# International Journal of Philosophical Studies

All manuscripts should be sent to the Editor:

Professor Dermot Moran
Department of Philosophy
University College Dublin
Dublin 4
Ireland

Review copies to Dr Maria Baghramian at the above address.

Subscriptions rates for Vol. 2, 1994 (two issues)

| | | |
|---|---|---|
| UK/EC | Institutional | £48 |
| | Personal | £27.50 |
| North America | Institutional | $85 |
| | Personal | $45 |
| Rest of World | Institutional | £55 |
| | Personal | £30 |

Subscription orders should be sent to the publishers, Routledge Journals, Cheriton House, North Way, Andover, Hants SP10 5BE, UK.

Published by Routledge
11 New Fetter Lane
London EC4P 4EE
UK

© Routledge, 1994

ISSN 0967-2559

The Editor gratefully acknowledges the support of the
National University of Ireland

Printed in Great Britain by Bell and Bain Ltd, Glasgow

---

# International Journal of Philosophical Studies

Volume 2 Number 1
March 1994

Editor: Dermot Moran

ROUTLEDGE

NEW SERIES

Key Issues

*Series Editor: Andrew Pyle, University of Bristol*

Until the appearance of *Mind* in 1876, there was no British journal specifically devoted to philosophy. Articles on philosophical subjects competed for space in the pages of the *Edinburgh Review*, the *Quarterly*, and the *Westminster*, and later in the *Fortnightly*, the *Contemporary*, and the *Nineteenth Century*. The result is a body of philosophical literature that is both popular and profound, addressing the great issues of the day in a manner accessible to any thoughtful and literate reader. The issues with which these writers dealt – political liberty, population growth, evolution and ethics, the subjection of women, agnosticism – remain very much alive a century or more later. Until now, however, these articles have only been available to the most intrepid scholar.

*Key Issues* is an exciting new series of titles, aimed at bringing together the original contemporary responses to particular works or the ideas of particular thinkers, as found in these now largely inaccessible journals.

**LIBERTY**: *Contemporary Responses to John Stuart Mill*

Mill's *On Liberty* has turned out to be, as he predicted, the most widely-read and long-lasting of his writings. It has proved, however, extremely difficult to pin Mill down to any definite political doctrines. His contemporaries clearly had the same problems as have beset modern commentators. Some portray Mill as a dangerous revolutionary, a latter-day Jacobin; others see him as peddling mere platitudes. Was he aiming to preach Pyrrhonism, undermine the Church, and dissolve social bonds; or were his principles already informing the thought and practice of Victorian Britain? This volume traces the reception of *On Liberty* in the periodical literature, from the 'rave' review by Buckle in *Fraser's Magazine*, by way of the furious denunciations in such Tory journals as *Blackwood's* and the *Quarterly*, down to later liberals like John Morley and Leslie Stephen.
New Introduction by Andrew Pyle
ISBN 1 85506 244 5 Hbk £45.00 ISBN 1 85506 245 3 Pbk £14.95
Publication January 1994

**POPULATION**: *Contemporary Responses to Malthus*

The second volume in the series deals with the contemporary responses that met Thomas Malthus' *An Essay on the Principle of Population*, published in 1798. It has been said of Malthus' *Essay* that it was a work much more widely discussed than read. This selection represents a wide range of the arguments that followed this contentious work – from those who vehemently attacked Malthusian notions to those who passionately defended them. Including articles by William Cobbett, William Hazlitt and Thomas de Quincey, this volume brings together some of the most lively contributors to the debate.
New Introduction by Andrew Pyle.
ISBN 1 85506 344 1 Hbk £40.00 ISBN 1 85506 345 X Pbk £13.95
Publication June 1994

Proposed future topics in the *Key Issues* series include Mill's *Subjection of Women*, Hobbes' *Leviathan*, *Mind and Brain: Philosophical Discussions 1860–80*, *Mind and Body: Theory and Therapy*.

*Access/Visa/MasterCard/American Express/Switch accepted*

**For further details or for a free copy of our 1994 catalogue listing over 300 titles on Philosophy and the History of Ideas, please contact:**
**Deborah Mann, Thoemmes Press,**
**85 Park Street, Bristol, U.K.**
**Tel: (0272) 291377 Fax: (0272) 221918**

---

# Consequentialism and Moral Psychology

## *Philip Pettit*

### 1. The Flight from Consequentialism

Consequentialism aspires to tell part of the story, and only part of the story, about what makes an option the right one for an agent to choose or to have chosen. Why was it right, if it was right, for Mother Theresa to have devoted her life to work among the poor in India? Why was it right, if it was right, for the United Nations to have imposed an economic blockade on Serbia? Consequentialists say that if those choices were right, then they were right because they 'promoted' the good better than any of the alternative courses of action open to the agents; for the moment we may ignore what exactly 'promotion' involves. Consequentialists tell us something, then, about what made the options right: they tell us that the options must have promoted the good. But they do not tell us everything relevant to the rightness of the options: they do not tell us, in particular, what the good is – what the values are – that should be promoted. Consequentialism does not amount to a total moral vision; it is only of limited significance in the development of such a vision.

But though it is of only limited significance, consequentialism has invited, and continues to invite, a very intense degree of hostility. Traditionally, consequentialists have often been utilitarians about the good; they have thought that all values boil down, in the end, to the one value of happiness or pleasure, however exactly that is interpreted. It is now well understood that consequentialism does not entail utilitarianism and that it can be linked with any of a variety of theories of the good, any of a variety of value-commitments. But the recognition that consequentialism is not shackled in this way to utilitarianism – the recognition that it only tells part of the story about what makes an option right – has hardly reduced the degree of hostility with which its opponents regard it.

It is easy to understand why utilitarianism should invite hostile and impassioned responses. Received wisdom is deeply challenged by the claim

*International Journal of Philosophical Studies* Vol. 2(1), 1–17

that the one and only thing that is of any value in the universe is the experience of happiness or pleasure. But why should consequentialism receive such a bad press? Its opponents often treat the doctrine, not just as a mistaken hypothesis about what makes an option right, but as an expression of some deep misconception about the meaning of life: as a malaise of the spirit to which only the morally wicked or the intellectually arrogant could succumb. Why should this be so? Some opponents of consequentialism are manifestly confused about the doctrine and this confusion may explain their disdain. For example, certain opponents seem to think that consequentialism holds, in the old adage, that the end always justifies the means and they may rant against the doctrine on that ground. But, clearly, this is an outrageous misrepresentation of consequentialism: if we are considering how far an option promotes the good then what is done by way of realizing a certain option – what means is adopted – is just as relevant as what end results from its realization. I do not concern myself with opponents who are guilty of confusion about such straightforward matters but only with sophisticated critics.

So why, then, does consequentialism arouse such hostility, even in philosophically enlightened circles? The reason, I suggest, is that opponents see consequentialism as culpably and fundamentally misrepresenting the moral psychology of human beings, as it is often described. The moral psychology of human beings refers to the commonplace psychology of human beings, so far as it bears on moral matters: the commonplace psychology, as distinct from particular empirical theories. Consequentialists, so their opponents allege, carry on in the criticism and prescription of choice as if human agents were of a different character from that which is revealed in our everyday experience of ourselves and which is charted in this commonplace psychology. Consequentialists carry on, so it is alleged, as if human agents were cast in an alien, unfamiliar mould or were made of alien, unfamiliar material. Anxious to develop a systematic theory of the right, they ignore the basic facts of our moral psychology; they fail to keep in touch with the realities of motivation and deliberation, affection and reason, that are salient to ordinary subjects.

The opponents of consequentialism, as we would expect from this line of criticism, consciously take their starting-point from a picture of how things stand in the realm of commonplace psychology. Drawing on a variety of traditional sources – often literary as well as philosophical sources – they sketch a picture of the human being in action and they argue that that picture gives the lie to the model of the human agent that is allegedly presupposed in consequentialist thought. The agent of the consequentialists, they suggest, is a theoretical fiction and consequentialism is incapable of intellectually surviving exposure to the real thing. The real thing – the human being in action – is often described in terms associated with virtue-theory, as it is called, and this sort of description is contrasted with the more austere, economistic representation that is supposedly associated with consequentialism.

Widespread though it is, I do not think that this line of criticism is on the right track. Let us endorse all the moral–psychological claims that opponents of consequentialism wish to put in place. If the criticism pressed by opponents is correct, then this fidelity to moral–psychological facts ought to undermine consequentialism. It ought to remove the question to which consequentialism offers an answer; or it ought to make clear that the correct answer to that question is not the consequentialist one. But neither of these things happens. Or so at any rate I wish to argue.

My argument is developed in two sections. In Section 2 I introduce the principal psychological claims that opponents of consequentialism have wanted to stress. And in Section 3, I try to show that even if these claims are admitted, consequentialism remains a plausible moral doctrine. I describe the attempt to be faithful to the moral psychology of human beings, for want of a better term, as moral–psychological realism: the approach is realist to the extent that it seeks to describe the psychology of moral agents as it really is; it is committed to avoiding all theoretical simplification and distortion. Moral–psychological realism, I argue, does not marginalize the question addressed by consequentialists and neither does it make the consequentialist answer to that question particularly implausible.[1]

The argument of the paper is of practical import. It bears on the question of what presence consequentialism ought to have in the deliberations of the virtuous agent. I conclude, in an epilogue, with a brief sketch of the answer suggested. Consequentialism ought not to make an impact, explicit or implicit, on every decision. All it ought generally to enjoy is what I describe as a virtual presence in the deliberation that produces decisions.

## 2. Moral–Psychological Realism

I shall mention three broad claims that are emphasized by the opponents of consequentialism in their search for moral–psychological realism.[2] I identify each claim in a negative way, by reference to what it denies. The first I describe as anti-atomistic, the second as anti-economistic and the third as anti-rationalistic. I am happy to endorse each of these claims and, since they are common ground with the opponents of consequentialism, I will not go in any detail into the reasons for defending them. Suffice it here to describe the claims involved. There are other moral–psychological claims that anti-consequentialists defend but we will restrict ourselves here to the three I mention. The lesson that emerges with these claims is likely to survive any further claims that anti-consequentialists can plausibly introduce.

The first, anti-atomistic claim is that the deliberative considerations that motivate people – the considerations that give people intrinsic (though perhaps defeasible) reason to act – often involve other individuals essentially; they are considerations that obtain and motivate agents only in the presence of those other individuals. The springs of human action are not solipsistic or atomistic. They involve other people immediately.[3]

Suppose that I go to some trouble to help a friend, Mary, who is in financial difficulty, and that I do this for reasons of friendship. There are two possible pictures of how my deliberation goes in such a case, one atomistic, the other not. The first picture of my reasoning is this. I am concerned with the general value of having friends and of helping those who happen to be my friends – this gives me intrinsic reason to act – and, seeing that Mary is a friend, I judge that I can best promote that value by now helping her: she falls, as luck would have it, on the trajectory of my general friend-helping project. Under this picture, Mary does not figure essentially in the consideration that gives me intrinsic reason to act: her part is not to motivate me, as it were, but only to provide me with the occasion for satisfying the independent motivation to help anyone who happens to be a friend.

This picture is distinctively atomistic. I am supposed to be resourced with all the motivation I need by the consideration of values that do not require the existence of any particular individual, let alone Mary: I am fully resourced by the consideration of the good of having friends and of helping those who happen to be my friends. If another individual breaks in on my consciousness, then, it is only by way of offering me an opening for the pursuit of projects that do not essentially involve the other. The project may be that of helping whoever are my friends but equally it may be that of enjoying company, achieving sexual satisfaction, or having a certain degree of status or power.

The anti-atomistic claim denies, surely plausibly, that people's reasonings are generally like this. It adopts a different picture of my reasoning in regard to Mary, holding that if people pursue the good of their friends, that is normally because they care for those particular individuals, and not because of the independent appeal of the general value of having and helping friends. The reasoning involved is not of the form: helping needy friends is good; Mary is a needy friend; so I should help her. Rather it takes the form: my friend, Mary, is in need; so I should help her. The consideration that moves me essentially involves Mary, as it essentially involves friendship. Mary does not just happen to provide the opportunity for satisfying my friend-helping project. Mary provides the opportunity, as only Mary can, for satisfying a project that essentially engages her.

Under the first picture of my reasoning, the consideration that intrinsically moves me is there also to move the counterfactual self which confronts, not Mary, but some other indiscernible individual, or even an illusion. My counterfactual self is moved, as I am moved, by the value of having and helping friends. Under the second picture, the common motivation fails. The consideration that intrinsically moves me is that my friend, Mary, is in need and this is not a consideration that can move the counterfactual self envisaged. It is a consideration that involves the presence of that other person, Mary, in my mind; it is a consideration that extends my mind to the world of that other person. The thought that binds me to Mary, the thought that moves me to action on her behalf, could not exist – could not logically exist – if she did not exist. It is not a thought that I can share with the counterfactual self which inhabits a Mary-less world.

The anti-atomistic claim is clearly true to the everyday phenomenology of motivation. People do not generally pursue friendship or company or even sex in such a way that others are well described as offering them contingent occasions for the satisfaction of the projects involved. People enter one another's minds and lives in the more intimate manner described. If anyone ever thought that things are otherwise – or ever thought anything that supposes that things are otherwise – then that can only have been, plausibly, because of adopting a false picture of the mind: a picture under which the contents of our thoughts, and therefore the considerations that ultimately move us, cannot presuppose the reality of entities outside our skin. But that is matter for speculation elsewhere.[4]

I said earlier that there are other moral–psychological claims made by anti-consequentialists besides the three under consideration here. To many eyes the most striking omission from my list will be the sort of anti-impersonalism which Bernard Williams has defended in a number of places.[5] But that omission may be intelligible in the light of this discussion of the anti-atomistic claim. What the discussion shows is that though it is the fact that someone is a friend which explains and justifies my partiality towards her, the abstract value of having friends and helping whoever are my friends is not at the motivational source of my commitment. The commitment is given by the more or less brute fact of our acknowledged attachment to one another, the brute fact of our acknowledged affection: it is not derived on either side from a contemplation of the value of having and helping friends. Were the commitment derived on either side from that abstract sort of starting-point, then it would not really be friendship.

Anti-impersonalism can be seen as a more general doctrine than anti-atomism: one which anti-atomism exemplifies but does not exhaust. The essential anti-impersonalist message is that every human agent finds himself or herself with more or less brute commitments already in place – more or less brute projects and desires – and that the abstract categories that may be invoked to explain and justify continuing with those commitments should not be thought of as abstractly motivating the agent. Human beings have to live with a certain contingency as they embark on moral reflection and moral reform. They should not think that any commitments they make should be capable of being derived from the consideration of abstract values. The sort of agent who could reconstruct his or her every commitment on an abstract basis would not display the moral psychology of the human subject; he or she would be an angel or a devil. The point is salient in the case we have discussed. The sort of agent who came to have and to help friends, not on the basis of more or less passive feeling, but on a strategic, value-oriented basis would be something less, or perhaps something more, than a friend in the ordinary sense of the term.

So much for the anti-atomistic claim which opponents of consequentialism often emphasize. A second claim that they rehearse is anti-economistic rather than anti-atomistic in import. It holds that not only do other individuals enter essentially into the considerations that move human agents, those intrinsically moving considerations are often pluralistic and incommensurable in character. The idea is that it may be an intrinsically moving consideration for me – it may be, by my lights, a value – that something will help Mary, or that it will accord with a promise, or that it will reduce the pain and suffering in the world, without it being the case that I can meaningfully rank that consideration against all other evaluative considerations. There is no guarantee that among the values I countenance, I can establish rankings of importance. There is no guarantee that my values can all be connected by a unified system of weights.[6]

The claim here is that for all we can know value is essentially pluralistic. More strongly indeed, the claim is that when we consider the irresolubility of many of the moral dilemmas that people face, we must admit that for at least certain pairs of values, there is no meaningful way of establishing that one is more important than the other. We have no doubt, and reasonably have no doubt, that the saving of another's life is more important than the enjoyment of a quiet evening. And equally we may have no doubt that the preservation of a friendship is more important than seeing the first night of a new play, or that helping one's children through a difficult patch is more important than getting an article finished on time. But there are other cases, so the anti-economistic claim goes, where such intuitions run out: cases where there is no saying, by any criteria of what is reasonable, which of two values is the more important.

I describe this second claim as anti-economistic, because it rejects the most striking assumption built into the model of *homo economicus*: the assumption that there is a single function which the economic agent is expected to maximize and that there are no discontinuities of the kind that a value-pluralism would support. The anti-economistic claim would fail dramatically if there were a single property to engage the agent: a property like the pleasure or happiness which utilitarians would have him or her promote. But it would also fail under weaker versions of the axioms of consumer choice theory. Under these axioms there may be different properties that can count as values for an agent but they must all satisfy conditions – conditions often expressed with the help of the notion of indifference curves – which mean that every value is weighted against every other. The anti-economistic claim rejects the picture associated with utilitarianism but also this latter sort of vision.

The third moral–psychological claim that is often stressed by opponents of consequentialism is anti-rationalistic in character. The claim is that human agents are not on the whole a rationalistic or calculative species and that it is an essential part of realizing many of the things they cherish that they do not calculate rationalistically about what they should do. Granted that the values which count with human agents may involve other people essentially. Granted also that those values may not be entirely commensurable with one another. Still, it could be that so far as certain values go, human agents typically concern themselves with those values in a calculative, actuarial spirit: they itemize alternatives, draw up an inventory of possible consequences, assign appropriate probabilities to those consequences, and then try to work out which alternative does best by the values they embrace. The third anti-rationalistic claim denies that this is what human beings typically do and, more importantly, denies that this is something they could do consistently with satisfying their concerns.[7]

Two points will serve to illustrate the anti-rationalistic thought. The first is that many of the virtues which people recognize and admire in one another are such that their exercise is incompatible with rationalistic calculation. The point will be obvious with a virtue like that of spontaneity. But it also applies with more characteristically ethical virtues. Take generosity, for example. To be a generous person is, presumably, to be someone who tends spontaneously to want to help when a suitable cause presents itself: to be someone who does not first have to consider the cost before being prepared to make a commitment. But this means that to be a generous person is to be someone who is disposed not to calculate about all relevant costs and benefits when confronted with certain demands.

A critic may respond that being generous still requires calculation, since the generous agent will at least calculate about how far a given option is generous and about how important generosity is in the case on hand; what makes him or her generous, so it may be said, is the fact that he or she gives generosity a high ranking. But this will not do either. Someone who thinks of options in terms of whether they are generous or not, and who takes account of their generosity-scores in determining which to enact, is not a generous person. Rather he or she is an agent who aspires to behave as a generous person would; an agent who aspires, not to exemplify generosity, but to mimic the generous person.

I promised to make a second point in illustrating, and indeed in vindicating, the anti-rationalistic claim. That is that many human relationships are built upon the common belief that each will treat the other in a manner that is incompatible with rationalistic calculation. If we count one another as friends then we will individually assume, and assume that the other assumes – and so on, perhaps, up the familiar hierarchy – that if either of us is in certain sorts of difficulty, then the other will be prepared to offer whatever help may be called upon, without having to calculate about the cost–benefit ratio. The sign of affection and friendship is precisely that calculation is not going to be needed in a case like this. Again, if we count one another as mutually respectful members of a well-ordered society, then we will individually assume, and assume that the other assumes, that if either proves to be an obstacle in the other's pursuit of a project, then the other will be prepared, without calculating about whether this is really to his or her advantage, to respect the rights of the obstructing person. As affection requires unthinking consideration, so a dispensation of respect requires a more or less unthinking recognition of rights. I am not truly befriended if the other must do sums before being prepared to help me out. I am not truly respected if the other must do a check-list of the advantages and disadvantages before being prepared to refrain from interfering with me.[8]

This completes my sketch of the sorts of psychological points to which opponents of consequentialism direct us. I think that these points are fundamentally correct, though no doubt they need to be developed and qualified in various ways. The psychology of human beings, in particular their moral psychology – their psychology, so far as this bears on moral matters – is non-atomistic, non-economistic and non-rationalistic. It allows other subjects to enter directly into people's individual motivation. It allows the values that motivate people to be only partially commensurable with one another. And it allows and encourages people to be spontaneous, non-calculative servants of those values.

## 3. Back to Consequentialism

The fact that opponents of consequentialism rehearse the psychological points discussed tells us about their understanding of consequentialism. If they rehearse the anti-atomistic point, or something akin to that point, this is because they conceive of consequentialists as believing that the moral agent brings a moral project, a project of promoting the good, to his or her experience of other people, rather than allowing those people to enter directly into the formation of that developing project: those people do not figure in the considerations that motivate him or her but represent opportunities to pursue the preconceived commitment. Again, if anti-consequentialists rehearse the anti-economistic point, this is because they think that consequentialists are committed to the view that every moral decision is subject to more or less algorithmic adjudication; there are no indeterminacies, so they think, that consequentialists can leave open in matters of right and wrong. And finally, if anti-consequentialists rehearse the anti-rationalistic point, this is because they imagine that consequentialists picture moral decision-making as an exercise in actuarial computation. F. H. Bradley gave vivid expression to that picture more than a century ago. 'So far as my lights go, this is to make possible, to justify, and even to encourage, an incessant practical casuistry; and that, it need scarcely be added, is the death of morality.'[9]

If we accept the moral–psychological realism of the last section, then we must certainly reject the sort of doctrine which opponents have wished upon consequentialists. I do not think that the most important consequentialists have ever defended this sort of doctrine but I will not try to argue that here. What I want to do instead is to accept the moral–psychological points made by anti-consequentialists and then to ask whether those points really undermine the core, consequentialist message. The issue breaks down into two sub-issues. First, does moral–psychological realism leave in place the question that consequentialism addresses? And, second, does it leave the plausibility of the consequentialist answer to that question unaffected? I shall argue that the answer in each case should be 'Yes'.

Take the first sub-issue. The question addressed by consequentialism, as I stressed in the introduction, bears on what it is that makes an option the right one for an agent to choose or to have chosen. Does this question survive the endorsement of moral–psychological realism? The second, anti-economistic assumption suggests that there may not always be a fact of the matter about whether this or that option is the right one; there may be incommensurable value-sets involved on the different sides so that, whatever our theoretical inclinations, we have to regard the issue as indeterminate. But no one thinks that such incommensurability is total and most will agree that in the general run there is indeed a fact of the matter as to which of a given set of options is the right one for the agent to choose or to have chosen. So does moral–psychological realism leave in place, in those cases, the question addressed by consequentialism? Does it leave in place the question as to what it is that makes one option right and another wrong?

Of course it does. No moral–psychological realist would want to say that whatever the ordinary sort of agent does is right: right, so long as he or she is not atomistic or economistic or rationalistic in his or her approach to things. But if some of the things such an agent does are right, and some are not, then we are naturally left with the question as to what it is that makes one option right and another wrong. To reject that question would be to reject the very idea of moral evaluation. It would be to endorse a sort of moral quietism: a view to the effect that provided people conform to the everyday realist image of agency – provided they are not misled into adopting an atomistic or economistic or rationalistic posture – then whatever they do is fine, whatever they do is right.

But if moral–psychological realism leaves in place the question addressed by consequentialism, the next issue – the second sub-issue that we distinguished – is whether it allows us still to consider consequentialism as a plausible answer to the question. This is a slightly more tricky matter and will require us to make a detour. We need to say more on what exactly the consequentialist answer to the question is: in particular, more on what the consequentialist answer is, given that the answer cannot entail any form of atomism, economism or rationalism.

The consequentialist answer, as I described it in the introduction, is that the right option is that which promotes the good, that which promotes value, better than any of the alternative options. In defending this answer, consequentialists do two things. First, they offer a gloss on what constraints must be satisfied by any property for it to count as a value; although they are not committed to any particular theory of the good, they do make important assumptions about the sort of thing that the good is.[10] And second, they give an account of what it is for an option to promote the good, what it is for an option to promote value.[11]

Consequentialists assume, like most others, that if it is of value to promote a state of affairs that involves a particular individual, say a particular person, then it is of value to promote any state of affairs that differs from the original only in the identity of the individual or individuals involved. If it is of value, *pro tanto*, to promote Mary's talents, then it is of value to promote the talents of anyone who is exactly – or indeed relevantly – similar to Mary. If it is of value, *pro tanto*, for me as a parent to promote the welfare of my children, then it is of value for any parent to promote the welfare of their children. And so on. This is to say, as it is often put, that value – the good – is essentially universalizable: if any state of affairs involving a particular is of value, then so is the universalized counterpart of that state of affairs. Alternatively, no particularized state of affairs is of value except so far as the universalized counterpart is of value. The particularized state need not be desired by the agent because of a desire for the realization of the universalized counterpart – that is the lesson of anti-atomism and, more generally, anti-impersonalism – but consistently with deriving and valuing the particularized state, the agent cannot deny the value of the universalized counterpart; consistently with pursuing the welfare of his or her friends, for example, he or she cannot deny the value, *pro tanto*, of any agent's pursuing the welfare of their friends.

Having laid down this constraint on value, consequentialists go on to say that an option is right if and only if it does better than alternatives in promoting value, universalizably conceived. But what exactly does promotion involve? Here there are two broadly contrasting explications available. One would hold that to promote value is to act in the way that, as it happens, realizes more *actual value*: it realizes more value than alternative actions would actually have done. The other would say that to promote value is to act in the way that, given subjective or at least reasonable expectations at the time of choice, maximizes *expected value*. The idea will be familiar from Bayesian decision theory: roughly, the agent chooses that option such that if we give a value-score to each possible outcome, discount that score by the probability associated with the outcome, and then add up the discounted scores, the option does better than alternatives in this regard.[12]

I believe that promotion should be defined, not by the maximization of actual value, but by the maximization of expected value; I ignore in the present context the difference between maximizing reasonably and maximizing subjectively expected value.[13] Maximizing expected value gives a more plausible reading to the notion of promotion, because it means that whether an agent promotes value and does right is fixed at the time of choice, and that an agent cannot be vindicated just by how the world happens to go in the wake of his or her action. It means, for example, that the doctor who gives a life-threatening drug to someone for the relief of a minor pain will not have done right, just because he or she is lucky and no serious harm is caused in the course of relieving the pain.

Back now to the main line of argument. We are concerned with agents who are explicitly conceived in the moral–psychological image of the last section: other people are essentially involved in the values that move these agents; the values that move, and ought to move, them are not at all necessarily commensurable; and the agents do not pursue those values in a rationalistic, calculative manner. We are focused in particular on choices where the values involved are not, as it happens, incommensurable with one another: that is, we are focused on more or less run-of-the-mill choices, not on tragic dilemmas. The issue we face is this. What is it about a particular option in such a choice that makes it the right option for the agent to take or to have taken? And, specifically, does moral–psychological realism mean that there is something particularly implausible about the consequentialist idea, as just explicated, that what makes a particular option right is that it does better than alternatives in the promotion of value, universalizably conceived?

Intuitively, the answer is that moral–psychological realism does not render the consequentialist answer implausible. Such realism means that I will and should carry on in my everyday life without abstracting from the other people to whom I happen to be committed, without aspiring to identify a complete, commensurable set of values, and without rationalistically calculating about how best to satisfy the values I recognize in my every action. But the realism allows that, carrying on in this way, I may or may not actually do what is right. What criterion is relevant, then, to determining whether any choice is right? What criterion might others reasonably use in adjudication of my behaviour? What criterion might I reasonably use myself, when I review my performance? The consequentialist gives us an answer and nothing that is involved in being a moral–psychological realist impacts on the plausibility of that answer.

The consequentialist says, to be a little more vivid, that while it is fine in principle for me to be faithful and partial towards my friends – assuming friendship is a good thing – I must recognize that this is true of any individual and that my own fidelity and partiality will be valuable, therefore, only if it does not impact negatively on friendship in the event – the unlikely event, as it must surely be – that there is something else that I can do instead of cherishing my own friends by way of realizing this value overall. It will impact negatively on value in general in the event – again, surely the unlikely event – that what I can best do to promote value requires disowning my friends or distancing myself from them. One reason why it is unlikely that there is something I can do to promote friendship or value that involves detachment from my own friends is that any alternative that requires this detachment is an alternative under which I must make myself into something less, or something more, than an ordinary human being. The likelihood of my being able, under this transformation, to promote friendship or value in the required measure cannot be very high.

This may be enough to show that the consequentialist answer to the question of rightness remains plausible – certainly it retains whatever plausibility it was previously assigned – under moral–psychological realism. But it may be useful to do one thing more in supporting this claim. It may be useful to consider alternatives to the consequentialist answer and to ask if any of these is particularly favoured by moral–psychological realism.

As there are two elements in the consequentialist answer to the question of rightness, so there are two broadly differing ways of rejecting consequentialism.

The first way of rejecting it would be to deny that there are universal values relevant to every human choice – strictly, to every run-of-the-mill human choice – and to deny, therefore, that the consequentialist answer bears on such choices. Perhaps Philippa Foot means to direct us to this sort of approach, when she argues that even if universalized states of affairs have values, one non-consequentialist viewpoint would deny that the moral agent should be concerned with such states of affairs; it would see the agent as properly concerned only with the options before him or her.[14]

Whatever there is to be said for this approach, however, moral–psychological realism does not particularly support it against consequentialism. For there is nothing in such realism to deny the plausibility of the following line of argument. No matter what the basis for judging that an agent's choice is right, every theorist is committed by that judgment to the value of any relevantly similar agent making the corresponding choice; and for any theorist, therefore, it is relevant to look at how the agent's choice does in relation to that universalized state of affairs: it is relevant to consider the impact of the choice on that universal value. If I am said to have done right in helping a friend, to return to our example, then it is implied, under universalizability, that it is good for anyone in my position to help their friends and it is surely relevant, then, to consider how my action relates to the realization of that universal value. What if my helping this friend means, for whatever complicated reason of circumstances, that it becomes more difficult and more unusual for people in my position to help their friends? Surely that is at least a *prima facie* relevant consideration? This line of argument

remains forceful, even if moral–psychological realism is admitted. And so the admission of moral–psychological realism does nothing to favour the first way of rejecting consequentialism.

The second way of rejecting consequentialism would be to admit that there are always universal values that are relevant in determining which option before an agent is right but to deny that those values have the sort of relevance envisaged by consequentialism: to deny, even in run-of-the-mill cases where commensurability obtains, that an option is right in virtue of promoting the relevant values. What might a non-consequentialist of this character say? The general line, though not always expressed in this way, goes as follows. What makes an option right is not the fact that it promotes the relevant values better than alternatives – though the right option will often do that – but rather the fact that it does best by way of exemplifying a concern for that value: the fact that it does best in bearing witness to the value or, as we may say, in honouring it.

What might it be to honour a value without promoting it? Assume that promise-keeping is a value and consider a case where it is the only relevant value in a decision I face. Suppose that if I keep the promise then, for whatever reason, it is to be expected that others will keep their promises less rather than more: breaking this particular promise would promote the universal value of promise-keeping better than keeping the promise. It is plausible in such a case that if I am to honour the value of promise-keeping then, despite the promotional costs, I should still keep my promise. Why would this seem to be the best way to honour value? Perhaps because the way for me to promote promise-keeping if I look only to my own affairs – if I keep my nose out of other people's business – is indeed by keeping my own promises. Or perhaps because the way for me to show others what it is best for each of us to do in the promotion of promise-keeping is by keeping my own promises.

The second way of rejecting consequentialism would be to hold that while universal values are always relevant to a choice, what makes an option right is the fact that it best honours the relevant values, not the fact – if it is a fact – that it best promotes them. This way of rejecting consequentialism is endorsed by those who say that the concern of the moral agent should often be with agent-relative values: with values like the value of keeping personal promises or looking after personal friends; the moral agent need not always focus on the corresponding agent-neutral values: with the value of everyone's keeping their promises or with the value of everyone's looking after their friends. The question for us is whether moral–psychological realism does anything to support this line and, in that way, to render consequentialism more or less implausible.

I see nothing in moral–psychological realism to support the view that rightness is determined by the honouring of relevant values rather than by their promotion. Unlike consequentialism, this view gives rise to a variety of difficulties and moral–psychological realism serves in no way to redress the balance of difficulty.[15] The honouring-values view is not well equipped to deal with cases of risk and uncertainty, cases where there is no knowing which properties will be realized by which option. In this respect it contrasts with consequentialism, at least when promotion is defined as maximizing expected value, for the decision-theoretic notion of expectation is designed, specifically, to take account of such cases. Again, the honouring-values view offers us a model of how best to respond to matters of concern – universal values – which, unlike the promotional model, is not replicated elsewhere: in other cases we generally suppose that the way to respond to matters of concern is by seeking to promote the things we care for; thus, the honouring-values view is *sui generis* and, to that extent, implausible. There is nothing in moral–psychological realism that would tend to eliminate or reduce such difficulties and thereby to shift the burden of proof towards consequentialism.

And so to our conclusion. I believe that the opponents of consequentialism have done us a service in developing a variety of fruitful lines of observation and argument in the area of moral psychology; I hope that the discussion in the last section will have given a sense of that achievement. But I think that they are mistaken in suggesting that consequentialism cannot survive the development of a detailed moral–psychological picture of human agents. Even if we are explicit in endorsing the anti-atomistic, the anti-economistic and the anti-rationalistic points that have been well made in moral psychology, the question addressed by consequentialism remains in place and consequentialism itself remains a plausible answer to that question. If consequentialism is to fall, it will not be for an intolerance of such moral–psychological foundations.

## Epilogue: Virtual Consequentialism

The argument that we have conducted suggests that the virtuous agent ought in general to remain faithful to his or her instincts and ingrained habits, only occasionally breaking with them in the name of promoting the best consequences. But how are we to conceive of this pattern of spasmodic resort to consequentialist reasoning? There are two common images of how consequentialism might be present in practical reasoning and neither is attractive. One is the picture according to which it should be explicitly present all the time; this is the image that anti-consequentialists have traditionally wished on their opponents. The other is the picture according to which it should be at least implicitly present all the time: it should be there, more or less unconsciously, in the day-to-day reckonings of the virtuous agent.

If the image of the consistently explicit consequentialist resembles that of a moral actuary – that of someone who practises, in Bradley's words, 'an incessant practical casuistry' – the image of the agent who remains always implicitly consequentialist in his or her reasonings resembles that of a person who displays an incessant false consciousness. It suggests that the virtuous agent ought to be at once looking and not looking at the expected consequences of his or her doings: ought to satisfy the demands of moral psychology by being generally spontaneous while keeping his or her consequentialist conscience clean by continually running quick checks on how the moral sums are coming out.

Happily, the argument in this paper does not support the idea that the virtuous agent ought always to be at least implicitly consequentialist in his or her deliberations. There is a third way in which consequentialism can be present in the deliberations of such an agent and it fits very nicely with our argument. Explicit and implicit presence are both forms of real presence; the difference between the two is a difference of the conscious–unconscious or conscious–preconscious variety. The third alternative is something that I describe as virtual presence. Consequentialism will be virtually present in the deliberations of the virtuous agent to the extent that while it is actually absent, while the agent conducts herself or himself without any explicit or implicit reference to consequentialist calculation, there are various triggers in place whose activation will cause the agent to reflect in a consequentialist manner on his or her mode of behaviour.

The triggers I have in mind will be associated with certain thresholds of aspiration. I assume that any virtuous agent will have an overall sense of how in general her or his action ought to be impacting on the world: with what effects and to what good. The agent will continue in unreflective, spontaneous mode – punctuated, no doubt, by periods of self-examination – so long as those thresholds are being attained or surpassed. But as soon as his or her behaviour fails to satisfy those constraints, the red lights will go on and the agent will reconsider performance in a more reflective, consequentialist spirit. The metaphor of the red lights may be misleading, for I do not mean to suggest that there is going to be an easy formula or algorithm whereby an agent can tell when it is appropriate to suspend his or her spontaneous inclinations. I believe that it may require an unusual degree of practical wisdom, an unusual degree of sensitivity and resolution, to be able to know when it is time to reflect and consider.

At this point the tradition of consequentialism rejoins the tradition of virtue-theory. To be virtuous, by all accounts, is to be a lover of the good. Virtue-theorists suggest that it is difficult and perhaps impossible to be a lover of the good – or at least to be a good lover – in the absence of certain habits of the mind and heart: the habits we traditionally designate as virtues. The virtues are certainly useful for combating failures of practical reason: failures such as weakness of will, compulsiveness, and the like.[16] But virtue-theorists generally suggest that the virtues are useful, and even necessary, in other ways too.[17] The consequentialist is well disposed, in at least one respect, to agree: to agree, not just in a spirit of concession, but with a view to learning from their work. He or she can associate the sensitivity which the model of virtual consequentialism postulates with the sort of practical wisdom – the *phronesis* – that virtue-theorists celebrate. And the consequentialist can look to the tradition of virtue-theory to see if there are lessons to be learned in the many discussions that have been devoted to the nature and function of this faculty.[18]

*Research School of Social Sciences,*
*Australian National University, Canberra*

## Notes

1 The paper is conceived as a review, in good part, of arguments developed elsewhere and so I shall often be referring to published work of my own. But though the paper serves as a review of earlier arguments, it recasts and reconnects those arguments in what I hope will be an interesting manner.

2 For lists of references on these individual claims I refer in what follows to other discussions of my own, discussions that develop the arguments reviewed – and recast – here. A good selection of the relevant references are to be found in Sam Scheffler (ed.), *Consequentialism and its Critics* (Oxford: Oxford University Press, 1988). See also many of the essays in Philip Pettit (ed.), *Consequentialism* (London: Dartmouth Press, 1993).

3 For references on this first claim, and for further discussion, see Philip Pettit, 'Social Holism and Moral Theory', *Proceedings of the Aristotelian Society*, vol. 86, 1985–86, pp. 173–97 and, in particular, 'The Paradox of Loyalty', *American Philosophical Quarterly*, vol. 25, 1988, pp. 163–71. Perhaps the most important paper relevant to the theme is Michael Stocker, 'The Schizophrenia of Modern Ethical Theory', *Journal of Philosophy*, vol. 73, 1976, pp. 453–66.

4 See Philip Pettit and John McDowell (eds), *Subject, Thought and Context* (Oxford: Oxford University Press, 1986) for essays in introduction of this issue. See also Philip Pettit, *The Common Mind: An Essay on Psychology, Society and Politics* (New York: Oxford University Press, 1993).

5 See, for example, his discussion of integrity in J. J. C. Smart and Bernard Williams, *Utilitarianism: For and Against* (Cambridge: Cambridge University Press, 1973); see too Sam Scheffler, *The Rejection of Consequentialism* (Oxford: Oxford University Press, 1982).

6 This claim appears in many writers but it is particularly associated with Isaiah Berlin, Stuart Hampshire and Bernard Williams. See Scheffler, op. cit. note 2. And see also many of the papers in Amartya Sen and Bernard Williams (eds), *Utilitarianism and Beyond* (Cambridge: Cambridge University Press, 1982).

7 For references on this third claim see Philip Pettit and Geoffrey Brennan, 'Restrictive Consequentialism', *Australasian Journal of Philosophy*, vol. 64, 1986, pp. 438–55. See too Peter Railton, 'Alienation, Consequentialism and the Demands of Morality', *Philosophy and Public Affairs*, vol. 13, 1984, pp. 134–71.

8 On the issue with rights and with respect see Philip Pettit, 'The Consequentialist Can Recognise Rights', *Philosophical Quarterly*, vol. 35, 1988, pp. 537–51 and 'Consequentialism and Respect for Persons', *Ethics*, vol. 99, 1989, pp. 116–26. A crucial paper is David Lyons, 'Utility and Rights', *Nomos*, vol. 24, 1982, pp. 107–38. For a recent review of relevant literature see Alan Hamlin, 'Rights, Indirect Utilitarianism, and Contractarianism', *Economics and Philosophy*, vol. 5, 1989, pp. 167–88.

9 *Ethical Studies* (1876) (Oxford: Oxford University Press, 1962), p. 107.

10 Wouldn't it be better policy for the consequentialist to abstain from all comment on the theory of the good, and not just from particular commitments there, and to say that an option is right just in case it promotes whatever states of affairs we think are good? No, it wouldn't. The problem with this looser formula is that other positions in the theory of the right can be made to conform to it; they can be represented, under that formula, as consequentialist in character. For example, the honouring-values view discussed later can be represented as holding that the right option for an agent is that which promotes the non-universal good of that particular agent's honouring the values in question: of that agent's keeping his or her hands clean.

11 For a critique of my perspective see David McNaughton and Piers Rawling, 'Honoring and Promoting Values', *Ethics*, vol. 102, 1992, pp. 835–43.

12 On these issues see Frank Jackson, 'Decision-theoretic Consequentialism and the Nearest and Dearest Objection', *Ethics*, vol. 101, 1991, pp. 461–82 and Peter Menzies and Graham Oddie, 'An Objectivist's Guide to Subjective Value', *Ethics*, forthcoming.

13 I also ignore theories which say that promotion need only involve, not maximization, but satisficing: that is, doing enough, by some contextually given criterion, in the realization of value. See Michael Slote, *Beyond Optimizing: A Study of Rational Choice* (Cambridge, MA: Harvard University Press, 1989). I have criticized that view in 'Satisficing Consequentialism', *Proceedings of the Aristotelian Society*, supp. vol. 58, 1984, pp. 165–76.

14 See Philippa Foot, 'Utilitarianism and the Virtues', *Mind*, vol. 94, 1985, pp. 196–209. See too Will Kymlicka, 'Rawls on Teleology and Deontology', *Philosophy and Public Affairs*, vol. 17, pp. 173–90.

15 On such difficulties see Philip Pettit, 'Consequentialism' in Peter Singer (ed.), *A Companion to Ethics* (Oxford: Blackwell, 1991). For an account of the difficulties that arise in political thought see John Braithwaite and Philip Pettit, *Not Just Deserts: A Republican Theory of Criminal Justice* (Oxford: Oxford University Press, 1990), pb. edn 1992, and Philip Pettit, 'The Contribution of Analytical Philosophy' in R. E. Goodin and Philip Pettit (eds), *A Companion to Contemporary Political Philosophy* (Oxford: Blackwell, 1993).

16 See Philip Pettit and Michael Smith, 'Practical Unreason', *Mind*, vol. 102, 1993.

17 We can distinguish three claims: that virtue is essential or useful in practice for doing what is right; that virtue is epistemologically necessary or useful, being associated with the perception of what is right; and that virtue is ontologically involved in the constitution of rightness: an action is right, as an a priori matter, just to the extent that it is the sort of thing that a virtuous agent would choose. The tradition of invoking virtue always defends the first claim, as everyone should, often maintains the second, and occasionally supports the third. The second and third claims are naturally associated with one or other of the two varieties of anti-consequentialism discussed in the third section of the paper. But the connections are not straightforward, as indeed the present discussion should indicate.

18 This paper was written during a period when I held successive visiting appointments in Paris: first, in September 1992, at the *Ecole des Hautes Etudes en Sciences Sociales* (CEMS) and then, in October 1992, at the *Ecole Polytechnique* (CREA). I am grateful to both institutions for their support. I am also grateful to Monique Canto-Sperber for helpful remarks on an earlier draft. A French version of this paper is forthcoming in the *Revue de métaphysique et de morale*.