

Contemporary Consequentialist Theories of Virtue¹

Ben Bradley

The theory of virtue has recently been dominated by Aristotelians. This is unsurprising given Aristotle's status in the profession, the extensive treatment of virtue found in the *Nicomachean Ethics*, and the centrality of virtue to Aristotelian ethics (as compared to its subsidiary status in e.g. the utilitarian tradition). But from the standpoint of the philosophical understanding of virtue, Aristotle's dominance is arguably lamentable, since the "Doctrine of the Mean" is virtually devoid of content (Aristotle 1976, Book II). Consequentialism about the virtues offers a potentially more informative alternative. Roughly speaking, consequentialist theories of virtue explain a character trait's status as a virtue or vice by appealing to the value of the consequences of the trait, unlike the dominant Aristotelian and Kantian views according to which an agent's virtue is determined largely or entirely by the intrinsic quality of her psychological states. I will begin by tracing the development of virtue consequentialism and spelling out some ways to develop a consequentialist theory of virtue. Then I will discuss some of the advantages and shortcomings of consequentialist theories.

I. The Development and Varieties of Virtue Consequentialism

The roots of virtue consequentialism can be found as long ago as 1689 in Locke's *Essay Concerning Human Understanding*. According to Locke, God has "by an inseparable connexion, joined *Virtue* and publick Happiness together; and made the Practice thereof, necessary to the

¹ Thanks to Dave Sobel and Nancy Snow for helpful discussion and comments.

preservation of Society, and visibly *beneficial* to all, with whome the Virtuous Man has to do” (Locke 1975/1689, I.iii.6). We can find virtue consequentialist ideas also in Frances Hutcheson (1991/1725, 271) and David Hume (1975/1777, 181). Here is a nice statement from Hutcheson’s *Inquiry concerning the Original of our Ideas of Virtue or Moral Good*: “these four qualities, commonly called *cardinal virtues*, obtain that name, because they are dispositions universally necessary to promote *public good*, and denote *affections* toward *rational agents*; otherwise there would appear no *virtue* in them. (Hutcheson 1991/1725, 271). In his *Enquiry Concerning the Principles of Morals*, Hume claims that good consequences are at least part of the explanation of what makes a character trait a virtue, or at least what makes benevolence a virtue: “upon the whole, then, it seems undeniable, that nothing can bestow more merit on any human creature than the sentiment of benevolence in an eminent degree; and that a part, at least, of its merit arises from its tendency to promote the interests of our species, and bestow happiness on human society” (Hume 1975/1777, 181).

A closely related view, which Robert Adams calls “Motive Utilitarianism” (Adams 1976), can also be found in the utilitarian tradition, in Jeremy Bentham and Henry Sidgwick: “If [motives] are good or bad, it is only on account of their effects: good, on account of their tendency to produce pleasure, or avert pain: bad, on account of their tendency to produce pain, or avert pleasure” (Bentham 1907/1789, X.12); “the doctrine that Universal Happiness is the ultimate *standard* must not be understood to imply that Universal Benevolence is the only right or always best *motive* of action... if experience shows that the general happiness will be more satisfactorily attained if men frequently act from other motives than pure universal philanthropy, it is obvious that these other motives are reasonably to be preferred on Utilitarian principles” (Sidgwick 1981/1907, 413).

None of these views is quite the view we are after. Locke's claim that the virtues are necessary to the preservation of society seems overblown; couldn't a virtue be good for people, and thus a virtue on consequentialist grounds, without being necessary for there to be a society at all? Hutcheson's view seems too weak; many dispositions might be *necessary* for bringing about good outcomes without being *sufficient* to do so. Hume endorses at best a partial version of virtue consequentialism, since he claims merely that benevolence is the greatest virtue (not the only one), and that its status as a virtue is at least *partly* explained by its consequences. Bentham and Sidgwick are consequentialists about *motives*, but motives and character traits are not quite the same thing. We must look just a bit later than Sidgwick to find a real statement of virtue consequentialism.

A clear statement of a virtue consequentialist view is found in G.E. Moore's *Principia Ethica*: "a virtue may be defined as an habitual disposition to perform certain actions, which generally produce the best possible results" (Moore 1903, 172). We will take Moore's view as our starting point in formulating varieties of virtue consequentialism. Moore's view about virtue is adapted from his consequentialism about moral duties, according to which an act is morally required if and only if its consequences are better than those of any of its alternatives (Moore 1903, 148). Moore's view is that virtues are dispositions to perform actions we are morally required to do.

Moore's view is only one of at least two ways to adapt consequentialism about moral requirement into consequentialism about the virtues. Another would be to say that a character trait is a virtue if and only if its possession maximizes utility, i.e. possessing that character trait has consequences at least as good as possessing any alternative character trait.² This view is

² This view is suggested by a version of "character utilitarianism" formulated by Peter Railton: "an act is right just in case it would be done by someone having a character, the general possession of which would bring about at least as

preferable to Moore's. On Moore's view, it could turn out that there are no virtues, since it could turn out that there simply are no character traits that always or generally dispose the bearer to perform acts with better consequences than any alternative. Furthermore, suppose there are two character traits with the following features. The first disposes the bearer to perform utility-maximizing actions in situations where the choice is of little consequence, but in certain rare but very important situations disposes the bearer to perform acts with terrible consequences. The second *always* disposes the bearer to perform actions that are very nearly, but not quite, as good as the best alternative. The virtue consequentialist should rate the second trait higher. It has better consequences overall, even though it is not a disposition to perform the best option.

So here is a first pass at formulating virtue consequentialism:

Maximizing Virtue Consequentialism: a character trait is a virtue if and only if its consequences are at least as good as any alternative character trait.

MVC employs the notion of an "alternative" character trait. Let us understand this as follows: two traits are alternatives if and only if it is impossible to have both at the same time.

One important feature of MVC is that whether a trait counts as a virtue is not determined in isolation. It is determined by a comparison with alternative traits. This is a departure from many other views about virtue; for example, if virtue were a disposition to love the good, then a trait's status as a virtue would be independent of facts about other traits.

Understanding virtue consequentialism in a maximizing way is probably a mistake. It is overly demanding; it entails that any character trait that is less than the best possible trait is not a

much utility as any alternative" (Railton 2003, 227). This is a view about morally right action, but it presupposes a maximizing consequentialist view of the virtues. See Sinnott-Armstrong 2014 for a similar formulation.

virtue, which seems wrong (Adams 2006, 53). Among consequentialists about permissible action, it is controversial whether an act must maximize utility in order to be permissible. But maximizing views at least have some plausibility in the context of permissible action since permissibility does not come in degrees: either an act is permissible or it is impermissible, and arguably it does not make sense to say that one act is *more* permissible than another. Virtue is not like that. One trait can be more of a virtue than another. MVC gives us no way to evaluate whether one trait is more of a virtue than another if neither is maximally good.

Since virtue seems to be scalar, the virtue consequentialist should perhaps formulate her view accordingly:

Scalar Virtue Consequentialism: one character trait is more of a virtue (vice) than another if it has better (worse) consequences.³

SVC, however, does not tell us which character traits are virtues; it does not specify a threshold separating the virtues from the vices. One trait could be more of a virtue than another – could be higher on the virtue scale – even though both are vices. We would need an additional principle to specify this threshold, and MVC has already been ruled out.

We might wish to say that a character trait is a virtue if it has *good* consequences (not necessarily better than any alternative – just good). Julia Driver's version of virtue consequentialism takes this route. According to Driver, a trait is a virtue if and only if it has on balance good consequences (Driver 2001, 82). The question is, what does this mean? It cannot mean that its consequences are on balance good *in themselves*, for there are virtues whose value

³ This view is similar to views proposed by Railton and Adams (Railton 2003, 228; Adams 1976, 470), though neither Railton nor Adams provides any motivation for formulating the view in a scalar way.

comes not from bringing about what is intrinsically good but rather from *preventing* what is intrinsically *bad*. This is especially important in cases where background conditions render it impossible for any trait to have consequences that are on balance intrinsically good, and the best traits will merely make things less bad for people (Bradley 2005).

In attempting to identify a sub-optimal threshold trait, from among the set of alternative traits one might have, such that any trait that has better consequences than that trait is a virtue and any that has worse consequences is a vice, we run into the same sorts of well-known problems that consequentialists encounter when attempting to formulate “satisficing” versions of consequentialism about permissible action. For example, where is the threshold to be drawn, and why? Such questions will be no easier to answer concerning character traits than concerning actions.

The best route for the virtue consequentialist is to say that there simply is no *absolute* fact of the matter about whether any character trait is a virtue or a vice, even in a given society at a given time. To call a trait a virtue is to say that it is better than some alternative character trait you are comparing it with, and in different conversational contexts we may have in mind different alternative traits. This contrastive view gives us a way to provide truth-conditions for statements attributing virtue status to traits, without requiring us to determine a threshold. Just as we can truly call someone “tall” even though there is no absolute threshold of tallness, we can truly call a trait a virtue even though there is no absolute threshold of virtue (Bradley 2005; Bradley forthcoming).

Questions remain about which consequences we are to consider in evaluating a trait. When comparing two traits, it might turn out that one has better consequences if possessed by a particular individual, while the other has better consequences if possessed by all of humanity or

if generally possessed by members of a given society. For example, a certain degree of honesty might have good consequences if possessed by society generally but bad consequences if possessed by a diplomat. If so, is honesty a virtue or a vice for the diplomat?

On the one hand, it may be unintuitive to say that what is a virtue for someone depends only on the consequences of that person's having it. As Robert Adams says, "an industrialist's greed, a general's bloodthirstiness, may on some occasions have better consequences on the whole than kinder motives would... But we want to say that they remain worse motives" (Adams 1976, 480). We may want to say this even if the industrialist's greed has better consequences than kinder motives, not just on some occasions, but throughout the industrialist's life. On the other hand, it may seem odd to say that whether it is a virtue for one person to have a character trait depends at all on the consequences of some *other* person or people having it. Furthermore, as Adams notes, it seems we should leave room for there being a necessary or useful *diversity* of virtuous motives in a population (Adams 1976, 480). Perhaps it is best for the virtue consequentialist to say that virtues and vices are relative to an individual or a population, so that it is at once both a virtue and a vice for the diplomat to be honest, relative to different populations (Moore 1903, 173; Bradley 2005, 289).

Another question is whether we are to look at the consequences of a trait in counterfactual situations in order to determine whether it is actually a virtue. The concern here is the extent to which luck might determine whether a trait is a virtue. We might wish to refrain from calling a trait a virtue if its actual good consequences are merely an accident, and therefore look at its consequences across a range of realistic but non-actual situations to determine whether it is a virtue.

Driver argues that we need not look to such consequences. According to Driver, the consequences of a trait across actual situations are sufficient to determine its status as a virtue or vice, and looking at counterfactual situations yields some wrong results. Here is her main argument:

Suppose, for example, that Sally would have had bad traits if she had not been raised by her mother, who, it turns out, did raise her only through amazing luck – the mother was almost run over by a truck but avoided death through an amazing fluke. Well, in worlds very close to this one, Sally is a bad person. Her high spirits become disruptive... Sally's high spirits are still virtuous. (Driver 2001, 81)

Here it may be helpful to recall that for the virtue consequentialist, a virtue may be relative to a population or to a person. Even if high spirits have bad consequences when possessed by Sally in nearby worlds, they might have good consequences when *generally* possessed, and thus might be virtuous relative to the general population. Still, we might want to say that it is virtuous *for Sally* to have high spirits given their actual good consequences when possessed by her. The counterfactualist can accommodate this thought by pointing out that which worlds are “closest” to the actual world is not an absolute matter (Lewis 1986). Worlds in which Sally's mother is run over by the truck, Sally is raised badly, and so her high spirits have bad consequences, differ from the actual world in some ways; worlds in which Sally's mother is not run over, Sally is raised well by her mother, but Sally's high spirits nevertheless accidentally have bad consequences despite her mother's good guidance, differ from the actual world in other ways. Which worlds count as “closer” depends on which similarity relation we employ.

One consideration that might lead us towards a counterfactual view is that if actualism were true, then any character trait that failed to have any actual consequences would fail to be a virtue – even if the reason it has no actual consequences is that nobody possesses the trait (Bradley 2005, 293). It is odd to say that a trait is not a virtue merely because nobody has it; it seems natural, for example, to lament the fact that nobody has a certain virtue. If actualism were true, we could at best lament the fact that nobody has a trait because the trait *would be* a virtue if someone were to have it.

Consequentialism about morally permissible action comes in objective and subjective varieties. Objective consequentialism is the view that an act is morally permissible iff it maximizes utility, while subjective consequentialism is the view that an act is morally permissible iff it maximizes *expected* utility. The expected utility of an act is determined by the values of the consequences that *might* result from the act, weighted by the *subjective probability* that those consequences would result. Subjective probability is determined by the credences of the agent performing the act. The move to subjective probability is again motivated, in part at least, by concerns about moral luck. If an agent takes a great gamble and brings about the best possible consequences only by luck, while knowingly risking calamity, the subjective consequentialist can part ways from the objective consequentialist and give what seems to be the correct account of the situation: the agent acts wrongly. We might wish to formulate virtue consequentialism in a subjective form, so as to minimize the extent to which a trait may be a virtue as a result of luck. However, it is unclear how such a view could be formulated. Whose subjective probabilities would be used, and at what time? In evaluating the permissibility of an action, the answers to these questions are: the agent of the act, at the time it is performed. Such

answers are not available when evaluating a trait, since it is possessed by many people and exemplified at many times.⁴

Finally, virtue consequentialism may be pure or impure (or partial). I have focused on pure versions; but one might wish to say that while the values of the consequences of a trait are part of what make it a virtue, other features are also relevant. A common modification to the basic virtue consequentialist idea is the notion of virtue as a corrective. According to Philippa Foot, virtues “are *corrective*, each one standing at a point at which there is some temptation to be resisted or deficiency of motivation to be made good” (Foot 1978, 8). This idea is widely endorsed throughout the history of virtue theory, including for example by Bernard Mandeville (1991/1714, 233), G.E. Moore (1903, 172), Franz Brentano (2009/1952, 309), and G.H. Von Wright 1963, 147-150). The suggestion here is that we should add a clause to virtue consequentialism so that in addition to having good consequences (however this is understood), a virtue must also involve resisting temptation, or controlling natural inclinations to bad behavior. Adopting this more complicated proposal would help rule out certain traits as virtues. For example, suppose that in a certain society, people can bring about good consequences merely by acting in their own narrow self-interest. By engaging in greedy or wasteful consumer behavior, people would generally improve economic conditions so that everyone is better off than if people were generally more altruistic. On the simple version of virtue consequentialism, these people’s greedy dispositions would count as virtuous; but since they would not be resisting any temptation by acting in these ways, such dispositions would fail to count as virtues on the corrective

⁴ See Bradley 2005, 295 for an attempt to formulate such a view; also see Adams, who suggests that the most plausible version of “motive utilitarianism” will appeal to something he calls “average probable utility” (Adams 1976, 480), which sounds like an expected utility view.

account.⁵ But perhaps such a view is too Kantian for the consequentialist. If being helpful and truthful has good consequences, why wouldn't it be virtuous to be helpful and truthful, even for someone who found it easy or even pleasant to be that way?

Another notable theory of virtue that may be regarded as consequentialist is the neo-Aristotelian view defended by Rosalind Hursthouse (1999). According to Hursthouse, what makes a trait a virtue is that it promotes the agent's survival, the survival of the agent's species, the "characteristic" enjoyment of the agent, and the functioning of the agent's social group (Hursthouse 1999, 202). It seems fair to call Hursthouse a kind of virtue consequentialist, since the features of a trait that make it a virtue are not intrinsic features of the trait; they are causally downstream from the trait and largely external to its possessor. On the other hand, it is not because of the *value* of the consequences of the trait that Hursthouse calls the trait a virtue; rather it is because having traits with those consequences is what it is to be good *qua* human being (Hursthouse 1999, 167). Nevertheless many of the arguments for and against virtue consequentialism will apply to Hursthouse's theory as well.

1. The Case For Virtue Consequentialism

Why be a consequentialist about the virtues? Some of the motivations are the same as those for being a consequentialist about permissible action or about social policy. It is plausible that the purpose of morality and moral rules is to make people (or sentient beings generally – but *someone*, at least!) better off. If so, then we should not call a character trait a virtue if it does not help anyone or if it harms more than it helps. For instance, just as we might criticize racist, sexist

⁵ In this way the virtue consequentialist could avoid Adams's objection that competitiveness is not a virtue despite its good consequences (Adams 2006, 56).

or homophobic actions, laws, or institutions on the grounds that they harm a group of people without helping anyone, we might also criticize certain racist, sexist or homophobic *character traits* on the grounds that their possession harms people.

The consequentialist about virtue need not think that welfare is the only goal of morality. Some acts and rules have consequences that are bad in a way that goes beyond negative effects on the total well-being – for example, they result in unjust distributions of goods, or in unjust punishments or rewards. Just as acts, rules, and institutions can have such results, so can character traits. The consequentialist about virtue thinks that such bad results are at least part of the explanation of the fact that those traits are vices, and this is a plausible explanation.

According to Driver, a main point in favor of virtue consequentialism is that it entails that what counts as a virtue can change over time as circumstances change (Driver 2001, 84-86). Thus virtue consequentialism can explain why our attitudes towards certain traits have changed over time. For example, chastity was once considered to be an important virtue for women in Western societies, but this is no longer the case. “If women were not chaste, men would have no confidence in paternity and would not support children. The social consequences of this would be disastrous” (Driver 2001, 84). As societal attitudes about such matters change, and e.g. paternity testing develops, chastity ceases to be required to prevent disaster, and thus chastity ceases to be a virtue. Of course, attitudes about what is a virtue can be wrong. But it is a nice feature of virtue consequentialism that it does not require us to attribute widespread mistakes to people about what the virtues are, at least in cases like this.

This is not a decisive consideration in favor of virtue consequentialism. Robert Adams argues that given the impact of AIDS, if people regarded the consequences of sexual activity as relevant to chastity’s status as a virtue, chastity ought to have made a comeback and been

regarded as a virtue; the fact that this didn't happen undermines Driver's argument (Adams 2006, 57). Todd Calder points out that there are alternative explanations of why a trait may go from a virtue to a non-virtue (Calder 2007, 211). This can happen in the case of a trait that is not a *fundamental* virtue, but is a virtue only because it is a way of having some other virtue.

Chastity can lose its status as a virtue because at one time, being chaste is a way of loving the good (one's own or others'), which is a fundamental virtue, while at another time it is not a way of loving the good. In general, the anti-consequentialist is not committed to the view that all virtues are *necessarily* virtues. Luke Russell suggests a different reply. He says that while in the case of chastity we would be content to say that chastity simply ceases to be a virtue when it stops having good consequences, this would not be the case with other virtues such as honesty or sympathy (Russell 2007, 477). In those cases we would be conflicted about whether honesty and sympathy remain virtues. He thinks this supports a "disjunctive" view according to which virtues are *either* intrinsically or instrumentally good.

According to Driver, it is also a point in favor of consequentialist theories of virtue that they do not require knowledge – e.g. knowledge of the good – for virtue. This is important because she thinks that some virtues, such as modesty, actually *require* ignorance. Modesty necessarily involves ignorance of one's own excellence: to be modest is to believe that one is not as good as one really is (Driver 2001, Ch. 2). Whether this is an accurate account of modesty is debatable. Consider Erik Wielenberg's example of Amazing Bob, who is actually the strongest and most intelligent person in the world but believes himself to be only the hundredth-strongest and hundredth-most-intelligent, and brags constantly about his status as the hundredth-best person in the world (Wielenberg 2005, 105-6). Amazing Bob is not modest, but counts as modest on Driver's account.

Even if Driver is wrong to say that modesty requires ignorance, virtue may still be *compatible* with lack of knowledge. Driver argues that it is not necessary to have any particular psychological state, such as knowledge of the good or good intentions, to be virtuous. Good intentions are no guarantee of virtue, and bad intentions are no guarantee of vice. Consider Huck Finn, who intentionally performs actions benefiting Jim *while believing he acts wrongly*. Arguably, his intentions fail to be good, but this does not make him vicious; on the contrary, he is virtuous in spite of his intent not to do what is good (Driver 2001, 54-55). The virtue consequentialist can explain this by appeal to the good consequences of his character. Of course, this does not show that virtue consequentialism is true, since the intention to do what is right or good is not essential to non-consequentialist theories of virtue. Furthermore, it is possible that even if Driver is right that *some* virtues do not require good intentions or other particular psychological states, other virtues do (Russell 2007). Nevertheless Driver's argument is an important move in defending virtue consequentialism against the charge that it is essential to virtue that one have one's heart in the right place; if this amounts to the claim that intention to do good is essential to virtue, it is a highly contestable claim.

Experimental philosophy has provided some recent support for virtue consequentialism. According to a recent study by Adam Feltz and Edward Cokely, the folk are more inclined to call a trait a virtue if it has good consequences. Feltz and Cokely claim that their study supports the view that "while other factors may be relevant to virtue attribution, the predominant factor is the consequences that those character traits bring about" (Feltz and Cokely 2013, 713) – not pure virtue consequentialism, but closer to it than any other theory. Of course this is only one study, asking fairly abstract questions to participants about a scenario involving "Pat" who has a trait characterized only by the extent to which people feel good or protected, or ashamed or unsafe,

when it is exercised (Feltz and Cokely 2013, 705). Perhaps further studies asking more specific questions would yield different results. And of course whatever such studies show, they do not demonstrate the truth of any theory of virtue, only what people believe about virtue. So whatever support virtue consequentialism gets from this study is limited; but it can at least provide some defense against claims that virtue consequentialism is objectionably unintuitive, since that claim has no experimental support yet.

A final important motivation for virtue consequentialism is its potential for use in indirect consequentialist explanations of morally permissible action. While a direct consequentialist determines the permissibility of an act by looking at its consequences and those of its alternatives, an indirect consequentialist does so by looking at the consequences of some other thing associated in some way with the act. Rule consequentialism is one well-known version of indirect consequentialism. According to the rule consequentialist, the permissibility of an act is determined by whether it accords with or violates a correct rule, and the correctness of a rule is determined by its consequences.⁶ Consequentialism about the virtues can be employed in an analogous way. There are at least two possible ways to do this. One is suggested by Peter Railton: “an act is right just in case it would be done by someone having a character, the general possession of which would bring about at least as much utility as any alternative” (Railton 2003, 227). Whether an action is permissible is determined by what a virtuous person would do under the circumstances, and whether someone is virtuous is determined according to Maximizing Virtue Consequentialism. A slightly different view is suggested by Walter Sinnott-Armstrong: “whether an act is morally right depends on whether it stems from or expresses a state of character that maximizes good consequences and, hence, is a virtue” (Sinnott-Armstrong 2014).

⁶ See Hooker (2000) for a recent defense of this view.

Here we do not look counterfactually at what a virtuous person *would* do, but rather at what trait actually motivates the action; but again we evaluate that trait according to MVC.

2. The Case Against Virtue Consequentialism

In discussing Hume's view about virtue, after agreeing with Hume that virtues are generally beneficial, Adam Smith gives the following argument against virtue consequentialism: "it seems impossible that the approbation of virtue should be a sentiment of the same kind with that by which we approve of a convenient and well-contrived building; or that we should have no other reason for praising a man than that for which we commend a chest of drawers" (Smith 1991/1759, 246). To call someone virtuous is to praise her; if virtue consequentialism is true, this praise is for her beneficial qualities. But just about any sort of thing can have beneficial qualities. Though we might praise a chest of drawers, the praise is different from the praise we give a person. To praise a person for her virtue is to express admiration. Nobody sees a chest of drawers and thinks: *I'd really like to be like that*.

But Smith's humorous example is perhaps not decisive. The important difference between a person and a chest of drawers is that a person has mental states and acts intentionally; and we have already seen that a primary motivation for virtue consequentialism is that mental states such as intentions to do good are *not* necessary for virtue. So perhaps the virtue consequentialist can plausibly claim that we should not see such a big difference between the virtues of people and the virtues of furniture.

If Smith is right, then there should be some character traits that we do not find admirable yet that count as virtues on the virtue consequentialist view: having good consequences should

fail to be sufficient for a trait to count as a virtue. Contemporary criticisms of virtue consequentialism have focused mainly on establishing this insufficiency claim. For example, Robert Adams argues that even if a trait like competitiveness has good consequences, it is not a virtue since it involves wanting to do better than others (Adams 2006, 56). Perhaps, though, it is not so counterintuitive to say that a certain degree of competitiveness is a virtue. It is a trait for which we sometimes praise people. Why couldn't it be a virtue to want to do better than others, as long as it does not lead to, e.g., cheating or harming others?

Other traits might pose more of a problem. Given the right circumstances, traits like malevolence or dishonesty could have good consequences, and would therefore count as virtues (Slote 2004, 30; Adams 2006, 54-55; Calder 2007, 204). Of course, such circumstances would be farfetched. In order for malevolence to cause good consequences systematically, some magical demon would have to systematically interfere with people's actions, bringing about the opposite of what they intend. Still, the objection goes, malevolence would not be a virtue in such circumstances. The malevolent person would still be intending to cause harm to his "victims." "The problem for consequentialist theories of virtue and vice is that, although maliciousness in such a world systematically promotes the good, it does not make the characters of malicious people, such as the one described above, any better, and thus, maliciousness is still a vice. Yet on a consequentialist account of virtue and vice, maliciousness would be a virtue in these worlds because it promotes the good" (Calder 2007, 205).

Should the virtue consequentialist be moved by this argument? Under these bizarre circumstances, wouldn't it be more virtuous to be malevolent? Calder argues that even though a virtuous person would choose to be malevolent in such cases, this is not because malevolence is virtuous in those circumstances; rather, it is because *benevolence* is a virtue, and being

malevolent would be the best way to benefit people. Malevolence would be chosen as a means, not as an end (Calder 2007, 206). So we have two competing ways to think about this situation. Either malevolence has become a virtue in light of its good consequences, or it remains a vice that would nevertheless be chosen as a means by the virtuous person because it serves another virtue. Calder argues that the second way is the better way to think about things, but it is unclear why it is better. The virtue consequentialist should say: we have reached a stalemate with clashing intuitions about the status of malevolence. As seen in the previous section, the folk may well side with the virtue consequentialist in this clash of intuitions (for what that is worth).

Calder also objects that virtue consequentialism doesn't tell us anything about the intrinsic properties of virtues and vices (Calder 2007, 208). Thus, it does not tell us anything about the "common structure" of the virtues (Calder 2007, 209). This is certainly true and follows from the basic virtue consequentialist idea that it is the consequences of a trait that determine its status as a virtue or vice. The question is, why think that virtues have a common intrinsic structure? Only if we thought that virtues are defined by internal psychological states would we think this to be so; but a large part of the motivation behind virtue consequentialism is the denial of that Aristotelian/Kantian idea. It seems question-begging to argue against the virtue consequentialist on these grounds.

Virtue consequentialism may face some of the same kinds of problems that consequentialism about permissible action faces. For one thing, virtue consequentialism seems to entail that it is very difficult to be sure that any given character trait is a virtue. A trait that we think of as being a virtue might, in the long run, have bad consequences. Virtue consequentialism requires a good deal of epistemic modesty. Non-consequentialist views do not require the same

degree of modesty, since events taking place in the distant future will have no bearing on a character trait's current status as a virtue.

Consequentialist views about permissible action face well-known problems involving cases where a better consequence can be brought about via means that seem impermissible, such as executing an innocent person to prevent a riot, killing an innocent person to distribute his organs to several others who need them, or torturing one person for the amusement of billions. These sorts of objections do not carry over straightforwardly to virtue consequentialism, since someone whose character disposes her to perform such actions would likely be disposed to kill or torture people in other circumstances where doing so would not have the best consequences. Still, if someone were disposed to perform such acts *only* when it really would maximize utility, she would count as virtuous. This may be hard to accept.

The virtue consequentialist may wish to say that being disposed never to harm innocents is a virtue, even though in these rare cases harming innocents would have good consequences, because it *generally* has good consequences to be disposed never to perform such actions. Of course, as I pointed out in Section 1 above, the most plausible version of virtue consequentialism will not entail that a character trait is a virtue or a vice full stop; the most fundamental virtue facts are of the form 'V1 is more of a virtue than V2.' So we must be a bit more careful in formulating the objection. Virtue consequentialism entails that being disposed never to harm innocents is more of a virtue than being disposed to punish or harm arbitrarily, haphazardly or self-interestedly (which is probably not a virtue to *any* degree); but it also entails that being disposed never to harm innocents is *less* of a virtue than being disposed to punish or harm only in cases where it brings about the best outcome. It is the latter claim that seems counterintuitive.

A final problem concerns the place of virtue in moral theory. Virtue has been alleged to play many roles. In the previous section we saw that it has been employed in theories of morally permissible behavior. Another role is in axiology and the theory of well-being: virtue is among the things alleged to make the world, or a human life, better just by being a part of it (Ross 1988, 134). A more indirect role in axiology is also possible: the world has been alleged to be made better by the virtuous getting pleasure, and worse by the vicious getting pleasure (Ross 1988, 138; Feldman 1997, 162; Kagan 2012, 6).

If virtue consequentialism is true, virtue probably cannot play these axiological roles (Bradley forthcoming). As explained above, the best version of virtue consequentialism is scalar, and on the most plausible version of the scalar view there is no absolute threshold between virtue and vice. This fact wreaks havoc on our ability to employ virtue in normative ethics, axiology and the theory of well-being. For example, suppose we are trying to determine the contribution a particular character trait makes to the value of the world. If the scalar view is true, there is no absolute fact of the matter about whether that trait makes a positive or negative contribution to the value of the world! Thus in some cases there will be no fact of the matter about whether the world has on balance positive value. The same difficulty arises if virtue contributes directly to individual well-being: there will be cases where there is no fact of the matter about whether such a life is worth living. In order for virtue to be theoretically useful in these projects, the virtue consequentialist must provide a plausible account of the threshold between virtue and vice, and this seems impossible.

So where does this leave us? Virtue consequentialism's theoretical utility seems limited in some ways. But this is perhaps unsurprising given that on the virtue consequentialist approach, virtue is a *derivative* notion, not a fundamental one. Within a consequentialist framework

considerations of *intrinsic value* are fundamental; virtues are understood in terms of their conduciveness to the value of the world. For this reason, those who think virtue is a crucially important theoretical notion are not likely to be attracted to virtue consequentialism. Thus, whether one is at all persuaded by the defenses of virtue consequentialism offered here is likely to depend on one's attitude concerning the theoretical importance of virtue.

References

- Adams, Robert. 1976. "Motive Utilitarianism." *Journal of Philosophy* 73: 467-81.
- Adams, Robert. 2006. *A Theory of Virtue: Excellence in Being for the Good*. Oxford.
- Aristotle. 1976. *Nicomachean Ethics*, trans. J.A.K. Thomson. Penguin.
- Bentham, Jeremy. 1907/1789. *An Introduction to the Principles of Morals and Legislation*. Oxford: Clarendon Press. Library of Economics and Liberty. 15 September 2015. <<http://www.econlib.org/library/Bentham/bnthPML10.html>>.
- Bradley, Ben. 2005. "Virtue Consequentialism." *Utilitas* 17: 282-298.
- Bradley, Ben. Forthcoming. "Character and Consequences." In Iskra Fileva (ed.), *Perspectives on Character* (Oxford University Press).
- Brentano, Franz. 2009/1952. *The Foundation and Construction of Ethics*, trans. E. Schneewind. Routledge. First published 1952; Schneewind translation first published 1973.
- Calder, Todd. 2007. "Against Consequentialist Theories of Virtue and Vice." *Utilitas* 19: 201-219.
- Driver, Julia. 2001. *Uneasy Virtue*. Cambridge University Press.
- Feldman, Fred. 1997. *Utilitarianism, Hedonism, and Desert*. Cambridge University Press.

- Feltz, Adam and Edward T. Cokely. 2013. "Virtue or consequences: the folk against pure evaluational internalism." *Philosophical Psychology* 26: 702-717.
- Foot, Philippa. 1978. *Virtues and Vices and Other Essays in Moral Philosophy*. University of California Press.
- Hooker, Brad. 2000. *Ideal Code, Real World*. Oxford University Press.
- Hume, David. 1975/1777. *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, Third Edition, Ed. L.A. Selby-Bigge and P.H. Nidditch. Oxford: Clarendon Press.
- Hursthouse, Rosalind. 1999. *On Virtue Ethics*. New York: Oxford University Press.
- Hutcheson, Frances. 1991/1725. *An Inquiry concerning the Original of our Ideas of Virtue or Moral Good*. In Raphael (ed.), *British Moralists 1650-1800, Vol. I*, pp. 261-299.
- Kagan, Shelly. 2012. *The Geometry of Desert*. New York: Oxford University Press.
- Lewis, David. 1986. *On the Plurality of Worlds*. Blackwell.
- Locke, John. 1975/1689. *An Essay concerning Human Understanding*, ed. P. Nidditch. Oxford: Clarendon Press.
- Mandeville, Bernard. 1991/1714. *An Enquiry into the Origin of Moral Virtue*. In Raphael (ed.), *British Moralists 1650-1800, Vol. I*, pp. 229-236.
- Moore, G.E. 1903. *Principia Ethica*. Cambridge University Press.
- Railton, Peter. 2003. *Facts, Values, and Norms*. Cambridge University Press.
- Raphael, D.D. (ed.) 1991. *British Moralists 1650-1800, Vols. I & II*. Hackett.
- Ross, W.D. 1988. *The Right and the Good*. Indianapolis, Ind.: Hackett Publishing Co. First published 1930.

Russell, Luke. 2007. "What Even Consequentialists Should Say About the Virtues." *Utilitas* 19: 466-486.

Sidgwick, Henry. 1981/1907. *The Methods of Ethics*. Hackett.

Sinnott-Armstrong, Walter. 2014. "Consequentialism." *The Stanford Encyclopedia of Philosophy* (Spring 2014 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2014/entries/consequentialism/>.

Slote, Michael. 2004. "Driver's Virtues." *Utilitas* 16: 22-32.

Smith, Adam. 1991/1759. *The Theory of Moral Sentiments*. In Raphael (ed.), *British Moralists 1650-1800, Vol. II*, pp. 201-254.

Von Wright, G.H. 1963. *The Varieties of Goodness*. London: Routledge & Kegan Paul.

Wielenberg, Erik. 2005. *Value and Virtue in a Godless Universe*. Cambridge.