation is compatible with act consequentialism.

This response would not be available if act consequentialists proposed their system as a public guide. But we already saw that what act consequentialists propose is conditions of moral rightness. Act consequentialists can agree that no acceptable public guide would ever allow students to cheat or doctors to cut up healthy patients. Nonetheless, even if no act consequentialist moral system should be publicly announced or taught in medical schools, acts that are forbidden by the best public guide still might be morally right in abnormal emergencies where those acts have the best consequences. That is all that act consequen-

---

**Notes**


3. This formulation dominates earlier editions without the morally decisive question, such as Bernard Gert, *The Moral Rules* (New York: Harper and Row, 1966), 92.


5. Austin might respond that some acts that maximize utility also belong to other classes whose general instantiation would reduce utility. But that is true of any act that causes any harm, since, if “causing harm” were “generally done,” “the probable effect on the general happiness” would be bad. So this cannot show that an act is morally wrong. For more on reduction, see David Lyons, *Forms and Limits of Utilitarianism* (Oxford, Eng.: Clarendon Press, 1965).

6. Gert’s emphasis on public permissive rules is shared by John C. Harsanyi, “Moral Theory of Rational Behavior,” reprinted in *Utilitarianism and Beyond*, ed. Amartya Sen and Bernard Williams (New York: Cambridge University Press, 1981), 39-62. On page 59, Harsanyi says that rule utilitarians will always have to ask the question, “What would be the social implications of adopting a moral rule permitting that promises should be broken under conditions A, B, C, etc.—assuming that all members of the society would know that promise breaking would be permitted under these conditions?” Thanks to John Gert for this reference.

7. In personal communication, Gert does cite consequences as reasons to include some features but not others on his list of morally relevant features.


12. Quoted by Gert on his pages 15-16 from John Stuart Mill, *Utilitarianism*, ed. Roger Crisp (Oxford, Eng.: Oxford University Press, 1998), chapter V, paragraph 14. Mill also calls punishability the “real turning point of the distinction between morality and simple expediency.” Nonetheless, I will assume that some acts morally ought to be done even though they are not morally required and omitting them is not morally wrong. Also, if Mill claims that an act is morally wrong whenever it ought to be punished, then sometimes one morally ought to do what is morally wrong. To avoid this odd implication, I will assume that Mill means to claim that an act is morally wrong if and only if it both morally ought not to be done and also morally ought to be punished (if the agent is responsible and punishment would not be impractical).


14. These judgments about excuses, blame, condemnation, and responsibility show some ways in which act consequentialists can “recognize the fallibility of humans” despite what Gert says (205-6; cf. 245). We will see another way below.


17. This distinction is implicit in Mill and Sidgewick, but recent discussions seem to derive from R. E. Bales, “Act-utilitarianism: Account of Right-making Characteristics or Decision-making Procedure?” *American Philosophical Quarterly* 8 (1971): 257-65. I avoid the more common phrase “decision procedure,” because this is often taken to imply completeness, and Gert denies that there can be any complete moral decision procedure (11). I also refer to conditions instead of criteria, because conditions of oughtness and rightness need not be as easily accessible as Wittgensteinian criteria. Finally, there can be conditions and guides for both moral oughtness and moral wrongness. I will focus on moral wrongness, because here I am answering objections that are usually formulated in terms of what is morally wrong.


20. Act consequentialists do not even refuse to do this if they count only foreseeable consequences. They can also let fallibility affect standards of responsibility, as I mentioned in note 14.


25. For helpful comments on drafts, I thank Robert Audi, Marcia Baron, Julia Driver, Shelly Kagan, David Phillips, Bill Throop, and, especially, Bernie Gert.