Introduction

JULIA DRIVER
Dartmouth College

The evaluation of character has taken on new significance in moral theory, and, indeed, some advocate a shift in focus away from evaluating action to evaluating character. This has been taken to pose special challenges for utilitarian and consequentialist moral theory. Utilitarianism's commitment to impartiality and its seeming failure to accommodate virtue evaluation have led to problems, some of which are developed in the essays in this volume.

Utilitarianism has been charged with failure to account for the true value of human relationships – friendships, family love, and the like. These relationships are characterized by *partiality* in one's response to others. Utilitarianism, on the other hand, seems committed to impartial norms. If one's moral obligation is to promote overall well-being, then the well-being of friends and family does not warrant any extra weight. Yet this is strongly counter-intuitive. Friends are taken to have special obligations to each other in virtue of their friendship. Friends are taken to value each other, at least in some sense, more than others. If Ann is my friend then I really ought to take a *particular* interest in Ann's well being. Likewise with one's children. Further, there is the related objection that consequentialist practical reasoning seems incompatible with how one ought to *think* about one's friends. It would seem that the consequentialist is committed to holding that something – including a friend – has value to the extent that the friend promotes *overall* well being. Thus, the criticism goes, the consequentialist is committed to viewing the friend as having mere instrumental value. Yet this way of thinking is utterly incompatible with friendship.[1] This is indicative of its failure to account for a wide variety of partial values. This problem I'll term the 'partiality problem.'

Related to this problem, though also distinct, is the problem that both utilitarianism and the more general consquentialism have in accommodating or modelling the virtues. Some virtues, such as those involved in being a good friend, or parent, will be characterized by attitudes of partiality, and thus fall under both the partiality and the virtues problem. However, others aren't directly problems which involve taking a partial attitude towards others. For example, one classic criticism made by Bernard Williams is that utilitarianism is incompatible with integrity because it requires of agents that they

[1] See Neera Badhwar, 'Why it is Wrong to Always be Guided by the Best: Consequentialism and Friendship,' *Ethics*, ci (1991).

© Edinburgh University Press 2001

https://doi.org/10.1017/S0953820800003101 Published online by Cambridge University Press

maximize the good.[2] This will mean that there will be situations in which the agent must renounce deeply held convictions and values for the sake of generating more overall good.

The more general complaint is that the utilitarian tries to reduce decision-making to a single rule or principle, and this cannot possibly capture the richness and complexity of actual moral decision-making one finds employed by virtuous persons. The generous person isn't necessarily, or even characteristically, trying to maximize the good. Neither is the courageous, the trustworthy, the just, or the kind. Indeed, they may knowingly act against the greater good, as when a compassionate person cannot bring herself to harm one individual even for the sake of saving others – or, as in a case proposed by Peter Railton – a loving husband spends money to see his wife rather than send the money to Oxfam.[3] With this latter example we can see how the partiality problem and the virtue problem overlap. However, the criticism here isn't one of partiality so much as a general complaint that the sort of decision-making advocated by utilitarians and consequentialists completely fails to represent virtuous decision-making, which may well not include any reference to overall good.

The authors in this volume have all developed novel arguments which expand and develop these criticisms of consequentialism. All but Hurka and Ridge believe that these problems cripple the theory. Hurka believes that the consequentialist can accommodate the intuition that virtues are intrinsically good, and vice intrinsically bad, by holding virtues to consist in attitudes toward the good that are themselves good. Ridge tries to show that consequentialism can – though at some cost – provide a plausible account of integrity.

Some of the other authors, however, are not as sanguine about consequentialism's ability to handle difficulties relating to our intuitive commitment to partial norms. One standard way of putting the problem is to point out that intuitively we do have *options* to pursue our own projects, even at the expense of overall utility. Further, there seem to be impersonal constraints on our behaviour which conflict with a demand to promote overall utility. The latter problem isn't a partiality problem, though one could argue, for example, that a doctor who fails to respect the constraint relating to bodily integrity, and who cuts 'up one healthy patient against his wishes in order to distribute his organs to five of his patients who are in need of transplants and would otherwise die',[4] exhibits a vice, and if this is

[2] Bernard Williams, 'Consequentialism and Integrity,' *Consequentialism and its Critics*, ed. Samuel Scheffler, New York, 1988.

[3] Peter Railton, 'Alienation, Consequentialism, and the Demands of Morality,' *Philosophy and Public Affairs*, xiii (1984).

[4] Brink, 156.

https://doi.org/10.1017/S0953820800003101 Published online by Cambridge University Press

demanded by utilitarianism, it would seem that the theory would demand vicious behaviour.

However, in his essay for this volume David Brink argues that it isn't so much consequentialism's inability to accommodate constraints and options, which are themselves problematic. Instead, he argues that the real problem is its inability to account for associative duties. These duties are not impartial – that is, they are by their very nature, partial. The parent who is not partial to his children would be regarded as failing in a very crucial respect. This is distinct from an option, since parents would seem to be obligated to behave in various partial ways towards their children. Nor is this an impersonal constraint on parental behaviour, since associative duties by their very nature are not impersonal and impartial.

It is important to note – as Brink does – that impartiality does not evaporate with associative duties. Parents ought to treat their children impartially. It would be completely inappropriate for a judge to not recuse himself from presiding over a trial involving one of his children, relatives, or friends – precisely because partial concern is considered inappropriate, even just wrong, in this context.

However, even aside from these considerations, there is a deeper way in which impartiality is not jettisoned. One could argue that there is an impartial norm at a higher level. For example, in evaluating parental behaviour it would be inappropriate to hold that it's permissible for Arthur to show favouritism to his children, and not allow the same for Mary. This would indicate that higher level impartiality applies across the board, and offers a strategy for allowing 'partial' norms justified at a higher level of impartiality, which a utilitarian or a consequentialist could endorse. Consequentialists incorporate this higher level impartiality using four general strategies.

(1) Indirect consequentialism would hold that the moral quality of an action – it's rightness or wrongness – is determined by the consequences of something else, that is, it is determined indirectly. Rule consequentialism is the most popular form of indirect consequentialism, maintaining that an action is right if and only if it is performed in accordance with a rule which maximizes utility.[5] However, one could

[5] This can, in turn, be spelled out in a variety of ways. For example, Richard Brandt and Brad Hooker opt for a view which holds that the right action is the one performed in accordance with a set of rules which would maximize utility in an ideal world – the world in which most people accept the rules. In the real world, most people don't. So this view is less demanding than one which holds that the right action is the one performed in accordance with the set of rules which maximize utility in the real world – since few are living up to the utilitarian ideal in the real world, individual moral demands become more severe. See Richard Brandt, 'Some Merits of One Form of Rule-Utilitarianism,' in his *Morality, Utilitarianism, and Rights*, New York, 1992. Also, Brad Hooker, *Ideal Code, Real World*, Oxford, 2000.

https://doi.org/10.1017/S0953820800003101 Published online by Cambridge University Press

also develop this as a form of virtue consequentialism – the right action is that action performed as a result of the agent's virtue, and what counts as a virtue is determined on consequentialist grounds. With this strategy one would be able to say that, for example, if certain virtues demand partial concern for one's children then the right actions will be those that exhibit this partial concern. Nurturing my children – which can be given a consequentialist justification – requires of me that I display partial concern towards them. Thus, if I must choose between the well-being of a group of children and the well-being of my own child, the right action is the one which promotes the well-being of my own child. The danger with this strategy is that one might be presented with conflicting demands rendered by conflicting rules or virtues. Then the picture gets more complicated, because there will have to be a hierarchy of rules and virtues, and a separate consequentialist argument for the ordering of the hierarchy. Nevertheless, one can see the general strategy here for incorporating partial concern within an overall impartial framework.

(2) The objective consequentialist strategy is distinct from (1) because it simply maintains that the right action is the one which results in the best actual outcome. This is contrasted with subjective consequentialism which holds that the right action is determined by the subjective states of the agent. For example, one version holds that the action is right if and only if the agent performing it *expects* it to produce the best, or the best expected, overall outcome. Another possibility is that the right action is the action the agent performs in *trying* to produce the best overall outcome. However, the objective consequentialist is not at all committed to the agent's consciously trying to maximize the good in performing right actions, or even the agent's simply expecting that the action she wants to perform will produce the best overall outcome (regardless of what she is trying to do). The right action is simply the one which produces the best outcome, but what is going on in the agent's mind when she decides to perform that action is left open. Arguably, if it really is better overall that parents concern themselves primarily with their own offspring, then the right action will be the one exhibiting the partial concern. The more plausible approach along these lines would involve acceptance of the fact that in many cases partial concern does not result in the right action; however, because the partial concern is the result of character traits, which are good on consequentialist grounds, we wouldn't want to blame a parent who exhibited this concern. In this way we can see a real distinction between indirect and objective consequentialism.

Another possibility (3) is to pursue a strategy suggested by Amartya Sen which holds that consequentialism is perfectly compatible with

https://doi.org/10.1017/S0953820800003101 Published online by Cambridge University Press

situated decision-making.[6] In this way one denies that it is committed to agent-neutral values. This has the consequence of committing one to the view that consequentialism simply refers to any theory which determines moral quality in terms of consequences. Ethical egoism would count as consequentialist, but then so would a whole host of other theories which are also not agent-neutral. Thus, the fact that I am a parent is a legitimate factor affecting my decision-making in certain contexts – e.g., in a case in which I must choose whether to save an entire group of children, or my son, my situation demands saving my son (or at least allows it). A critic of this strategy might charge, however, that this really guts consequentialism – true, it would help alleviate problems, but at the expense of depriving the theory of much of its force. Impartiality is a way of demanding of people the sorts of sacrifices of self-interest that are also necessary for morality. Once this is breached, and once agent-neutrality, at least at some level, is dispensed with, then how can one in principle restrict the scope of permissible self-interest? Ethical egoism becomes just as justified as any other version of the theory. If one appeals to common sense morality to rule this out, then why not just stick with common sense morality?

Other consequentialists, like Hare, opt for (4), which holds something like the following: consequentialism is the correct theory, but in real life situations, since we are not fully informed archangels, it is very difficult to deploy this decision procedure and we must fall back on heuristics which inform our intuitions, and which are, overall, good-producing.[7] The Hare strategy is a 'two-level' strategy – there is the critical level – the level of the archangel, and then the level which employs heuristics.[8] Our intuitions, for example, might favour preference to the near and dear, as Frank Jackson argues, because such a preference itself maximizes overall utility, and our intuitions are responsive to this.[9] So, impartiality operates, but at the higher level. It is permissible for me to show favouritism, but it is for anyone else as well, and the reasons have to do with considering the overall consequences of allowing it.

[6] Sen has spelled this option out in a variety of contexts. See, for example, 'Rights and Agency,' *Philosophy and Public Affairs*, xi (1982).

[7] R. M. Hare, *Moral Thinking*, Oxford, 1982.

[8] On Hare's view, the archangel level will be objective – the archangel does see the actual consequences, though we ordinary humans cannot always see them. Thus, there is overlap here between a two-level strategy and the objective consequentialist strategy. But one could adopt a different sort of two-level view, in which the level determining actual rightness wasn't objective.

[9] Frank Jackson, 'Decision-theoretic Consequentialism and the Nearest and Dearest Objection,' *Ethics*, ci (1991).

https://doi.org/10.1017/S0953820800003101 Published online by Cambridge University Press

However, it should be noted that Brink, in developing his criticism of consequentialism, is not focusing on permissions, but on duties. It is not simply that the parent is allowed to favour his children – the parent *must* favour his children.

Consider the following familiar case adapted from Hare:[10] a father and son are travelling on an aeroplane and the plane crashes. His son is sitting on one side, and to the father's other side is a world-renowned medical researcher on the brink of discovering a cure for cancer. The father can save only one person as he struggles from the crashed plane. Whom should he save? Brink's point is that it isn't simply permissible for him to save his son, he has an obligation to do so even though overall impartial utility is best served by saving the medical researcher. This is a rather dramatic example, but Brink's point would still be made with a more modest one, where overall utility differences are marginal. This is an example of an associative duty. This distinction having been noted, it seems nevertheless as if the same consequentialist considerations could be brought to bear in justifying these duties. Is it better for parents to be obligated to take care of their children? Of course there is a built-in motivation to do so. But it also seems that this is the sort of responsibility that occurs because of one's situation relative to the child. It is perfectly fine for the consequentialist to make use of the fact that we are naturally motivated to take particular interest in our own children and their well-being. Harnessing this motivation to a good end will appeal to the consequentialist. Thus, we have a powerful argument for why this particular concern, as long as the context is very restricted, is very efficient.

If the consequentialist adopts the indirect strategy of virtue consequentialism, then the father's saving his child is the right action because, for example, it displays an instance of the sort of nurturing concern which is a parental virtue. The difficulty is that this move might preclude one from making nicely split moral judgements that reflect our ambiguous evaluations of hard cases. If virtue determines rightness, then it will be harder to make sense of claims such as 'He did the right things, but only because he is ruthless' or 'He did the wrong thing, but out of a generous nature'.

With the objective version we get a nice split between decision-procedure and evaluation. For the objective consequentialist, whatever decision-procedure works is fine – we evaluate based on objective criteria, and whatever works may vary, even, from context to context. Thus, partial norms are perfectly well justified at the higher level of impartiality. Here, though, one would say not that the father did the

[10] *Moral Thinking*, pp. 138 f. The case Hare discusses was originally presented by Bernard Williams.

https://doi.org/10.1017/S0953820800003101 Published online by Cambridge University Press

right thing, only that what he did was praiseworthy. While the indirect version seems able to accommodate Brink's concern, the objective version does not do so directly. The objective version instead views associational duties as contingent on the fact that when we are in relationships with others, better actual outcomes occur with partial concern for the well-being of others, and this justified adopting partial attitudes towards others. Brink might argue that this misses the point of the objection, since the critic holds that these duties would exist in any case. Suppose that there were no difference – in terms of good produced – between the world of the impartial and that in which relations have attitudes of partiality; wouldn't it still be good to have the partial concern – wouldn't it still be a standard by which we say relationships are good? Additionally, it would be quite odd to evaluate a friendship, for example, in terms of that friendship's contribution to the overall good. That seems irrelevant to evaluating *that* friendship. While I believe that the objective consequentialist can avoid some of these problems – for example, by appealing to the systematic goods resulting from relationships of this type, Brink has posed a tough challenge to consequentialism.

His positive thesis is that we should understand the norms of these relationships on analogy with prudence, since there will be a similarity between interpersonal connections or continuity of relationships and *intrapersonal* psychological continuity: 'one's relations to associates are similar in kind to, if different in degree from, those that hold between oneself now and oneself in the future. But this suggests that one has the same sort of reasons to be concerned about associates as one does about ones own future self' (168). This serves the function of showing that associative duties are not problematic, since most people agree that prudence makes normative demands. If prudence is held to involve the development of one's deliberative capacities, this seems quite plausible. I have a concern for developing my deliberative capacities that needs no further justification. This is a crucial aspect of agency. Just as my own future well-being is of concern to me, so is the well-being of those close to me – those with whom I have formed associational bonds (which are similar to intrapersonal psychological bonds). There is a psychological connection to others that underlies these bonds – like the psychological connection that exists between the past, present, and future self.

Brink has provided another account of the normative basis for associational duties, one that does not appeal to utilitarian impartial norms. However, suppose that Angela has been told that she needs a brain operation in order to live, but that as a result of this operation she will have absolutely no memory of her former life. There is absolutely no psychological connection in that sense, though there is same-

https://doi.org/10.1017/S0953820800003101 Published online by Cambridge University Press

ness in structure, if you will (we can assume that they will have the same *sorts* of preferences, etc.). Though new Angela doesn't know anything about old Angela, let's say that someone meeting new Angela, who knew the old one, would be instantly struck by the similarity and have no problem regarding her as really Angela. Does old Angela have any reason to be concerned for new Angela? Yes, though there is no *psychological* connection between the old and new. Arguing that there is in virtue of similarity of preference won't solve the problem, since the judgement would hold even if Angela's personality changed.

Further, one can imagine an even wilder thought experiment: suppose that reincarnation is true and that when we die we are reborn though we have no memory of our previous lives, and we come into the world as blank slates. Also suppose that virtue in this life is rewarded in the next. Angela believes in this account. Does she have any reason to be virtuous in this life? Yes, though there is no psychological connection.

This analysis may also have some difficulty accounting for some special obligations; for example, the obligation of a father to his newborn child. A man who abandons his child in preference to others is generally taken as failing in his duty, though he had not yet had time to form any psychological bond with his child. Interpersonal bonding that one sees between family and friends doesn't seem to provide a basis for associational duties. But Brink could respond that it provides a basis, though there may be other factors at work as well. The new father who abandons his child may not have any psychological connection, but it may be that relationships of dependency in general impose such duties. And when one is causally responsible for the dependency then another basis for such duties may exist. Evidence for this would be the fact that we believe the father is not failing in a duty if he arranges alternative care and nurturing for the baby, by, for example, giving it up for adoption. The dependency issue is dealt with and he has no psychological bond which would impose associational duties. So there are ways for Brink to avoid this objection, though they impose extra complexities.

Michael Smith also believes that consequentialism will have a fairly serious problem in trying to accommodate certain crucial features of friendship. First of all, a friend will be one who, when confronted with the choice of giving slightly more of a benefit to a friend than to a stranger, will choose to give it to a friend, all other things being equal. This is 'The Friendly Motivation Thesis'. The second is 'The Friendly Justification Thesis', which holds that 'a good friend is disposed to offer the fact that her friend is *her friend* as a justifying reason' for giving the slightly greater benefit to the friend. So a good friend is motivated

https://doi.org/10.1017/S0953820800003101 Published online by Cambridge University Press

to give slightly more of a benefit to her friend, as opposed to strangers, and further she feels it appropriate to use the fact of the friendship itself as a justification of so acting. Smith argues that the consequentialist has more of a problem with the justification thesis than the motivation thesis. This is because the most intuitively plausible or palatable version of consequentialism – modest global consequentialism (MGC) – fails because it can't be spelled out plausibly in such a way as to preserve a distinction between consequentialism and alternative theories. Further, the only other alternative – immodest global consequentialism (IGC) – can't make sense of the justification thesis. The MG consequentialist agrees with the non-consequentialist that it is better to give a slightly greater benefit to the friend, but disagrees in that she holds this extra value to be non-moral value, whereas the non-consequentialist holds it to be moral value. Smith attacks MGC on the basis of this distinction. IGC, in contrast to MGC, holds that neutral values are the only ones that determine rightness and further that neutral values have much greater significance than relative ones. So, while the IG consequentialist is quite free to hold that it may be good for people to believe it is better to give slightly more to a friend, and thus be motivated to do so – since such a belief will maximize neutral value, the belief is false. On this view it would not be true that it is better to give slightly more to a friend because she's my friend. It might be useful to believe this, and it might be useful for people to provide that justification and believe it to be warranted, but it isn't true on this theory – instead the proper justification is just that giving slightly more to friends, overall, maximizes neutral value. Thus, IGC conflicts with a deeply held moral conviction about the nature of friendship. However, Smith does note that the severity of this problem is open to dispute.

The challenge is a deep one, and offering a solid defence of consequentialism is beyond me here. However, one way to meet Smith's challenge would be to reject the FJT as presented. One could hold, instead, that there is a sense in which the FJT is true – but that the explanation for its truth is consequentialist. The belief is justified by consequentialist considerations – but this has no bearing on the issue of whether or not someone using FJT need be aware of the consequentialist justification in order to be acting as a true and good friend. Consider an analogy with evolutionary explanations for our emotions. It may be true that jealousy, for example, is very useful in cementing family bonds and keeping couples together, and actually promotes love and affection. Given we think these things are good, we may then think that jealousy is a good emotion because it promotes those ends. However, someone in the throes of jealousy may have no awareness of the emotion serving those ends at all – and, indeed, such awareness

https://doi.org/10.1017/S0953820800003101 Published online by Cambridge University Press

may be incompatible with the emotion. 'I'm jealous just because he's my husband!' will be the characteristic thought, not 'I'm jealous because jealousy had adaptive benefits for human society!', though the latter may well be true. In this spirit, IGC may accept FJT as an account of what an agent appropriately has in mind when helping her friend, but reject it as offering a final justification of the behaviour.

Thomas Hurka gives us an account of vice (and contrasting virtue) which is compatible with consequentialism, but not committed to it. A consequentialist adopting Hurka's account will have resources to respond to many standard criticisms of consequentialism. Hurka's approach is to argue that virtue need not – even to the consequentialist – be understood as having purely instrumental value. Virtue is taking a positive attitude towards the good – loving the good. As such, it is intrinsically good.

In this particular essay, Hurka focuses on developing his account of vice. Vice consists in failure to love the good. As such, vice is intrinsically bad. This manoeuvre allows Hurka to argue that vice is bad even when it does not result in a bad outcome. For example, even ineffective malice is vicious. This is because the malicious person has an attitude toward another's pain which is itself intrinsically evil. Thus, Hurka would be able to analyse Moore's deluded sadist as someone who does exhibit vice – because he believes he is hurting others and he delights in that belief – even though he is not actually hurting anybody. Actual bad outcomes don't matter on this account.

Some might regard Hurka's account as too narrow. The tendency is to view virtue and vice as dispositions to act as well as dispositions to feel. But the objection from the objective consequentialist will focus on the issue of ignoring actual outcomes. An implication of Hurka's account is that loving the good is both necessary and sufficient for virtue. An incompetent or klutzy lover of the good would count as virtuous even though he systematically actually harms others. While we might be reluctant to call this person 'vicious', I don't think he would be regarded as virtuous either. One wouldn't think it appropriate to admire him, for example, and one might try to keep him from trying to help others (since he'd just end up hurting them inadvertently).

Christine Swanton pursues a quite different course, arguing that even accounts such as the one provided by Hurka grant consequentialism too much. On her view a consequentialist account of virtue cannot be correct since it leaves out crucial ways in which we recognize value. Consequentialist theories privilege one mode of valuing, or one attitude to value, over others. Thus they are committed to what she terms 'The Hegemony of Promotion': 'The only right-making relation is that of promotion.' Swanton argues that close consideration of

https://doi.org/10.1017/S0953820800003101 Published online by Cambridge University Press

virtues reveals that they are not consequentialist in nature, since it is not the case that the domain or field of a virtue is a good (or the domain of a vice an evil). Rather, the domains of specific virtues and vices will be much more concrete – 'pleasure ..., friends ..., failures ...'. So, things identified as goods aren't intrinsically good; pleasure is not intrinsically good – it is good only if handled well. As an example she uses failure – and the ability to handle failure well, as a good. Thus, failure is not intrinsically evil. Though handling failure well is a good, and is certainly virtuous, she does not want to claim we ought to seek out failure and the opportunity to handle it well and thereby display virtue.

A consequentialist might try to handle this type of criticism in a variety of ways. First she could point out that Swanton's intuitions which support the view that other attitudes, besides promotion, are appropriate, can be handled by the argument that they are instrumentally good. Thus, the good of 'honouring' value may be accounted derivative from promotion of value: the ability to handle failure well is a virtue because of its utility. Surely it would have tremendous social utility. For example, refusal to accept failure or handle it well in political contests could result in lack of social stability. Smaller scale refusals result in broken friendships, and competition that has a rougher more unpleasant element. We find it immediately pleasing because we realize the person who fails in a competition – which is an emotional experience – is giving up a chance to display his displeasure. This entails a sacrifice, but one which is really necessary.

Like Swanton, Michael Ridge is concerned that consquentialism cannot model crucial virtues, or come up with a plausible way of representing them. However, he also believes that the consequentialist can handle certain criticisms – particularly those having to do with its seeming inability to account for an agent's legitimate concern for his own integrity. By integrity Ridge means 'a person's commitment to following her own all-things-considered moral judgements' (236). Ridge rightly points out that the puzzle has to do more with the agent's being concerned for her own integrity, since it is open to a consequentialist to place 'integrity' on the list of goods which are intrinsically valuable. But any consequentialist account of the value of integrity – which is at some level committed to its impartial value – gets things the wrong way around, according to a critic like Stephen Darwall. That's because the consequentialist pursues an 'outside-in' approach – starting from states of affairs impartially characterized as good, and then explains the value of inner states such as integrity by appealing to the outside valuable states of affairs. However, an alternative is the 'inside-out' approach which begins by considering the internal valuable states – like integrity – and then determines the value of external

https://doi.org/10.1017/S0953820800003101 Published online by Cambridge University Press

states of affairs by reference to the internal value. This seems more consistent with integrity since the starting point is not agent-neutral, but agent-centred. The agent is concerned for her own integrity because it is her own. One ought not give this up even to save the integrity of others. However, Ridge points out that at least on one interpretation of the inside-out approach, one might be led to valuing a form of self-indulgence – should I really value my own integrity at the cost of others'? On another interpretation it merely holds that we are committed to doing what we believe we ought to do, and this isn't incompatible with consequentialism.

Ridge suggests a third way of accounting for the concern for integrity which avoids triviality and which, again, the consequentialist could endorse. One could appeal to the Humean idea of reflective endorsement and argue that when each person examines or reflects on his or her own moral commitments and values, those are supported by reflection and the content of those values and commitments is consequentialist. This is an inside-out approach, because the agent reflects on her own values and then works out to the value of external states of affairs, which she realizes through reflection ought to be promoted. Further, it doesn't commit one to moral self-indulgence. But in making this move the consequentialist is committed to something unusual. To start with decision-procedures that pass the test of reflective endorsement, and then construct objective standards by which to evaluate actions from those decision-procedures, will strike many consequentialists as putting the cart before the horse. But the beauty of this approach, as Ridge points out, is that reflective endorsement can be held to be important, but for reasons having to do with moral epistemology – how do I know what has value? Reflective endorsement. But this importance spares the agent from the charge of self-indulgence.

Rae Langton argues that the consequentialist has value all wrong – and that our attitudes towards harms and goods gives clear evidence of this. Our 'reactive attitudes' do not mirror appropriate consequentialist responses. Thus, if Mary steps on my foot on purpose I am more resentful of her than I am of Kate, who steps on my foot by accident, even though I might know that Mary will never do it again, whereas Kate, who is inveterately clumsy, will do so numerous times. If resentment is a proper guide to apportioning blame, then I should blame Mary's intentional though one-off harm more than Kate's accidental but habitual harm. Yet the consequentialist, citing usefulness as a guide, will argue that Kate is more deserving of the blame, since she's the one whose behaviour modification would have the best result. Further, applying this to virtue, an objective consequentialist would maintain that the virtue is the trait that systematically produces good;

https://doi.org/10.1017/S0953820800003101 Published online by Cambridge University Press

the vice one that systematically produces bad – thus, Kate is vicious because here clumsiness produces bad consequences, and she is thereby deserving of resentment on an accuracy measure, whereas Mary, while possibly vicious, is at least not as vicious as Mary, and not as deserving of resentment. This seems absurd. Langton points to the fact that there are two separate standards one could use for determining the goodness or appropriateness of the resentment in these cases – an accuracy standard would be sensitive to whether or not the resentment was responding to a genuine vice, or a usefulness standard which determines the resentment's appropriateness by its degree of usefulness. Thus, if Kate steps on my toe, and if resentment is useful, then I should feel it. But this holds even if Kate possesses no vice. Thus there is the potential for the two standards to come apart. When this happens, Langton believes that the usefulness measure will take precedence for the consequentialist, and usefulness will determine whether or not my resentment is appropriate, and this conflicts with common sense. I ought to resent Mary more, though she causes less harm and though the resentment directed towards her is, presumably, less useful than that directed towards clumsy Kate.

While Langton's points are directed against a particular consequentialist account of virtue – an objective consequentialist account, the criticism holds generally for those accounts which hold virtue's value to be instrumental. As Langton points out in her essay, however, there are still strategies for the utilitarian to pursue in countering these examples. The familiar one involves accepting the schizophrenia observation, and simply biting the bullet. My resentment is justified only to the extent it is useful, but it might not be good or useful for me to believe this. Or the consequentialist could make a distinction between what guides resentment and how one ought to assess it. And, perhaps, one could hold that one ought to assess resentment as appropriate just in case it is *useful*, yet the philosopher judging cases ought to assess it as appropriate only if it is *accurate.* But this does seem to pose problems – as Langton points out, it would seem incumbent upon the philosopher to give up accuracy where that serves utility. Thus, she would be giving up doing philosophy itself.

My own view is that the consequentialist has more options. The most promising is to argue for a plurality of goods, and maintain, as earlier consequentialists have, that something like knowledge or truth is itself a good. Indeed, one might try defining intellectual virtue as a disposition to acquire true belief (true belief being the relevant 'good' for this sort of virtue). Thus, it is not incumbent upon the philosopher to abandon the accuracy assessment – that itself represents a sort of good, and there is no fundamental contrast to be drawn between the two standards. Then the difficulty for the utilitarian is reduced to the

https://doi.org/10.1017/S0953820800003101 Published online by Cambridge University Press

classic problem of incommensurability. When there is tension between differing goods, how is it to be resolved? But this is a separate problem.

More generally, though, I believe there are problems associated with relying on reactive attitudes when providing an account of holding people responsible. A standard utilitarian manoeuvre can be employed by the consequentialist here. Just as intuitions about specific cases are formed under normal circumstances, so our reactive attitudes evolved in response to standard situations. Thought experiments positing unusual situations – for example, the situation in which I know the person stepping on my foot is doing it only this one time, that no one else will know of it and be encouraged to behave similarly, etc., would be quite unusual, as would the contrast case of clumsy Kate who try as she might can't seem to stop herself from squashing toes. But this observation might not be fair to Langton's strategy. She will of course point out the distinction between how we do feel and how we ought to feel, and the significance of reactive attitudes concerns how we ought to feel in a given situation – and, in so far as the consequentialist determines this based solely on forward-looking considerations, the account is going to be guilty of the charge that it gets things upside down. To hold the badness of intentions to be determined by badness of outcome is false; badness of outcome, rather, is determined by badness of intention. So, when Mary evilly squashes my toe the pain and anger I feel are due to the badness of her intentions. When Kate squashes my toe the badness isn't as bad because her intentions aren't as bad. But here Langton will have a tough balancing act. If she truly does want to claim that badness of outcome depends on badness of intention, the objective consequentialist will point to the unpalatable features of this alternative. Joe and Jimmy both intend to hurt Bill, but deluded Joe tries to hurt him with a wet piece of spaghetti while Jimmy uses a sawn-off shot gun. The outcome in the latter case is far worse than the former, though the intention is the same in both (i.e., to send Jimmy to the hospital). Langton could make her point slightly differently, however. She could argue that it isn't strictly speaking true that badness of intention determines badness of outcome; rather, badness of intention enhances badness of outcome so that if we have two outcomes that are equivalent in terms of simple harm generated – like two squashed toes, the squashed toe which is the result of bad intentions is worse than the squashed toe due to accident. But I might resent someone's insensitivity more than their intentional wrong-doing.

But placing such significance on reactive attitudes has other difficulties. Reactive attitudes diminish over time, and it seems that they ought. Thus, I resent the squashed toe today far more than the squashed toe of two years ago. And, if I did not most would regard me

https://doi.org/10.1017/S0953820800003101 Published online by Cambridge University Press

as defective. Maintaining the same high level of resentment for two years would indicate a problem – I would be *really* holding a grudge. Yet there doesn't seem to be any reason for this in terms of what the acts themselves deserve. Surely the squashing of my toe two years ago was just as vicious as the squashing of my toe today. The appropriate level of resentment is tied to instrumental considerations (that was two years ago, let it go – the resentment serves no positive function), whereas immediate resentment has a strong impact on the offender and the ones witnessing the offence. Virtue and vice evaluation depend simply on a consideration of the consequences produced by the character trait. Whether or not I resent someone's manifest viciousness separately depends upon instrumental factors.

Most of the papers in this volume raise important challenges to any consequentialist theory hoping to account for virtues, and associated partial and special concern for others. As I've tried to indicate in this introduction, however, I believe that consequentialism has the resources to meet these challenges.

julia.l.driver@dartmouth.edu

https://doi.org/10.1017/S0953820800003101 Published online by Cambridge University Press