Routledge
Taylor & Francis Group

International Journal of Philosophical Studies

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/riph20

A Morality Fit for Humans

Philip Pettit

To cite this article: Philip Pettit (14 Dec 2023): A Morality Fit for Humans, International Journal of Philosophical Studies, DOI: 10.1080/09672559.2023.2290547

To link to this article: https://doi.org/10.1080/09672559.2023.2290547

Published online: 14 Dec 2023.

INTERNATIONAL JOURNAL OF PHILOSOPHICAL STUDIES
https://doi.org/10.1080/09672559.2023.2290547

A Morality Fit for Humans

Philip Pettit

Philosophy, Australian National University, Canberra, Australia

ABSTRACT
There are a number of assumptions made in our accepted psychology of moral decision-making that consequentialism seems to violate:: value connectionism, pluralism and dispositionalism. But consequentialism violates them only on a utilitarian or similar theory of value, not on the rival sort of theory that is sketched here.

ARTICLE HISTORY Received 11 November 2023; Accepted 29 November 2023

KEYWORDS Consequentialism; moral psychology; robustly demanding goods; pluralism

In 'Consequentialism and Moral Psychology' (Pettit 1994) I begin with the core assumptions of a realistic moral psychology, as I call it, and point out that many critics of consequentialism have claimed that it neglects those aspects of our mentality. The main thesis of that paper is that the consequentialist theory of moral rightness can take a form that respects our psychology, provided that it endorses a particular view of the values that make consequences good. Thirty years later, I still hold by that thesis, although I would phrase things somewhat differently, give a different description of the elements in a realistic psychology, and draw explicitly on a view of values that does not figure in the earlier paper. In this short piece, I try to set out a case for the thesis that follows such lines.[1]

The paper is in two sections. The first offers an account of consequentialism and looks at the features of the human mentality or temperament that critics take to conflict with its assumptions. And then the second introduces a view of the values we ought to care about, arguing that insofar as consequentialism takes such goods into account, it will cease to conflict with that psychology.

1. Consequentialism and Psychology

A Consequentialist Theory

By all accounts, consequentialism is a theory that identifies the morally right option in a choice, personal or social. In the most prominent version, the right option in such a choice is that which promises to have the best consequences,

CONTACT Philip Pettit philip.pettit@anu.edu.au

given the probabilities rationally associated at the time of choice with the different consequences. The right option, as we may say, is that which promotes the best consequences. This formula can be spelled out more formally in different ways but it will serve, as it stands, for our current purposes.[2]

The formula leaves us with two questions. What is a consequence? And by what metric should we gauge consequences in determining which option is best, assuming that there is no tie for best? The formula is meant to provide only a criterion by which to judge which alternative in the choice facing a person or a society is the morally right one for them to take or to have taken, not necessarily a guideline for them to follow in making the choice; more on this in the second section. Even in such an assessment, however, the formula won't be of much use unless we are clear about what consequences are and about how to gauge consequences in determining the right option.

A consequence is anything that an agent can be said to bring about by what they do, to stick for simplicity with a personal rather than social choice. One sort of consequence is the causal result of the choice: say, the relief I give you by returning a loan. Another may be a result guaranteed constitutively by the choice: say, the keeping of a promise that I ensure by returning the loan. And a third may be a result that presupposes contingent circumstances: say, proving to be your only reliable debtor , as I do in returning the loan.

Consequences of these kinds are all situation-dependent. But there may also be consequences that are disposition-dependent, as we shall see. Examples might be the love I give you when I act out of loving care, the friendship I give you when I act out of a friend's concern, the respect I give you when I act out of a respectful attitude. That you enjoy my love in the first case, my friendship in the second or my respect in the third depends, not just on what I do-in-the-situation – not on my act, period – but also on the disposition manifested in that act. It is a serious failure in standard forms of consequentialism, as we shall see, that they neglect such consequences (Pettit 2018b).

So much on the question of what should count as consequences But, to turn now to the other question raised, how are the consequences promoted by the rival options in a choice to be ranked so as to establish that one option promises to bring about better consequences than another: to establish, in other words, that it is the morally right alternative for the agent to take or to have taken. The main issue here is whether consequences are to be ranked by neutral goods or values, as we may call them, or by goods or values that privilege a particular individual or indeed a particular of any kind.

Something will count as a neutral value insofar as it represents something that anyone might treat as a value, regardless of who they are and of whom or what they are connected to. A value will fail to be neutral in this sense if it involves a good for a particular individual or country or group: a benefit to Ireland or to the philosophy profession or whatever. And it will also fail to be neutral if it is identified by indexical reference back to the particular individual – or indeed the

society or polity – responsible for the choice in question: the good, in the words of the relevant party, will be 'my good' or 'our good', or 'the good of mine or ours'. The first sort of non-neutral value is a particular-specific good, the second a particular-relative good: as it is usually cast, an agent-relative, good.

Consequentialism has traditionally been distinguished by the fact that it takes the morally right option to be determined by neutral values alone, not values of such a particularistic kind. This fits with the fact that consequentialism appeared in the mid-twentieth century as a generalization of the utilitarianism associated with Jeremy Bentham and John Stuart Mill (Sinnott-Armstrong 2015).[3] It broke with utilitarianism in allowing that happiness may not be the only neutral good or value but continued to hold that neutral values determine how to rank the likely consequences of the options in a choice and that the morally right option is that which promotes the best consequences.[4]

Why should consequentialism take only neutral values to be relevant to determining what is morally right? Basically, because of the common assumption that the morally right option is that which has the best consequences from the point of view of all relevant parties: in a standard form, all relevant human beings. Morality contrasts with prudence and patriotism and the like in just this regard and for that reason, so the assumption goes, it must let the right option be determined by values that are independent of the good of particular people and places and so on.[5]

A Realistic Moral Psychology

There are three claims made by a realistic moral psychology. In the earlier paper I characterize those elements in a negative and rather opaque fashion as anti-atomism, anti-economism and anti-rationalism. Let me now give them more positive, if still somewhat technical names: value connectionism, value pluralism and value dispositionalism. Before we look at these assumptions, it will be useful to say something on what valuing and values involve.

By all accounts there is a difference between valuing something, even valuing it for prudential or sectarian reasons, and just desiring it. Human beings value things, I shall assume here, insofar as they desire them reliably enough, by their own lights, to be able to bet on themselves – whether in commitments to others, or in personal resolutions – to act on them in a range of contextually salient variations of circumstances. I will value reducing weight, or working steadily, or sticking by my friends, insofar as I explicitly or implicitly commit to taking those actions, not just when there is no rival desire in play, but even when I face a range of conflicting temptations: desires that I will have committed in the relevant context to resisting. If I value doing something in that way, then it will count intuitively as a failure if I do not do it because of a conflicting temptation. That is what distinguishes valuing from just desiring.

The difference here is important. It means that when I value something, and circumstances are appropriate – the only conflicting dispositions count as temptations – then as a commissive agent, I will have a motivating reason to choose it, albeit a reason that may generate a less intense desire than the temptation itself. But it also means that when I act on my valuation rather than on the temptation, I can hope to justify my doing so by reference to the commitment I had made – perhaps to others, perhaps just to myself – or to the feature that made the action fit for such a commitment on my part.

Valuing in this generic sense need not count as moral valuing. To value something morally, I shall assume, is to value it from a perspective, roughly, in which the interests of others, weighed in impartial terms, are taken into account equally with one's own. The others in question may or may not be taken to include other sentient creatures as well as ourselves, and the interests of such creatures may or may not be given the same weight in impartial terms. For current purposes, however, we may restrict relevant others to human beings.

With any form of valuing, there is a difference between what someone values as a matter of fact – what features are reliably attractive enough to make bearers fit for commitment – and what they value according to their beliefs; they may be ignorant or self-deceived about what actually matters to them in that way. But with moral valuing, and perhaps with other forms too, there is the further issue of whether what they actually value morally is what they ought to value: whether the features to which they give importance on the grounds of reflecting people's interests in general really do reflect those interests. We will return to this later.

On this view of value and valuing, it is important that something can count as worthy of being valued by an individual insofar as it has a suitably robust attractor, whether or not it attracts from the perspective associated with morality. That is to say, then, that moral values do not exhaust the category of values. There are many features for which an agent may value something that do not reflect the interests of others equally with their own. Examples would be features that make something attractive from the prudential perspective of the person or from their perspective as a member of a certain group or a citizen of a certain country.

Value Connectionism

The three elements in a realistic psychology amount to three assumptions, borne out in everyday experience, about the patterns in which we human beings value things. The first assumption is that our values are connectionist, in the sense that they often directly involve other people. If I love someone, Mary, what I value is the welfare of that particular person – something concrete that presupposes their existence – not anything abstract like the welfare of anyone I happen to love.

There is an old philosophical test of whether an attitude is connectionist – object-dependent, as it is often put – rather than not. Imagine a neuron-for-neuron replica of me living in a world that does not contain Mary. By hypothesis that counterpart will have all the attitudes that my and their make-up sustains, including the abstract valuing of care for anyone beloved. But despite that similarity within the skin, so to speak, my replica will not value Mary's welfare; that valuation will presuppose, not just the make-up we share, but also Mary's actual existence.

I take it that we are psychologically connected with other individuals and other groups, even particulars of other kinds, in this valuing way. We are invested in the good that may accrue to such concrete entities in a way that takes us, quite literally, out of ourselves. The investment shows up in the loves and loyalties that are dependent on the contingencies of where we are born, who it is we live with, and who are the children we parent. We value abstract goods aplenty, ranging from biodiversity to peace among peoples to the prosperity of those who will live in the future. But we also have values that bind us to the fortunes of the concrete individuals and particulars of our world.

Value Pluralism

The second assumption that I associate with being realistic about our psychology is value pluralism. Intuitively, we care for their own sake about a plurality of goods, ranging from goods in our own life like success or prosperity; to goods in the lives of others like the happiness of our children, our friends, or any group of others; or of course to collective goods like peace or cooperation among peoples. We may value such things for further effects that they promise to bring in their train, but we will value them also in themselves. And we may value them only conditionally on things being a certain way – say, conditionally on not living in a state of war – but in the world where things are that way, we value them unconditionally. We desire them for their own sake and we do so with the reliability that valuing requires.

Consistently with valuing things for such different features, we may desire and so value them with greater or lesser intensity: we may feel a stronger desire for the welfare of our children than the welfare of others. And consistently with desiring and valuing them, we may desire and value them with greater degrees of reliability: the contextually salient variations of circumstance over which we would continue to prioritize the happiness of our children may have a greater range than those over which we would prioritize the prosperity of a friend. In both respects, the value-making features in question may vary in their weights, as we say, so that the presence of one feature in an option may be enough to outweigh the presence of another feature in a conflicting option. Suppose that giving my child the satisfaction of being in the audience for their school play explains and

justifies why I choose to attend the play rather than go to an important business meeting. In that case I will give greater weight to the relevant feature of the first option, and give lesser weight – though presumably some weight – to the relevant feature of the second.[6]

While valuing things for a plurality of features, and giving them different weights, it is abstractly possible that for any choice between rival options or arrangements, the weights attached to those features are determinate and comparable enough to guarantee that in every choice there is going to be a best alternative, or at least a set of alternatives that are equally good. But that is just an abstract possibility. Given the variety of features that can support our valuing one or another alternative, it is likely that in many cases there are at least some alternatives where it is not the case that one is better than others and not the case that they are equally good. The issue is just indeterminate and the options, as Ruth Chang (2002) puts it, are on a par.

This is a salient possibility for an agent with regard to value-making features in general, personal or otherwise. But it is also likely to hold of those features that are relevant to the moral value of conflicting options. It would be amazing if the features that an agent takes to reflect the interests of other individuals as well as themselves had such determinate and comparable weights that there was no such thing as indeterminacy on the issue of which alternative in a choice – or which set of alternatives – is morally best for the agent to take. And it would be even more amazing if this were true of the features that really do reflect the interests of other individuals as well as the agent themselves.[7]

Value Dispositionalism

The third assumption of a realistic moral psychology, as I see it, is value dispositionalism. This is the view that there are many things we value such that in order to realize them, the agent has to act on a corresponding disposition, and not just target it as a result to bring about. Take love and friendship, the examples that we used in illustrating connectionism. In order to give you the good of my love, I have to act out of a loving disposition and in order to give you the good of friendship I have to act out of a disposition of friendship. There is no way of giving you those goods – and they surely are goods – other than by manifesting the corresponding disposition. Suppose I give you the concern associated with love, or the care associated with friendship, out of any other disposition, be it base opportunism or general benevolence. In that case I will act as if I loved you, or as if I were a friend, and may give you a good of a significant – I may help you out of a difficulty, for example – but I will not give you the real thing: the good of love or friendship itself.

In order to give you the real thing in cases like this, I cannot provide the contingent benefit associated in a given situation with love or friendship – say, helping you out of a difficulty – just for the opportunistic reason that it is in my self-interest or for the philanthropic reason that it contributes to improving the world. I must provide it on such a basis that even if it did not serve my opportunistic or philanthropic ends in that way, still I would have provided the benefit for you. I must provide it on the basis that you are someone I love or that you are a friend: that consideration must be so weighty in a range of circumstances that it would lead me to act in the same loving or friendly way. In other words, I must provide the benefit out of the disposition of a lover or a friend.

As this is true of love and friendship, so something similar is true in other cases too. In order to give you respect, I must let you make your own decision in certain matters out of a disposition of respect: that is, because you are a distinct, autonomous person. In order to give you my fidelity in promise-keeping, I must give it to you out of a disposition of fidelity: that is, because I made the promise. In order to give you my honesty, I must tell you the truth out of a disposition of honesty: that is, because it is the truth. And in order to do justice in dealing with you, at least if we are to trust the Roman Digest, I must give you your due out of 'a steady and enduring will to give each their due': that is, because doing so is what justice requires.[8]

2. Fitting Consequentialism to Psychology

Consequentialism, as already suggested, is best seen as a generalization from the 19th century doctrine of utilitarianism that drops the assumption that the only neutral good is happiness. Before looking at how consequentialism can take a form in which it coheres with the assumptions of a realistic moral psychology, it is worth remarking that on at least one standard construal, utilitarianism conflicts with those assumptions.

Utilitarianism and Moral Psychology

On the construal I have in mind, utilitarianism displays three distinctive features. First, it takes happiness to be, not just the good relevant to moral rightness, but the only thing that matters to people and the only thing worth seeking for its own sake. Second, it defines happiness in reductive terms: paradigmatically, in terms of pleasure; hedonism explains the connection between this claim and the first, arguing that pleasure is what people seek in everything they do.[9] And third, it conceptualizes happiness, not as something good for concrete people – that is, good for people who actually exist or will actually exist – but as an abstract good: a quantum whose promotion may call for expanding the number of people who actually exist.[10]

Such utilitarianism is consequentialist in identifying the morally right option in any choice with the alternative that promotes the neutral good of happiness in the sense defined. But it adds a hedonistic, quantized theory of the good – purportedly, the only good – to a consequentialist theory of the right. And insofar as it does this, as we shall now see, it runs into conflict with each of the three psychological assumptions we discussed.

It runs into trouble with value connectionism insofar as it takes happiness to be, not just the only good that is relevant to morality, but the only good that people care about in any mode: the only feature, allegedly, that leads them to think that something is worthy of being desired for its own sake. On the utilitarian view, the person who values the friendship or love that they give another – or indeed receive from another – must value it as a way of generating pleasure in the parties to that relationship or in third parties. They cannot value it for its own sake, as in the connectionist picture of human beings.

The good envisaged by utilitarians is also monist rather than pluralist. It represents a homogenous value that can allegedly serve to determine what value ought to be given to other ends that we seek such as love and friendship, respect and justice; these will presumably be valuable only insofar as they promote pleasure, and morally valuable insofar as they promote pleasure overall. Moreover, since overall pleasure can come at lower or higher levels, and materialize in fewer or more people, it can provide a basis in principle for ranking any alternatives against one another; it can rule out indeterminacy in matters of value, including moral value, of the kind that pluralism makes likely.

Finally, the good of happiness or pleasure cannot accommodate the complexity registered in dispositionalism. Whether someone promotes pleasure overall by taking a certain option in a choice is not going to depend essentially on the disposition out of which they act. Regardless of whether or not they are personally opportunistic, for example, or whether or not they are impartially benevolent, their actions will be capable of promoting or failing to promote pleasure overall; motive will be irrelevant. Pleasure, to revert to our earlier language, is not a disposition-dependent consequence of what an agent does; it depends only on the behavior displayed by the agent, the act they perform.

But if utilitarianism conflicts in this way with a realistic moral psychology, is there a form that consequentialism might take that would reduce or remove such conflict? I shall argue that there is. I do so by looking in turn at each of our three psychological assumptions and showing how an alternative form of consequentialism can endorse these.

An Alternative Form of Consequentialism

Any form of consequentialism must preserve the idea that the morally right option in a given choice is that which promotes the consequences that count as best by the standard of neutral values; for short, it is the option that promotes the

most neutral value overall. There may be two or more equally valuable options in any choice, of course, but we can put aside that complexity in the current context.

Consequentialist moral theories agree in taking the morally right option in any choice to be a function of how much neutral value it promotes but disagree in taking rival views of the values that are morally relevant. Utilitarianism takes happiness to be the only morally relevant value – and indeed the only intrinsic value of any kind – whereas other consequentialist approaches take competing views of what is valuable and valuable in a morally relevant way. I want to focus now on a class of approaches that I find independently appealing, and have supported elsewhere (Pettit 2015; 2018b, 2024). They each offer a picture of morally relevant values that rules out the sort of conflict with our human psychology that utilitarianism displays; they avoid any conflict with the psychological assumptions of value connectionism, value pluralism and value dispositionalism. That the image of values they project rules out that sort of conflict is a bonus feature but not the only or even the main thing to be said in its favor. So I shall assume – but not argue – for purposes of this paper.

According to connectionism, many of the goods that people treasure in ordinary life involve other individuals as a matter of their very constitution; we illustrated the point with the examples of love and friendship. Insofar as they involve other people in their particular identities, such goods will not be neutral, of course. But if I take it to be valuable that I act towards you in the manner of a lover or a friend then, assuming that I do not think of myself as special – and I can hardly do so if interested in the morally right – I must take it to be valuable more generally that people should act in that same way towards their lovers and friends. And that is to say that I must take it to be a neutral value that people should act in those ways towards those with whom they are connected as lovers and friends.

That same lesson applies to other people as well as to me and it generalizes to other goods of a connectionist character. It applies, for example, to the goods I confer on you in giving you respect, in being honest or faithful in dealing with you, or in doing justice by you. It means that among the neutral goods that a consequentialism might prize are the goods that consist in people being mutually respectful, in being honest and faithful with one another, and in treating one another justly.

This makes it clear that consequentialist theory is not in essential conflict with value connectionism and that all that is needed to avoid conflict is that it should take people's enjoyment of love and friendship and respect, as well as their enjoyment of honesty, fidelity and justice at one another's hands, to be neutral values. And that is hardly a counter-intuitive constraint to lay on the goods that the theory should countenance. For it is a matter of the merest common sense that the non-neutral form of those values are goods that human beings cherish in their relationships with one another and that they cherish them for their own sake, not just for the sake of an ulterior benefit like the pleasure they might bring.

If a consequentialist theory may and should countenance different goods of this kind, that alone is sufficient to imply that it must also satisfy value pluralism, which is the second assumption in a realistic moral psychology. But such pluralism is independently plausible in any case, since there are a range of other goods, unrelated to connectionist concerns, that human beings naturally seek for their own sake. Valuing their own pleasure or absence of pain, they must equally value the general enjoyment of those goods, as utilitarians hold. And the same is true of a variety of similar goods. Valuing their own freedom or security or prosperity, for example, they must also value the enjoyment of such goods on the part of people generally.

If consequentialism is to acknowledge a plurality of neutral values in these two categories, it is likely to have to face the sort of indeterminacy that we associated with pluralism. There are almost certainly going to be some sets of options that each option appeals on the basis of different values and where those values are not weighted exactly enough to enable the options to be ranked against one another. In such a case, no option will determinately score above or at the same level as another.

The third, dispositional assumption in our moral psychology is that there are many goods such that it is possible for an agent to promote them only by acting out of suitable motives: the goods may materialize as consequences of what the agent does but they can only materialize as disposition-dependent consequences. As we have already registered, many connectionist goods, and their neutral counterparts, are disposition-dependent in this sense. The only way of promoting my friendship with others will be by giving them the care of a friend out of the motive of a friend. And so the salient way for me to promote friendship in general – a neutral good – is by acting towards suitable others in that way. And the same is true of love and respect and indeed of the honesty or fidelity or justice that I may confer as benefits on others.

There are other ways, it is true, in which I might seek to promote the neutral versions of such goods. Rather than promoting that good in the neutral sense by my own personal relationships, I might in principle proselytize among others about the case for cultivating such relationships with one another. Or, again in principle, I might seek to promote the good by philanthropically providing relief from the sort of destitution that puts the relationships out of people's reach. But those alternative modes of promotion are much less assured of success than the salient disposition-dependent mode, and in any case are hardly inconsistent with it. If they are available, they do not argue for replacing the salient mode, only for complementing it.

There may also be other neutral goods that are disposition-dependent in this way. But it is enough for our purposes to register that any consequentialist theory that endorses the neutral values associated with connectionist benefits will cohere nicely with the psychological feature I described as dispositionalism. When the theory pronounces on what it is morally right

for an agent to do, it need not ignore or downplay the importance that people naturally give to the motives out of which they act in their relationships. On the contrary it is likely to argue in many cases that the right thing for the agent to do is to act on such motives.

Returning to a point made at the beginning of the paper, these observations make it clear that consequentialism can come into line with our moral psychology just insofar as it acknowledges the importance of disposition-dependent consequences as well as consequences of other kinds. There are many goods that we prize in our personal lives that require the benefactors to act out of suitable motives, and there are many neutral counterparts of those goods that ought to be taken into account in any plausible theory of the morally right. If consequentialism is to be a plausible theory, then it must take this lesson on board, acknowledging the relevance of disposition-dependent as well as situation-dependent consequences.

A Final Issue

But can a consequentialist theory internalize the lesson? Any such theory is going to require an agent to act so as to promote the best consequences. So how can it also require that in areas where disposition-dependent consequences are relevant, the agent should act out of the relevant dispositions? How can it require in such cases that instead of keeping an eye on the likely consequences an action may generate, the agent should let a disposition like that of love or friendship or respect, honesty or fidelity or justice, take over in their thinking? The response to this issue takes us back to the early observation that consequentialism offers a criterion for identifying the morally right option in a choice, not necessarily a guideline for the agent to follow in deciding on which option to adopt.

Let it be granted that on any consequentialist approach, the agent should control in their decision-making for taking the right option: they should take steps to increase the likelihood that they will do so. Our early observation amounts to the claim that while that is so, the exercise of such control does not require them to assess each option for how far it would promote the best consequences and then to make the choice on the basis of that calculation. How could an agent exercise such control, however, without doing such a calculation? The answer is that they may exercise virtual rather than active control. And it turns out that the notion of virtual control can help resolve the issue raised.

Suppose that I am put in charge of ensuring that no one enters a workplace by a particular door. I may actively control for ensuring this by keeping an eye on the door. But I may still control for that result while reading a book or using my phone. I will do so under two conditions: first, if I can be relatively sure that anyone seeking to enter will make a noise in doing so and that such a noise will alert me to their presence; and second, if I am ready take appropriate action in

that event: in the event that such a red light goes on. I will virtually control for ensuring that no one enters by being on standby, ready to assume active control on a need-for-action basis. And if no one seeks to enter, I will expect to be paid at the end of the day for, while I may have done nothing actively but read my book, I will still have discharged the control task allotted to me.

This notion of virtual or standby control can apply with any agent who acts in a suitable context out of a disposition of the kind envisaged. Assume that there are cases where love or friendship or respect or whatever is the salient good on which I as an agent should be targeted. And assume that in those cases I will generally find myself moved by the relevant disposition or, if you like, focused on the considerations it makes salient: for example, 'NN, my friend, is in need'. How can I act on that motive, promoting the good of this friendship – and, we may suppose, the good of friendship in general – while letting a concern for doing what is morally right control my response? The answer is: by resort to virtual control; by being ready to respond to any red light that may indicate that helping my friend in this case may fail to promote neutral good overall and by taking appropriate action in that event.

To recall an old joke, I will spontaneously agree to help if a friend asks me to move an apartment; the only issue will be, when I can best do so. But what if the friend asks me, without further explanation, to help them move a body? That will surely put on a red light and prompt me to take active control of what I do, trying to ensure in particular that I do not do something morally wrong. As I can let a concern for moral right control my disposition-dependent promotion of friendship in this standby or virtual manner, so in the same way I can let it control the promotion on my part of other goods like love and respect, honesty, fidelity and justice. There are lots of red lights that are primed to light up in cases where I pursue such goods, and it will require only a modicum of sensitivity for me to be ready to register and respond to them.

If this line of thinking is correct, then many of the traditional critiques of consequentialism are unsound. They may be appropriate as critiques of utilitarian consequentialism but, not applying to the class of consequentialist theories at which we have been looking, they do not challenge consequentialism as such. There are other critiques of consequentialism in the literature, of course.[11] But we can at least dismiss complaints to the effect that of necessity the approach is psychologically unrealistic: that it presupposes agents of a different stripe from ordinary human beings.[12]

Notes

1. In the argument of the paper I draw on lines of thought in two recent books: *The Robust Demands of the Good*, published in 2015, and *The Birth of Ethics*, published in 2018. See (Pettit 2015, 2018a). For earlier defenses of

consequentialist theory, see (Pettit 1997, 2012) and for a more recent statement and defense, see (Pettit 2024).

2. On any plausible rendering, the right option will be that which has the highest expected value in the decision-theoretic sense but even that sort of formula raises questions. Should 'expected' reflect the actual credences of the agent, for example, or the credences that they would have if they were rationally responsive to available evidence? And should it be understood in evidential or causal terms?
3. The word was first introduced by Elizabeth Anscombe (1958) but she used it to describe just about any position that assumes that any actions, no matter how unappealing, may turn out to be morally right; nothing is absolutely forbidden.
4. In recent years, some thinkers have opted for using the term for any doctrine in which the right is made a promotional function of the good, whether that good be neutral or non-neutral (Dreier 1993; Portmore 2014; Smith 2009). I reject that line here. For an argument that neutralism is the more important aspect of consequentialism on the grounds that it entail promotionalism but is not entailed by it, see (Pettit 2024).
5. Many will argue, following Peter Singer (1981), that it should be determined by values that are also independent of the good of our particular species. I set aside that issue here.
6. The variation in weights may itself vary between individuals or between different times in the life of any individual. I set aside that complexity here.
7. This is particularly likely on any account of ethics that is naturalistic in character and refuses to treat value-making properties as sui generis. See (Pettit 2018a).
8. *Justitia est constans et perpetua voluntas ius suum cuique tribuendi* (Watson 1985, I.1.10). I rely on my own translation.
9. Things are a little more complicated on the equation of goodness with preference-satisfaction, which was common for a time in the last century. But on this theory too, it is usually assumed that whatever people do, they do for the sake of preference-satisfaction, so that it should be take as the only good. For a critique of that claim, however, see (Pettit 2006)
10. In Jan Narveson's (1973) marvelous phrase, it may therefore call, not for making people happy, but for making happy people. This third feature is not important in the current context – it is crucial for issues of population ethics – and I mention it only for the sake of making the character of this utilitarianism vivid.
11. For a recent defense of the sort of consequentialism envisaged here, see (Pettit 2024).
12. I am indebted to the very helpful comments of an anonymous referee on an earlier draft of this paper.

Disclosure statement

No potential conflict of interest was reported by the author.

References

Anscombe, G. E. M. 1958. "Modern Moral Philosophy." *Philosophy* 33 (124):1–19.
Chang, R. 2002. "The Possibility of Parity." *Ethics* 112:659–88.
Dreier, J. 1993. "Structures of Normative Theories." *The Monist* 76:22–40.
Narveson, J. 1973. "Moral Problems of Population." *The Monist* 57:62–86.

Pettit, P. 1994. "Consequentialism and Moral Psychology." *International Journal of Philosophical Studies* 2:1–17.
Pettit, P. 1997. "A Consequentialist Perspective on Ethics." In *Three Methods of Ethics: A Debate*, edited by M. Baron, M. Slote, and P. Pettit, 92–174. Oxford: Blackwell.
Pettit, P. 2006. "Preference, Deliberation and Satisfaction." In *Preferences and Well-Being*, edited by S. Olsaretti, 131–53. Cambridge: Cambridge University Press.
Pettit, P. 2012. "The Inescapability of Consequentialism." In *Luck, Value and Commitment: Themes from the Ethics of Bernard Williams*, edited by U. Heuer and G. Lang, 41–70. Oxford: Oxford University Press.
Pettit, P. 2015. *The Robust Demands of the Good: Ethics with Attachment, Virtue and Respect*. Oxford: Oxford University Press.
Pettit, P. 2018a. *The Birth of Ethics: Reconstructing the Role and Nature of Morality*. Oxford: Oxford University Press.
Pettit, P. 2018b. "Three Mistakes About Doing Good (and Bad)." *Journal of Applied Philosophy* 35:1–25.
Pettit, P. 2024. "Direct Consequentialism Unlimited." In *Oxford Handbook of Normative Ethics*, edited by D. Copp, C. Rosati, and T. Ruli, Oxford: Oxford University Press.
Portmore, D. W. 2014. *Commonsense Consequentialism: Wherein Morality Meets Rationality*. Oxford: Oxford University Press.
Singer, P. 1981. *The Expanding Circle: Ethics and Sociobiology*. Oxford: Oxford University Press.
Sinnott-Armstrong, W. 2015. "Consequentialism." In *The Stanford Encyclopedia of Philosophy* , edited by E. N. Zalta. https://plato.stanford.edu/archives/win2023/entries/consequentialism/
Smith, M. 2009. "Two Types of Consequentialism." *Philosophical Issues* 19:257–72.
Watson, A. 1985. *The Digest of Justinian*. Philadelphia: University of Pennsylvania Press.