

# WILEY

---

Consequentialism and Cluelessness

Author(s): James Lenman

Source: *Philosophy & Public Affairs*, Autumn, 2000, Vol. 29, No. 4 (Autumn, 2000), pp. 342-370

Published by: Wiley

Stable URL: <https://www.jstor.org/stable/2672830>

## REFERENCES

Linked references are available on JSTOR for this article:

[https://www.jstor.org/stable/2672830?seq=1&cid=pdf-reference#references\\_tab\\_contents](https://www.jstor.org/stable/2672830?seq=1&cid=pdf-reference#references_tab_contents)

You may need to log in to JSTOR to access the linked references.

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



Wiley is collaborating with JSTOR to digitize, preserve and extend access to *Philosophy & Public Affairs*

JSTOR

JAMES LENMAN

## Consequentialism and Cluelessness

### I. INTRODUCTION

Shelly Kagan in *Normative Ethics* said:

Perhaps the most common objection to consequentialism is this: it is impossible to know the future. This means that you will never be absolutely certain as to what all the consequences of your act will be. An act that looks like it will lead to the best results overall may turn out badly, since things often don't turn out the way you think they will: something extremely unlikely may happen, and an act that was overwhelmingly likely to lead to good results might—for reasons beyond your control—produce disaster. Or there may be long term bad effects from your act, side effects that were unforeseen and indeed unforeseeable. In fact lacking a crystal ball, how could you possibly tell what *all* the effects of your act will be? So how can we tell which act will lead to the best results overall—counting *all* the results? This seems to mean that consequentialism will be unusable as a moral guide to action. All the evidence available at the time of acting may have pointed to the

Work on this paper was assisted by a year of study leave generously funded by the University of Glasgow and the United Kingdom Arts and Humanities Research Board. Audiences at the Universities of Dundee, Haifa, and Bar-Ilan furnished useful comments and encouragement. Peter Clark, Jim Edwards, Bob Hale, James Harris, Brad Hooker, Gary Kemp, Dudley Knowles, Thad Metz, John O'Neill, Michael Ridge, Adam Rieger, Elizabeth Telfer and Nick Zangwill read and commented on earlier drafts. Peter Clark and Mark Alford offered particularly valuable help in my efforts to get my comments on science in sections II and III right. Katherine Hawley and Rex Whitehead also helped me out here. John Broome and Wlodek Rabinowicz gave me useful advice when I was working on the reasoning of section IV. An anonymous reader for *Philosophy & Public Affairs* has also helped me to improve the paper. Many thanks to all the foregoing.

© 2000 by Princeton University Press. *Philosophy & Public Affairs* 29, no. 4.

conclusion that a given act was the right act to perform—and yet it may still turn out that what you did had horrible results, and so in fact was morally wrong. Indeed, it will never be possible to say for sure that any given act was right or wrong, since any event can continue to have further unseen effects down through history. Yet if it is impossible to tell whether any act is morally right or wrong, how can consequentialism possibly be a correct moral theory?<sup>1</sup>

We may call this the *Epistemic Argument* against consequentialism. It is often briefly aired and discussed but generally not taken very seriously and has received relatively little by way of sustained attention.<sup>2</sup> My aim in this essay is to contribute to putting this right. For if the arguments that follow are sound, the Epistemic Argument goes very deep and needs to be taken very seriously indeed.

By ‘consequentialism’ I mean the view that the rightness of an action or, more generally, let us say, a *policy*, is a matter—entirely a matter—of the goodness of its consequences. The sort of consequentialism with which I am concerned says of each person that *that* person should adopt policies with a view to the goodness of overall consequences. By ‘consequences’ I mean *all* the consequences: what we ought to do is maximally promote the *overall* good. ‘Policies’ here can refer to the performance (or omission) of actions, the adoption of rules, decision procedures (possibly not themselves consequentialist) or plans of life and the cultivation of traits of character. We might call this *pure consequentialism*. Thus, pure consequentialist theories include theories that recommend the adoption of some decision procedure that is not consequentialist at all, as well as those that recommend some subjectivized consequentialist deliberative strategy stressing the expected rather than the actual value of consequences, provided any such recommendation is made on, ultimately, pure consequentialist grounds. In what follows, in particular in section V, I will argue that such variant forms of pure consequentialism do not successfully sidestep the Epistemic Argument.

1. Shelly Kagan, *Normative Ethics* (Boulder: Westview Press, 1998), p. 64.

2. Notable exceptions in recent literature are Alastair Norcross, “Consequentialism and the Unforeseeable Future,” *Analysis* 50 (1990): 253–56; Robert L. Frazier, “Act-Utilitarianism and Decision Procedures,” *Utilitas* 6 (1994): 43–53; Frances Howard-Snyder, “The Rejection of Objective Consequentialism,” *Utilitas* 9 (1997): 240–48; and Dale Miller, “Actual-Consequence Act-Utilitarianism and the Best Possible Humans,” forthcoming in *Ratio* 15 (in press, 2002).





trothal to take shape. So did the friends who did not invite Richard's father out hunting the afternoon of Richard's conception because one of them was offended by something that Richard's father said in an argument. So did the man who made the casual remark that started the argument. So did *his* parents. And so on. Of course we know this already. For want of a nail...

The decision to spare Angie is an event with massive causal ramifications. It is highly plausible that almost all killings and engenderings and refrainings from these have similarly massive causal ramifications. These actions ramify in massive ways most obviously because they are, let us say, 'identity-affecting'. These are actions that make a difference to the identities of future persons and these differences are apt to amplify exponentially down the generations. A very high proportion of identity-affecting actions are, it is enormously plausible, reliably subject to such massive causal ramification. These will include all engenderings—reproductively efficacious sex acts as well as causally more unorthodox engenderings—and at least a very high proportion of killings, including abortions. It is reflection on identity-affecting actions that, above all, brings out the depth of the Epistemic Argument. There are two reasons for this. The first is that common-sense reasoning demonstrates most clearly that identity-affecting actions are reliably subject to massive causal ramification. Such massive causal ramification is perhaps, as I will go on to suggest, pervasive elsewhere, but in the case of identity-affecting actions, there is no "perhaps" about it. The second reason is that killings and engenderings are among the most intrinsically morally significant things we do, the kinds of action at which a large amount of our most serious moral thinking and theorizing is directed, and, as the same common sense reasoning shows, these are the actions about whose overall consequences the agents are most apt to know, relatively speaking, as good as nothing. So the Epistemic Argument bites hardest in the case of just those actions respecting which we are most likely to want to put any ethical theory to work.

To drive the problem home, consider another, homelier, example. Suppose I decide to have a child and have a daughter, Andrea. When she grows up, she marries Duncan and has a son, George. So my identity-affecting action of having sex with Andrea's mother ramifies into the next generation. But this is just the start of it. For perhaps if Andrea did not come along, Duncan would marry Sandra. Duncan being snapped up by

Andrea, however, Sandra will marry Howard, who would otherwise have married Patricia, who will feel so let down by losing out on Howard she will marry nobody. So Andrea's existence has considerable identity-affecting effects far beyond the existence of George. Even if she stays childless, she will have such effects by simply taking Duncan off the matrimonial market. Even if she stays single, she will have similar knock-on effects from any dating in which she participates. Even if she stays celibate, she will almost certainly make identity-affecting differences through the parties she hosts, the introductions she perpetrates, and in a host of less scrutable, indirect ways. Ten generations hence it is highly likely that the consequences of my engendering of Andrea in terms of the identities of the people alive will be vast. Some of these people will do terrible—or wonderful—things, and it is my seemingly innocent act of procreation that brings all this about.

Identity-affecting actions include most acts of killing or engendering people. They will also include, as we have already seen in developing these examples, many apparently less momentous actions. For some actions that seem insignificant have massive causal ramifications by virtue of being *indirectly* identity-affecting in unpredictable ways. Tony and Geraldine would have had sex and conceived a child that wet Tuesday night if Gary had not called and invited them out for a drink. We can scarcely conclude that all our actions are indirectly identity-affecting in this way, but it is certain that many of them are. Equally certainly, it is often quite impossible to know which actions these are. Given this, we can rely on continuing massive causal ramification for the vast majority of identity-affecting actions even when the bloodlines of those immediately affected are—as with the celibate Andrea—far shorter-lived than Angie's. All our lives are certain to contain a great many indirectly identity-affecting actions even when we ourselves, like her, perform no *directly* identity-affecting actions.

Indeed, it is arguably a very real possibility that very many actions that seem very insignificant are subject to massive causal ramification. For some causal systems are known to be extremely sensitive to very small and localized variations or changes in their initial conditions.<sup>5</sup> Such sen-

5. Interest in such systems has intensified as a result of the rise of chaos theory in recent decades, extreme sensitivity to initial conditions being partly definitive of chaotic systems. [On these, see James Gleick, *Chaos: Making a New Science* (London: Heinemann, 1988); Peter Smith, *Explaining Chaos* (Cambridge: Cambridge University Press, 1998)].

sitivity will make still more trouble for consequentialism if it is true in even a small number of domains that have a significant influence on the human world. One such domain is perhaps the weather: differences in the weather make extremely widespread differences to the behavior of huge numbers of people. Such differences affect, for example, people's moods, the plans they make for any given day, and the way these plans evolve as the day goes on. For any significant difference in weather over a large populated area, some of these effects are certain to be identity-affecting. Now if small differences in initial conditions could make great differences here, these might include my cooking my dinner, visiting the gym, or smoking a cigarette. Another example of a kind of system widely believed to behave like this is furnished by financial markets, and once again these are influenced by countless, often quite intrinsically insignificant, human actions, and probably—directly or indirectly—by a very high percentage of intrinsically more significant ones. And the effect of market movements on human life is again enormous and certainly often identity-affecting. And, of course, we do not need scientific theory but just common sense to tell us that any action that is, however indirectly, identity-affecting is liable to massive causal ramification. So while the Epistemic Argument is *strengthened* if such systems are pervasively instantiated in nature, it does not *depend* on this.

We may conclude that massive and inscrutable causal ramification is plausibly the norm for identity-affecting actions. And many of the most morally significant actions are patently identity-affecting. Such ramification will also infect actions that feature in the causal ancestry of identity-affecting actions, including a very large number of actions that seem relatively insignificant. And to the extent that the human world affects and is influenced by causal processes that are highly sensitive to initial conditions, the range of actions that are subject to massive and inscrutable causal ramification may be very large indeed. The question of the extent of pervasive extreme sensitivity to initial conditions quite *generally* is one I gratefully leave to scientists and philosophers of science. We can afford to be—and perhaps we ought to be—cautious here. For in the case of identity-affecting actions—the actions that often interest moral philosophers the most—the fact of massive causal ramification is inescapable. And this by itself makes *serious* trouble for consequentialism.

The seriousness of the trouble we can easily make clear. Massive causal ramification is inescapably the norm for identity-affecting actions. By



the same reasoning, even more astronomical causal ramification must reliably attach to actions that are identity-affecting on a large scale: actions such as mass murder. Hitler, for example, was responsible for the deaths of millions of people. But just how terrible—by consequentialist standards—were Hitler's crimes? The full consequences of each death are plausibly no less vast and impenetrable than the consequences of the sparing of Angie. How many Malcolm the Truly Appallings might have been among the descendants of his victims? Not that it would help us to know this. For the causal ramifications of what Malcolm the Truly Appalling himself does are so astronomically great that its moral value is—by consequentialist standards—utterly inscrutable. So we have only the feeblest of grounds, from an objective consequentialist perspective, to suppose that the crimes of Hitler were wrong. Here, if anywhere, surely, there is a considered moral judgment at stake that is well-enough entrenched not to be up for grabs in the cut and thrust of reflective equilibrium, a judgment far enough from the periphery of the web of our moral beliefs to furnish a compelling *reductio* of any theory that might undermine it.

We can now see how Kagan seriously understates the objection. The problem is, he says, that “you can never be *absolutely* certain as to what all the consequences of your act will be” and that you can never “say *for sure* that any given act was right or wrong” (emphases mine). This suggests that the problem is merely an absence of *certainty* about consequences, an absence consistent with our having a *pretty good idea* what the consequences will be. And this is just what he claims in dismissing the Epistemic Argument:

Although we may lack crystal balls, we are not utterly in the dark as to what the effects of our actions are likely to be; we are able to make reasonable educated guesses.<sup>6</sup>

However, that does not begin to do justice to the worry. The worry is not that our certainty is imperfect, but that we do not have a clue about the overall consequences of many of our actions. Or rather—for let us be precise—a clue is precisely what we do have, but it is a clue of bewildering insignificance bordering on uselessness—like a detective's discovery

6. *Normative Ethics*, pp. 64–65.

of a fragment of evidence pointing inconclusively to the murderer's having been under seven feet tall. We may not be *strictly* without a clue, but we are *virtually* without a clue.

The trouble for consequentialism then is that the foreseeable consequences of an action are so often a drop in the ocean of its actual consequences. All Richard knows about his action is that it makes the difference between life and death for Angie. That is, of course, tremendously important for Angie. But this contribution to the good is only a tiny detail in the overall consequences of Richard's actions. So it gives only the *weakest* of reasons for him to think his action, by consequentialist standards, right or wrong.

### III. THE MOORE/SMART RESPONSE

The problems I have raised have often been noticed but are typically dismissed too quickly to do them justice. Thus, before we return to Kagan, we may note G. E. Moore's claim that:

The effects of any individual action seem, after a sufficient space of time, to be found only in trifling modifications spread over a very wide area, whereas its immediate effects consist in some prominent modification of a comparatively narrow area... It does in fact appear to be the case that, in most cases, whatever action we now adopt, 'it will all be the same a hundred years hence', so far as the existence at that time of anything greatly good or bad is concerned.<sup>7</sup>

And J. J. C. Smart gives a brief discussion of the problem of the inscrutability of remote consequences but suggests that they "approximate rapidly to zero like the furthest ripples on a pond after a stone has been dropped on it."<sup>8</sup>

Robert Frazier, arguing for a similar conclusion to my own here, cites these passages and urges that the claims they make should carry little weight, because they are *empirical* claims that are very hard to test.<sup>9</sup> In

7. G. E. Moore, *Principia Ethica* (Cambridge: Cambridge University Press, 1903), p. 153.

8. At p. 33 of his "An Outline of a System of Utilitarian Ethics," in J. J. C. Smart and Bernard Williams: *Utilitarianism: For and Against* (Cambridge: Cambridge University Press, 1973), pp. 1-74.

9. Frazier aims to establish similar conclusions to mine largely by appealing to the transitivity of causation and noting that, "Most acts are parts of very large causal chains and

fact, I think we can be certain they are false of a great many actions. For common sense reasoning carefully conducted (or as this is sometimes known, *philosophy*) shows that they are certainly false for the very large numbers of actions that are identity-affecting.<sup>10</sup> That they are true in other domains is at least very much open to doubt. That the consequences of small actions are apt to peter out in the way Moore and Smart suggest is precisely what we cannot expect in dynamic systems that exhibit great sensitivity to variations in initial conditions. So, even if we disregard identity-affecting actions, the “plausibility in the light of total science”<sup>11</sup> of the understanding of causal ramification Smart proposes is nowadays, to say the least, seriously in dispute. Of course we can hardly disregard identity-affecting actions because so many of the actions that most concern moral philosophers—killings, abortions, maimings, incarcerations, and quite a lot of sex—are among the actions most liable to be identity-affecting.

#### IV. THE CANCELLING-OUT RESPONSE

Just how small is your clue? We can throw light on this by reflecting on another natural move to make against the Epistemic Argument, a move Kagan makes as follows:

Of course it remains true that there will be a very small chance of some totally unforeseen disaster resulting from your act. But it seems equally true that there will be a correspondingly very small chance of your act resulting in something fantastically wonderful, although totally un-

---

these causal chains can go on indefinitely.” (“Act Utilitarianism and Decision Procedures,” p. 46). This is true enough but very inconclusive. For the mere fact that consequences can feature in long causal chains is perfectly consistent with the sort of petering-out ripple effect described by Moore and Smart, where the consequences get less significant over time.

10. Smart himself goes some way to recognize this in his discussion of Adam and Eve that follows the passage just quoted. But this is highly unconvincing. In the case of Adam and Eve, he is able to say that “the ‘ripples on the pond’ postulate is not needed” (p. 33) only because he helps himself to the assumption that all their descendants will be happy. And in the case of “two actual parents” he alleges that “remote consequences can be left out of account” (p. 34), although the reasoning he offers for this claim seems quite inadequate to the task.

11. Smart’s own (and one I very readily endorse) “guide to metaphysical truth”— see p. 6 of J. J. C. Smart and John Haldane, *Atheism and Theism* (Oxford: Blackwell, 1996).

foreseen. If there is indeed no reason to expect either, then the two possibilities will cancel each other out as we try to decide how to act.<sup>12</sup>

I do not think this line of defence succeeds. First of all, notice that again Kagan underrates the nastiness of the problem with his talk of “a *very small chance*” of a disaster. But with an action that is reliably apt for massive causal ramification, as are many of the most morally significant actions, there is nothing small about this chance. In the countless careers of the countless descendants of Angie, we may expect to find the seed of countless disasters. And, simply because the scale of causal ramification is so great, we can be *confident* of this.

The thought, however, is that the good and the bad in this whole huge causal process may be expected to cancel out. Or rather that good and bad in the unforeseen consequences will cancel out, leaving the foreseen consequences to make the difference.

It is worth being clear what is involved in thinking about this. Just suppose we did have crystal balls. Consider the calculation that Richard would then have to perform to know if his action is right by consequentialist standards. He would need to know a vast amount about the future history of the world in which he acts (or fails to act) as he does. He also needs to know a similarly vast amount about the closest possible world in which he does not act (or fail to act) as he does. That is bad enough, but it already involves the supposition that his act is chosen from just two alternatives. If there are more, there may be a great many more future histories he has to know. Let us suppose, however, there are just two alternatives. He must then sum the goodness contained in each of these possible futures and select the action that leads to the future that is maximally good.

This may be too simple. Perhaps he must consider not two determinate possible futures but two futures that branch endlessly, assign probabilities to every branching, and seek to maximize expected goodness. Or perhaps there is no such thing as a determinate way things would have been in the future if he had not acted (or failed to act) as he did. Perhaps such talk of massively complex historical counterfactuals is metaphysical nonsense on stilts and there is nothing here for even God to know. Because all these possibilities make matters even more intrac-

12. *Ibid.*, p. 65.

table for consequentialists, let us stick to the most tractable case, in which there are just two possible futures to consider.<sup>13</sup> Even in this case the scale of what we do not know that is, by consequentialist standards, morally relevant is astronomical.

By way of analogy, suppose Benjamin is a financier and his aim in some transaction is to maximize the aggregate financial gains over losses of all the people affected by it. And suppose you tell him—let us suppose you are God—that the transaction will have the consequences that billions of people are either better or worse off in monetary terms than they would be otherwise but you do not tell him more than this. What should he do? He hasn't got a clue. Now suppose you tell him a tiny bit more. You tell him that Mrs. Jennifer Lawson of 17 Cedar Row, Macclesfield, will be £100 better off than otherwise if he carries out the transaction. In terms of what he needs to know to maximize what he wants to maximize, he *still* hasn't got a clue. Or, in any case, the clue he has is surely as good as worthless.

In the future where Benjamin makes his transaction, let 'x' represent the aggregate financial gain to others in pounds, excluding only Mrs. Lawson's £100, and 'y' the aggregate loss. These figures represent differences to how much people have, compared with the alternative future where the transaction is not made. Benjamin has no idea whether x or y is the larger number. All he knows is that Mrs. Lawson gets her £100 and that x and y sum to something *huge*. Is he entitled, as Kagan suggests, to suppose that the possibilities he knows nothing about somehow cancel out in a way that would entitle him to take the consideration furnished by Mrs. Lawson's prospective enrichment as no less decisive a reason than were he certain of its exhausting the consequences of his action? And is Richard similarly entitled? Let me make four points in response to these questions.

1. Benjamin—or Richard—can only begin to acquire such entitlement by making some indifference assumption whereby possible gains and losses are taken to cancel each other out. He might do this on either empirical or a priori grounds. He might have some empirical grounds to suppose that where we find cases of massive causal ramification in the economic sphere, the gains and losses tend to cancel out in the long run.

13. On these complications, see, e.g., Marcus G. Singer: "Actual Consequence Utilitarianism," *Mind* 86 (1977): 66–77, and Alastair Norcross, "Good and Bad Actions" *Philosophical Review* 106 (1997): 1–34.

But we can surely safely disregard this possibility. Clearly no such empirical reassurance is available to Richard or those suitably circumstanced—and it is cases such as these on which we want the case of Benjamin to throw light. There are no cases of massive causal ramification of the kind to which identity-affecting actions are liable where we have empirical data adequate to any such conclusion, for the simple reason that, even if such ramification were easy to trace (in fact it is quite impossible), there are no such cases in which we have good grounds to suppose the ramification has yet come close to running its course. No consequentialist can seriously suggest in our present state of knowledge that we are anywhere near having the sort of information that would support the kind of generalizations we would need to support such a cancelling-out assumption.

What about a priori grounds? The whole business of a priori probabilistic indifference postulates is, of course, notoriously vexed. The fundamental problem is that we cannot just carve up the various possibilities and assign equal probability to each of them without some basis on which to favour some one of the many ways of carving them up over others. We need some nonarbitrary way to partition the possibilities before us and there very plausibly isn't one—or rather there isn't one unless we appeal to the sort of prior information about probabilities which, *ex hypothesi*, we do not have. So a consequentialist is on an extremely sticky philosophical wicket if he or she wants to put such a priori indifference postulates to work. Yet, in the absence of the sort of empirical information nobody has got, there is nowhere else for a consequentialist to go.

2. Suppose we are generous at this point. Suppose we assume (an *extremely* shaky assumption, as we have just seen) that some probabilistic indifference postulate is warranted. It has to be stressed that this does not give us a straightforward case of  $x$  and  $y$  cancelling out directly. This would happen only were we to assume:

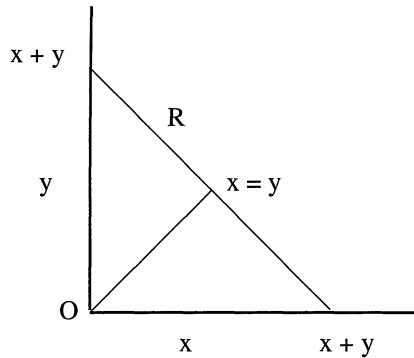
$$x = y$$

That is not a reasonable indifference postulate by anybody's standards. The most that even a friend of a priori indifference postulates is going to allow us is something like:

$$p(x > y) = p(y > x)$$

And this is consistent with its being hugely improbable that  $x = y$ , even

approximately. Suppose we represent the space of possibilities graphically as follows:



If, for simplicity, we hold  $x + y$  constant, the possible outcomes in terms of relative overall costs and benefits are represented by the line  $R$ . A natural indifference postulate will then be one that sees probability as distributed equally along  $R$ —so that the actual outcome is equally likely to be located in either of any two arbitrarily long equal intervals on  $R$ . And, of course, most of the length of  $R$  represents outcomes where, given that  $x + y$  is very large, the difference between them is far greater than Mrs. Lawson's windfall.

The good news for the consequentialist is that, given what we have assumed on his or her behalf:

$$p(x + 100 > y) > p(x + 100 < y).$$

The fact that the gain of which we have knowledge is positive and we know nothing else tips the balance of probability *slightly* in favour of optimism. And, if Benjamin can reason like this, so can Richard. If the expected consequences of his action are, on balance, good, that tips the balance of probabilities *slightly* in favour of a similar optimism.

The trouble is with “slightly.” We are here talking *very* slightly indeed. Consider further. Because Benjamin is a maximizer, he thinks he should make his transaction if gains exceed losses— $x + 100 > y$ —and he should not if  $x + 100 < y$ . (If they are equal it is a matter of indifference.) Do gains exceed losses? He has almost no clue. All he knows is that Mrs. Lawson

gets her £100. But, given that we know  $x$  and  $y$  sum to a very large number, a number astronomically greater than £100, that still leaves Benjamin looking pretty clueless. His information would be far more helpful if he could assume that  $x$  and  $y$  were equal. Or at least that they were nearly enough equal that Mrs. Lawson's £100 made a crucial, balance-tipping difference. That will be true if:

$$0 < (y - x) < 100$$

But, given that  $x + y$  is *very* large, this is *astronomically* improbable, even given the deeply questionable indifference assumption to which we have allowed the consequentialist to help himself. So it is astronomically improbable that Mrs. Lawson's profit makes a decisive difference to which of  $x$  or  $y$  is larger, which is the thing Benjamin, qua maximizer, needs to know. He thus remains, as we have urged hitherto, *virtually* without a clue. So perhaps the information he has about Mrs Lawson gives Benjamin a reason to favour making his transaction over not making it, but it is an *extremely* weak reason.

If Benjamin can reason like this, so can Richard. He should kill Angie if and only if the consequential gains of doing so exceed the losses. He does not know what these are. All he knows is *one item* on the loss side of the ledger—Angie gets killed. This is intrinsically a serious matter, of course, but, compared to the wider consequences of this action with its massive causal ramifications, it is a drop in the ocean, just as Mrs. Lawson's £100 is a drop in the ocean of the financial consequences of Benjamin's transaction. So once again the chance that the difference between gains and losses is so slight that the intrinsic evil of Angie's death is enough to bridge it is vanishingly small. That chance is not zero and the difference made by Angie's death can only be negative, so her death generates a reason against killing her. But it is an *extremely* weak reason.

3. Consequentialists might counter that all this is to miss the point. Consider all the unknown consequences of Benjamin's transaction *before* we factor in Mrs. Lawson. Given the indifference assumption we have allowed the consequentialist, the *expected value* of Benjamin's action, ignoring Mrs. Lawson, understood in the standard way in terms of the values of possible outcomes weighted by probability, is simply zero. Factor in the sure-thing gain to Mrs. Lawson and we get an expected value of exactly £100, just as if Mrs. Lawson's enrichment were the *only* consequence about which we had to worry. And this is just the cancelling out Kagan needs.



But this should not cause undue celebration in the consequentialist camp. The significance of the expected value of zero for an action of unknown astronomical consequences cannot at all be taken as the same as that of the zero expected value for an action with no expected consequences at all. Remember that we made our indifference postulate in the first place because we had *nothing* else to go on. It was no more than a *reflection* of our epistemic helplessness. Armed with this indifference postulate, we can then extract an expected value of zero for the unknown consequences. In effect we have reasoned like this: we have no idea what the expected value of Benjamin's action, before Mrs. Lawson is factored in, will be, so we just let it be zero. In this context, our expected value result, arrived at via an indifference postulate motivated in the way proposed, is just a rather fancy technical way of saying we have no clue what will happen. It is surely a sophistry to treat a zero expected value that reflects our knowledge that an act will lack significant consequences as parallel in significance to one that reflects our total ignorance of what such consequences (although we know they will be massive) will be.

Consider another analogy. Suppose you are in charge of plans for the Normandy landings for the Allied Forces in World War II. You are told by your staff that two very different plans of campaign have been worked out, plan A and plan B, of which you must choose one. You know, of course, that the plan you choose will have momentous consequences for the soldiers on the ground, for civilian populations, for the future course of the war. And you know too, let us suppose, that if you choose plan A, a certain dog, Spot, belonging to a harbourmaster in Cherbourg, will get his leg broken. If you choose plan B, you know Spot will be just fine. And let us suppose you know *nothing* else. What should you do? You have really no idea. And now suppose we suggest: Look. All these unknowns just cancel out in such a way that the expected value of B relative to A is otherwise zero but Spot tips the balance so you should pick B.

If this really were irremediably your epistemic position, perhaps you should. Perhaps we should concede that Spot gives you a reason to choose B, an extremely weak reason but nonetheless the only reason you have. Here I will concede you the *reason* part if you concede me the *extremely weak* part. Spot cannot weigh in here as a reason in at all the same straightforward way as he would if his broken leg were the *only* expected consequence of some action you contemplated. For in the plan A/plan B choice situation, the *relative* significance of Spot's injury in contrast to the other considerations you know there are (although you have no idea what they

are) is tiny. In the case in which there is no other consequence, this consideration is the whole story.

Mutatis mutandis, the same points apply to Benjamin and Richard. Of course killing Angie is, in itself, a rather bigger moral deal than injuring Spot or giving Mrs. Lawson a small windfall. But if it is swamped by other consequences to a proportionate degree, the significance of the reason it offers Richard is proportionately diminished.

4. There might seem to be a final move a consequentialist might make here. I have argued that the problem of invisible consequences is such that, if we accept consequentialism, we have only a *very* weak reason not to, say, kill an innocent person. However, might not a consequentialist grant that this reason is very weak but insist that a very weak reason is good enough to tip the balance? For there is *nothing* on the other side. Or if there is (I am proposing to kill Alphonse for his money and I will be made better off), it too is subject to a similar swamping in the ocean of invisible consequences.

After all, it could be argued, from a consequentialist perspective, when we say these are weak reasons, we cannot mean that they are weak *relative to other reasons* we sometimes have. For *all* our (practical) reasons are subject to the same swamping by an invisible hinterland of remote effects. So the *relative* strength of our reasons remains the same. We have simply scaled them all down. And when we scale *everything* down by proportional amounts, the shape of the landscape does not change.

When this move is made we can continue, it could be argued, to rely on aiming at maximizing benefits from proximate and scrutable consequences and motivate this by reference to the desirability of maximizing benefits from overall consequences. The only thing we would have to cope with would be a certain dizzying humility about our prospects for achieving this. We could comfort ourselves nonetheless that our efforts, however insignificant, were somewhat better than nothing.

This is an interesting move and perhaps a consequentialist might think the plausibility of his or her theory made the dizzying humility a price worth paying. But there are two problems—even if again we overlook the thoroughly dubious indifference assumptions on which the whole strategy must rest. The first, which I will defer to the final section of this essay, is that this response brings new life to the Integrity Objection to consequentialism. The second is this. In running through this reasoning, I suggested that “*all* our (practical) reasons are subject to the *same*

swamping by an invisible hinterland of remote effects.” But that is *not* the conclusion the argument of this essay justifies. What I have said is that we may be *certain* such swamping occurs with *most* identity-affecting actions and that many *more* actions *may* be indirectly identity-affecting than are obviously so. Taken to an extreme, my argument could be extended to suggest that, *if* the world we inhabit is sufficiently pervaded by causal processes sufficiently sensitive to initial conditions, *all* our actions might conceivably have this character. But it is at least debatable whether we are at present in a position to be confident about this latter conjecture. Certainly this does not follow from anything that is said here. It is significant here that it is many of the most morally significant actions that are most apt to be identity-affecting and so most straightforwardly apt to be subject to the swamping in question. Our confidence that such swamping occurs is plausibly not distributed equally across actions, and it is precisely with respect to many of the most morally significant actions that our confidence may often be greatest. It would be awkward if our reasons in these cases are apt to be scaled down more than our reasons with respect to other less significant actions. This final line of defence assumes the scaling down called for to be *uniform* but nothing in the epistemic considerations I have rehearsed supports this assumption.

#### V. DISENGAGEMENT: CRITERIA OF RIGHTNESS AND DECISION-PROCEDURES

Another natural move, also made by Kagan,<sup>14</sup> is to distinguish *subjective* and *objective* senses of rightness. Subjective rightness, for consequentialists, relates to a conscientious agent’s rational expectations about the overall consequences of his or her actions; objective rightness to the actual consequences. A consequentialist may then claim, very plausibly, that it is subjective rightness that is relevant to moral appraisal, and that properly informs an account of moral decision-making. On the more straightforward versions of subjective consequentialism, the thought is then that what we ought to do is always the subjectively right act, where this is the act with the best foreseeable consequences from

14. *Normative Ethics*, pp. 65–66. See also, e.g., Peter Railton, “Alienation, Consequentialism, and the Demands of Morality,” *Philosophy & Public Affairs* 13 (1984): 134–71; Frank Jackson, “Decision-Theoretic Consequentialism and the Nearest and Dearest Objection,” *Ethics* 101 (1991): 461–82.

the point of view of a rational and conscientious agent. But this does not help. For the problem is explaining, from a consequentialist point of view, *why* we should do the subjectively right act. For very often, and especially in the most morally significant cases, we are virtually without a clue how objectively right an action will be. At least in a great many morally significant cases, most of the consequences of an action will not be expected consequences. If the thought is that we should do the subjectively right thing as a means to approximating objective rightness and promoting objective goodness, that is just the thought the Epistemic Argument calls in question. It does not, by itself, constitute any kind of reply to the Epistemic Argument.

A more sophisticated and revisionist variant of this response is found in another common move whereby consequentialism is defended as offering a *criterion of rightness* but not a *decision procedure*. Consequentialism, this defence would go, is a *constitutive* story about the objective nature of goodness and rightness, objectively speaking, but a different story could be told by consequentialists about how we should deliberate, what our motives should be, and how we should appraise our own and each others' actions. This is revisionist in that it amounts to maintaining consequentialism as an account of objective rightness while advocating a non-consequentialist story about subjective rightness.<sup>15</sup> It would be odd, however, if these two stories were unrelated. If consequentialism is to be a theory of any real normative interest, it must at least furnish us with a regulative ideal to guide our choices either of actions or decision procedures; it must offer such choices a consequentialist rationale.<sup>16</sup> Either the consequentialist thinks the constitutive, consequentialist story has a role to play in motivating and making sense of the story about deliberation and appraisal or he or she does not. If the consequentialist does, the epistemic difficulties raised here have still to be faced and we have not yet seen how to face them. If he or she does not, the option perhaps remains open of sticking to objective consequentialism. Someone might say: Although we have no idea what actions of ours would have the best consequences, it remains true that

15. For this kind of strategy, see, e.g., Railton, "Alienation, Consequentialism, and the Demands of Morality," and Roger Crisp, "Utilitarianism and the Life of Virtue," *Philosophical Quarterly* 42 (1992): 139–60.

16. Here it may be worth reminding the reader that my concern is with *pure* consequentialism.

those actions are objectively the best actions we might take. So if we knew which actions of ours had the best consequences, we should govern our conduct in the light of that. Sadly we do not know this, so the truth of consequentialism is of no practical significance. Nonetheless, speaking from the disengaged perspective of pure philosophical inquiry, we can plausibly claim that it is true. Such *disengaged consequentialism*, as we might call it, might be true in spite of the Epistemic Argument. But if these arguments are sound, it would have, although true, nothing whatever to do with our practical thought in its application to the actual world.<sup>17</sup>

Disengaged consequentialism is naturally formulated as a *conditional* claim:

*The DC Conditional:* If massive causal ramification were not a fact of human life, or if it were but we had supernatural powers of foresight, then consequentialism would be the best ethical theory for human beings.

Claims such as the DC Conditional raise large questions about how far ethical theorizing should aspire to abstract away from the facts of human life and generalize across possible permutations of these facts. For, of course, there are all manner of ways in which these facts might have been very different. If we were asexual, solitary beings whose newborn offspring took care of themselves; if the circumstances of justice did not obtain; if we were immortal, our practical thinking would undoubtedly take very different forms. At issue here perhaps are two contrasting conceptions of ethical theory. The first is extremely general. It would seek, ideally, to vindicate a set of highly general claims about reason and value that abstract away from the specific and contingent facts of human life, claims from which more specific prescriptions about what we should do under all these various possible permutations of these facts could be derived, in principle, as theorems. The other approach, by con-

17. Compare Frank Jackson, "It is fine for a theory in physics to tell us about its central notions in a way which leaves it obscure how it can move from these notions to action, for that passage can be left to something which is not physics: but the passage to action is the very business of ethics." ("Decision-Theoretic Consequentialism and the Nearest and Dearest Objection", p. 467)

trast, starts from the specifics and is highly sceptical about the possibility, and perhaps the intelligibility, of the sort of axiological or normative theorizing that seeks to abstract away from them.

This is not the place to decide between these large conceptions. An ideal theory of the sort envisaged by the first of these conceptions might well be consequentialist, for all I have said here to the contrary, just insofar as it includes something like the DC Conditional among its theorems. In such an approach, the DC Conditional might make perfect sense and might even be claimed to be true, but, because its antecedent would be false, it would nevertheless not be very interesting. For massive causal ramification, and the inadequacy of our powers of foresight to anticipate where it will lead, is a fact of human life.

#### VI. TU QUOQUE

Kagan has one other point to make in defence of consequentialism here. It is to claim companions in guilt. The Epistemic Argument, if sound,

threatens not only consequentialism, but indeed all plausible normative theories. For if it is in fact impossible to get a grip on the consequences of an act, then this problem will be inherited by all theories that give this factor any weight at all and this will be virtually all theories. For... all plausible theories agree that goodness of consequences is at least *one* factor relevant to the moral status of acts.<sup>18</sup>

This is eminently questionable. Earlier in his book Kagan wrote:

It is important to understand that in saying that the moral status of an act is determined (at least in part) by its results, this is meant to include all its results. It is not only the immediate, or short term, results that matter: long terms results, side effects, indirect consequences—all these matter as well and they count just as much as short term or immediate consequences.<sup>19</sup>

18. *Normative Ethics*, p. 64.

19. *Ibid.*, p. 26.

It is important to notice that if we understand the claim that, “the moral status of an act is determined (at least in part) by its results” in this inclusively comprehensive way, it is not at all obvious that any plausible theory need agree to any such thing. Certainly many plausible moral theories quite explicitly do not give all the consequences of an action equal weight. Thus, we are very accustomed to moral theories that give weight to the distinction between intended and merely expected consequences and do so because they take the structure of the agent’s intentions to be crucially relevant to the moral appraisal of his or her actions.<sup>20</sup> And it might make good sense, under certain approaches—but not consequentialism—to give, in similar ways, moral weight to a distinction between expected consequences and unforeseen consequences.

A typical morally significant action will have a host of consequences that are neither foreseen nor foreseeable. Distant causal ramifications such as the atrocities performed by Angie’s distant descendants are of this sort. Let us call these the *invisible* consequences of an action as opposed to its *visible* consequences. The distinction is vague. Some consequences might be visible only to a heroically conscientious agent; others might be visible to the agent but only as more-or-less salient risks. There are some consequences we can confidently expect and others that are merely envisaged as more or less likely. Rather than bogging down in this complication, let us count as invisible only consequences that are, in practice, not visible to an ideally conscientious agent, where this is nonetheless a *human* agent and not some Laplacean fantasy; an agent we will suppose to have weighted all envisaged consequences by probability in some ideally reasonable but humanly manageable way. It remains likely that most of an action’s consequences will be invisible, especially for the more intuitively morally significant actions.

Here is what I suggest nonconsequentialists may plausibly say an agent should do about the invisible consequences of his or her actions. Rather than act on the profoundly shaky assumption that they cancel each other out, the agent should ordinarily simply not regard them as of moral concern. The agent should not think of himself or herself as maximizing the goodness of the consequences of his or her action by maximizing that of

20. Kagan, of course, does *not* think the Doctrine of Double Effect is plausible (see his *The Limits of Morality* (New York: Oxford University Press, 1989), ch. 4). But I take it his appeal to companions in guilt is intended to carry *ad hominem* force in a way that licenses reading “plausible” as “widely held to be plausible” rather than “plausible to Kagan.”

the expected consequences and hoping these manage to be somehow representative. Rather insofar as the agent's concern is with consequences at all, it is with *visible* consequences that he or she should be, even indirectly, concerned.

It is not an objection to this claim as I have formulated it that we often plausibly have a duty not to be negligent in exploring and anticipating the consequences of our action. For I have defined "visible" to mean, not *de facto* foreseen but foreseeable to an ideally conscientious agent. It is certainly plausible that we have duties to be conscientious in this way but that is nothing to do with invisible consequences as defined.

If you find consequentialism attractive, you will think this a surprising claim. For a consequentialist the point of maximizing the goodness in the visible consequences of our actions has to be as a means to maximizing the goodness in their overall consequences. But if the foregoing reasoning is sound, we have only the most feeble of grounds to suppose that means—or any feasible other—is a remotely reliable means to this end. It may, however, make far better sense for ethical theories for which the focus is on the character of agents and the qualities of their wills, for theories that are broadly Kantian or Aristotelian in spirit. Such theories would move us away from consequentialism to some radically less impersonal understanding of how best to live whereby we should be morally engaged not by the quite futile project of promoting good long-term results but by more local projects and concerns whereby, recognizing the fact of our epistemic limitations, we seek nonetheless to live virtuously, with dignity and mutual respect. What matters for theories such as these is the virtues of character our actions manifest and/or the forms of respect we show for others in acting; and perhaps in particular (many such theories are not ashamed to say) for certain others, those closest to us in a number of senses of "close," those most concerned in the intentions and warranted expectations on which we act.

What is common ground to all plausible ethical theories is the moral significance of visible consequences. When we can foresee harm to others in the outcome of our actions, we owe them the respect of taking this properly into account. And we owe it to others also to be adequately conscientious in foreseeing such harm. Of course, the invisible consequences of action very plausibly matter too, but there is no clear reason to suppose this mattering to be a matter of moral significance any more than the consequences, visible or otherwise, of earthquakes or meteor im-



pacts (although they may certainly matter enormously) need be matters of, in particular, moral concern. There is nothing particularly implausible here. It is simply to say, for example, that the crimes of Hitler, although they were a terrible thing, are not something we can sensibly raise in discussion of the moral failings or excellences of Richard's conduct. This could be justified on a number of theoretical perspectives. Thus a Scanlonian contractualist<sup>21</sup> might plausibly urge that, on principles that we could not reasonably reject, Richard owes it to Angie not to kill her but does not plausibly owe it to the Poles, Russians, Jews, and others of a distant generation not to perform actions with massive causal ramifications that might result in harm to them. He would have the first obligation because recognizing it is a quite fundamental way in which he shows respect for her; but he would lack the second because it is manifestly unreasonable not to agree to principles that limit the sphere of our responsibility to those harms and goods that are visible to us.

A consequentialist might seek to agree to this limiting of moral focus by again taking the line that, in matters of assessing the moral or rational merits of actions, of assigning praise and blame, we should concern ourselves with subjective rightness, for subjective rightness is precisely concerned only with visible consequences. The trouble is that, as I have noted, a consequentialist must understand this concern as motivated by the belief that maximizing value with respect to visible consequences is a reliable means to maximizing value with respect to overall consequences. And this belief does not appear at all secure.<sup>22</sup> Given this, we might prefer a theory that tells a different story about what the point is of our concern with visible consequences. And such a story would precisely *not* be consequentialist.

Perhaps the best normative theory is nonetheless just such a version

21. See T. M. Scanlon, *What We Owe to Each Other* (Cambridge, MA.: Harvard University Press, 1998).

22. Thus, Alastair Norcross wrote, "there may be a temptation to abandon consequentialist methods of reasoning on the grounds that what one is able to bring about, or able to be certain to bring about, is somehow dwarfed by the scale of what else is beyond one's control, or beyond one's epistemic reach. In each case, the cure is to focus on what is within one's control." (p. 256 of "Consequentialism and the Unforeseeable Future"). What I think Norcross does not here appreciate is just how ill-placed we are to offer a consequentialist—as opposed to, say, a virtue-ethical or contractualist—*motivation* for taking this cure. (Perhaps there is an *impure* consequentialist motivation but, as I noted at the outset, I'm not here concerning myself with that possibility.)

of consequentialism restricted to visible consequences and with some non-consequentialist rationale. In such a case all I have said here might prove strictly consistent with a view that keeps everything, for all practical ethical purposes, just where consequentialists want it. Perhaps. But if such a consequentialist-for-all practical-purposes view is strictly consistent with what I have argued here, its appeal is, very plausibly, greatly weakened. One promising starting point for understanding where an appreciation of the Epistemic Argument might leave us might be in terms of an influential distinction introduced by Philip Pettit as a shift in emphasis from the *promotion* to the *honouring* of values.<sup>23</sup> In standard consequentialist accounts, our ethical decision procedures legitimately demand that we honour values (in perhaps agent-relative ways) only where that is legitimately to be supposed an effective means of promoting them. The present arguments undermine this understanding of the matter. Rather than saying, “Where we are legitimately concerned to honour values that is, most fundamentally, because we are required to promote them,” we may instead say, “Where we are rightly concerned (as we clearly often are) within the domain of visible consequences to promote values, that is, most fundamentally, one of the ways in which we seek to honour them.” Once that thought is on the table, the supposition that there might be in the last (or second-last) analysis only one such way, that the promotion of value understood in some strictly agent-neutral way could be at some theoretical level the whole story about how we should engage with our ethical concerns, plausibly loses much of what might otherwise have seemed its considerable intuitive appeal.<sup>24</sup>

23. See Pettit’s “Consequentialism” in *A Companion to Ethics*, ed. Peter Singer (Oxford: Blackwell, 1991), pp. 230–40.

24. It seems worthwhile to air briefly a further way in which the present argument might affect our ethical theorizing that it would certainly be interesting to develop further elsewhere. I have been arguing that an appreciation of the extent to which our actions, and in particular some of our most morally significant actions, are subject to massive and inscrutable causal ramification may change the way we see consequentialism, making it look far less of a starter than initially it might. There is another view that such an appreciation might lead us to reevaluate and that is *ethical absolutism*—I mean the normative view that certain moral principles quite simply admit of no exceptions. The terrorists will bomb a city unless you torture and murder a child. Pedro will let the other Indians go if Jim shoots one himself. (The example is, of course, from pp. 98–99 of Williams’s “A Critique of Utilitarianism” in Smart and Williams: *Utilitarianism: For and Against*, pp. 75–150). “Tough,” says the absolutist, “There are things you just don’t do.” Absolutism is perhaps not nowadays a widely held view, especially among philosophers without religious faith. In the absence of a loving God in whom we might simply place our trust, it can easily seem just crazy. The bad

## VII. A THOUGHT EXPERIMENT: INTEGRITY REVISITED

A thought experiment might bring the concerns of earlier sections nicely together. Let us recall the line of defence envisaged at the end of section IV where the consequentialist suggests that a uniform scaling down of our reasons might leave everything unaffected and his or her theory in the clear. I argued there that such uniformity could not be assumed. But let us bracket this reply and suppose the scaling down *is* uniform. Another counter could still be made that is, it seems to me, no less compelling.

Suppose, first of all, you must spend your whole life alone in a single room seated at a console. This console has two buttons on it, one red and one green. God has explained to you that whenever you press either button, you cause an event with massive causal ramifications—that is, you initiate a massive chain of effects, many of them extremely significant in human terms, that cascade through history inscrutably. You know practically nothing else about what will happen except that, like Benjamin and Richard, you are given a tiny clue. You know for sure that, when you press the green button, of the countless billions of events that you initiate, at least one is a jolly good thing. And you know for sure that, when you press the red button, of the countless events that you initiate, at least one is a jolly bad thing. As I have argued, this gives you at best an extremely weak reason to press the green button and an extremely weak reason not to press the red button. That said, it is perhaps slightly more rational to press the green button than to press the red button or to do nothing at all and slightly less rational to press red than to press green or to do nothing at all. So, because there is nothing else for you to do with your time, you might as well just sit there and press the green button. You should be immensely humble about whether any of this does any

---

consequences of not killing the child or the Indian seem so clearly to outweigh the good that even those who are not consequentialists find it hard to resist letting the principle slip where such cases are in question. Only now, in the light of what has been argued above, absolutism might come to seem far less of a nonstarter than it may initially have done. For we can appreciate that we are simply not in a position to say, with any confidence at all, that we know such things about the overall consequences of the alternatives in such (identity-affecting) cases. This is not to say that absolutism is right. For I have said nothing that throws it in doubt that *all visible* consequences should carry at least *some* weight or that helps us much in the difficult business of reaching agreement on what weights those should be. But I may have given some grounds to abandon the belief that absolutism is *obviously* mistaken in the way it can very easily seem to be.

good overall, but it is still perhaps, by a slight margin, the best course of action available to you from an unappetising menu.

So let us improve the menu a bit. Suppose there is something else you can do. Leading off the room in which you are seated is another larger room to which you have access. And in this room there is, let us say, a Nintendo machine along with other materials for amusing solitary diversions suited to your tastes. Instead of sitting there pressing buttons and creating effects that you have almost no reason to believe will be better overall than those of doing nothing, you could simply adjourn to this neighbouring room and have a good time. And you can, we may suppose, be reasonably sure of having a fairly consistent good time (you like playing Nintendo and are not easily bored). To insist that you should persevere with your button pressing is surely just daft. Now, after all, you still get to do things that have one good consequence—your own pleasure—and now you get to have a life, too. Well, OK, relatively speaking you get to have a life—where a life is understood to be something more than what you had before: the mere pulling around of causal levers.<sup>25</sup>

Suppose at this point God tells you a little more. Suppose he tells you that the single good thing you can be sure of causing at any given time is significantly—although by no means massively—better than the good of the pleasure you could be having in the larger room next door. And suppose he also tells you that all the things you do in the larger room by way of having fun have massive causal ramifications of their own. But they have them in the big wide world outside—a world that is invisible to you—and only there. You know, of course, that in the causal chains you thus initiate there is still at least one good thing, your own pleasure. Only you know—God kindly having told you—that this is a lesser good than the single sure-thing good you can rely on producing by pressing the green button. In these circumstances a perfectly impartial consequentialist would stay put in the smaller room pressing the green button. But if there were ever a case of the supposed demands of morality being set absurdly high, it would be implicated in any insistence that that is what you should do.

Let us now offer you a little more of a life. Adjoining the larger room full of toys is another small room with another occupant similarly faced

25. I am here consciously echoing Bernard Williams, *Ethics and the Limits of Philosophy* (London: Fontana/Collins, 1985), p. 77.

with green and red buttons. Now you have the option, the pair of you, both to abandon your button pressing and adjourn together to the large room in pursuit of more interesting and personally rewarding pursuits: you can now have fun together.<sup>26</sup> And suppose the rules stay the same. God keeps you informed that the sure-thing but drop-in-the-ocean goods you get from button pressing are reliably a bit greater than the personal good you get from hanging out together in the large room. And you know as before that your joint enjoyments are reliably apt massively to ramify causally in the invisible world outside and only there. You have now, of course, all the more reason to stick to the life of the larger room. For the life you can lead there has been ethically enriched. This remains from a consequentialist perspective a slightly suboptimal thing to do, but you now plausibly have even less of a good reason to see that consideration as at all decisive.

Now let us give the larger room, progressively, many more such smaller occupied siderooms. And let us make it, progressively, much larger. And let us suppose that when you stay in the larger room you develop all manner of economic and social needs and interests that you are able to satisfy very much the better by engaging in all manner of cooperative and social activities with the people with whom you share it. But let us still keep the rules the same as in the previous paragraph—the visibly good things you can accomplish in the larger room are reliably somewhat less good than the sure-thing, drop-in-the-ocean goods you can produce externally by button-pressing. You can now have a lot more of a life. Indeed, you can now have something very close to a full human life. Or you can sit in a sideroom and press the green button. Because the rules are the same, the latter would be, by a tiny margin, the optimal solution. It would also very plausibly be crazy.

As you live your life in the large room with your companions, you may find that you need all manner of principles to govern your lives together and you may deliberate together to arrive at principles that are fair and reasonable. And as your lives together progress, you may find you come to share a conception of what sort of motivations and character the most flourishing and valuable members of this community are apt to have and thus develop a conception of a virtuous human being, a conception informed by and informing of the moral principles that you have come to

26. Stop that sniggering in the back row...

share. This conception and these principles might very plausibly require that you act with a due concern for the needs and interests of those with whom you share your life in this community. Indeed they might plausibly require that you act with a due concern for all those affected by the visible consequences of your actions. But if they also insist that you concern yourself with those affected by the invisible consequences—which we are still thinking of as taking shape in, to you, an invisible world outside—and insist you do this with such determined impartiality that it becomes rational for you to return to your button pressing, then they are plausibly quite crazy principles, principles that are not informed by any remotely adequate conception of the ethical shape of human lives lived together.

It only remains to lose the walls and relocate the ramifying invisible consequences of your actions within the social world in which you live, albeit to a large extent far in the future. With this last modification, we leave the realm of fantasy for the way things are. But this plausibly changes nothing of sufficient consequence to compel you to pay much greater attention to remote and invisible consequences. You and your fellow members of this newly wall-less community have constructed moral principles and ideals of character, the point of which is to inform and shape your lives together in mutually beneficial and harmonious ways. And this is a more plausible conception of what you were—and should have been—about when you constructed these principles and ideals than one that makes it their point maximally, impartially, and timelessly to promote the interests of all sentient beings. If that was what mattered to you, you would have stayed in your little room pressing the green button and would never have bothered to build a life in society with others. That cannot be what morality is about because it is so plausibly not what living a human life is about. All of which, perhaps, is merely to reemphasize what Bernard Williams taught us some time ago<sup>27</sup>: that if you expand the ethical circle to its outermost possible limits and flatten the contours within in, you will find you disappear in it. The Epistemic Argument and the Integrity Objection are close kin.

27. In numerous writings but especially “A Critique of Utilitarianism.”