Utilitas (2019), 1–17
doi:10.1017/S095382081900030X

CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

Subjective Consequentialism and the Unforeseeable

Christopher Jay*

University of York
*Corresponding author. Email: christopher.jay@york.ac.uk

Abstract

As is already well known, subjective consequentialists face a challenge which arises from the fact that many (perhaps even most) of the consequences of an action are *unforeseeable*: this fact makes trouble for the assignment of expected values. Recently there has been some discussion of the role of 'indifference' principles in addressing this challenge. In this article, I argue that adopting a principle of indifference to unforeseeable consequences will not work – not because of familiar worries about the rationality of such indifference principles, but because subjective consequentialist theories which adopt such principles end up entailing either deontic indeterminacy or arbitrary deontic variance. This is because of another well-known fact: that possibilities do not 'agglomerate'.

> Don't say 'face facts' to me . . . . Everybody keeps saying it just now; but the fact is, it's impossible to face facts. They're like the walls of a room, all around you. If you face one wall, you must have your back to the other three.
>
> (E. M. Forster)[1]

This article starts from the 'cluelessness objection' to consequentialism. I consider a recent proposal about how subjective consequentialists might avoid the objection, namely by endorsing a principle of indifference to unforeseeable consequences. I argue that even if this proposal works with respect to cases involving some unforeseeable consequences, in at least some cases there are consequences to be dealt with which are foreseeable but are still problematic in a particular way. I explain that *ultimately unforeseeable* consequences (a technical notion) are those which are individually foreseeable, but not *jointly* foreseeable in combination with other consequences of the same action. And I argue that with respect to cases involving foreseeable but ultimately unforeseeable consequences, subjective consequentialism modified by an indifference principle to deal with the standard cluelessness objection entails either deontic indeterminacy or arbitrary deontic variance – neither of which are plausible.

[1]Quoted in P. N. Furbank, *E. M. Forster: A Life*, vol. 2 (London, 1977), p. 2. I came across the comment as the epigram of Stephen Neale, *Facing Facts* (Oxford, 2001).

Section I introduces consequentialism and the debate about how consequentialist theories should or could deal with our epistemic limitations. Then Section II explains the way in which subjective consequentialists might solve what I call the standard cluelessness objection by invoking a so-called principle of indifference. Section III discusses foreseeability, and the importance of the fact that foreseeability does not 'agglomerate'. Section IV employs the point about non-agglomeration introduced in the previous section to argue that subjective consequentialism, if it is to avoid the standard cluelessness objection, entails at least one of two problematic conclusions: either that the deontic status of some actions is indeterminate; or that the deontic status of some actions varies arbitrarily, which is to say that whether some actions are permissible depends upon facts of no moral significance, which vary between people or modally. Section V explains that the problem I discuss in the previous section is not simply a version of a more basic and easy-to-solve problem. And Section VI explains why treating deliberation as itself subject to consequentialist evaluation will not solve the problem.

As it happens, *I* do not think that the epistemic objections to objective consequentialism which, as I describe in Section I, sometimes motivate subjective consequentialism are at all persuasive. And I would not be inclined to accept the various claims made in the course of the argument running through Sections IV and V about the importance of a moral theory's ability to provide action guiding answers to questions about the deontic status of particular actions. I agree with those objective consequentialists, and others, who argue that consequentialism should be seen and judged as a theory about what grounds the deontic status of actions, rather than as a decision procedure which tells us which particular actions to perform on the basis of calculations: it is worth knowing the *criterion* of permissibility and related notions, even if we are unable to use that knowledge to determine what in particular is permissible or impermissible.[2] But many philosophers (and others, if only implicitly) are strongly committed to the idea that moral theory should be action guiding, in ways that it only could be if it provided knowledge of the deontic status of particular actions, and if they are right then the argument I develop in this article, extending the already familiar cluelessness objection, is a very serious one for consequentialism. So, since I might be wrong about the importance of the action guiding constraint, it seems worth presenting this argument.

I. From objective consequentialism to subjective consequentialism

Consequentialists believe that the deontic status of actions – facts about their permissibility or their being required – depend upon the value those actions promote.[3] Even if they believe that actions have intrinsic value which partially determines their deontic status, they hold that the value in question is not only the intrinsic value of

[2]For different forms of 'indirect' consequentialism which explain the difference between a moral theory's principles and the principles which ought, according to that moral theory, to be appealed to by agents deliberating about what to do see, e.g., R. M. Hare, *Moral Thinking: Its Levels, Method and Point* (Oxford, 1981) and Peter Railton, 'Alienation, Consequentialism and the Demands of Morality', *Philosophy & Public Affairs* 13 (1984), pp. 134–71.

[3]In this article, an action's 'deontic status' is always its *moral* permissibility, or its being *morally* forbidden or required. For the characterization of consequentialism in terms of 'promoting' value, see Philip Pettit, 'The Consequentialist Perspective', in Marcia Baron, Philip Pettit and Michael Slote, *Three Methods of Ethics: A Debate* (Oxford, 1997). In the remainder of this section, I present a very sketchy summary of the varieties of consequentialism which is not supposed to be comprehensive but picks out some salient points for what follows.

the action itself, but the value of the action's consequences too.[4] According to (act) consequentialism, actions are ranked evaluatively, according to the amount of value or good they promote; and the deontic status of actions is determined by a principle which says, roughly, that the only permissible actions are those which are at least as highly ranked as any other action(s) – which means that it is impermissible to perform any action(s) other than the action(s) which promote the most good (and it is assumed always to be permissible to promote the most good).[5]

The brief characterization above best captures so-called *objective* consequentialist theories, according to which it is the value of the *actual consequences* (and perhaps of the actual nature of the action) which determines the evaluative ranking of actions, and therefore determines the deontic status of those actions. As is well known – and obvious – objective consequentialism has a feature which many take to be the basis for a decisive objection to it: the deontic status of an action, according to objective consequentialist theories, might be unknowable to us, for that status depends upon facts which are themselves unknown and in many cases unforeseeable by us. Since I cannot know what all the consequences of an action (my own or someone else's) will be, and since all its consequences might contribute to determining the evaluative ranking of the action, I cannot know how to evaluatively rank that action correctly; and since its deontic status depends upon its evaluative ranking, I cannot know its deontic status either.

A variety of objections are raised on the basis of this seemingly unavoidable feature of objective consequentialism. Frank Jackson argued that since we often lack sufficient evidence to know what the best thing to do is, even with respect to short-term and in principle foreseeable consequences, such as the more or less immediate effects of a particular drug, objective consequentialism fails to be action guiding, and any plausible moral theory must be action guiding. Frances Howard-Snyder has argued that objective consequentialism violates the principle that *ought* implies *can* in virtue of our limited knowledge. And James Lenman has pointed out the sheer scale of the epistemic mountain we would need to climb to know what is permissible or required according to objective consequentialism: we don't simply lack evidence about the effects of particular actions; we lack more or less any conception at all of the complete variety of facts we would *need* to know. The consequences of actions are not only far reaching; they are also bewilderingly complex. Some actions unpredictably change the course of history in respect of which agents exist, with all the complex ramifications that involves. And the problem isn't just that *our own* actions are mysterious to us in respect of their complete consequences (so the problem isn't just that objective consequentialism fails to be action guiding); it is that *every* action, past present and future, has this feature, meaning that we cannot even make properly informed moral judgements about the

[4]See, e.g., Derek Parfit, *On What Matters*, vol. 3 (Oxford, 2017), §57 on 'act-including act consequentialism' for a discussion of types of consequentialism which take account of the intrinsic value of actions, as well as their consequences.

[5]This is true of both traditional 'single dimension' consequentialism, such as utilitarianism, and (with some complications which aren't relevant for this article) more sophisticated 'multi dimension' consequentialism, such as the view developed in Martin Peterson, *The Dimensions of Consequentialism: Ethics, Equality and Risk* (Cambridge, 2013). In this article I am only concerned with *act* consequentialism. Rule consequentialism (see, e.g., Brad Hooker, *Ideal Code, Real World: A Rule-Consequentialist Theory of Morality* (Oxford, 2000)) is a very different type of theory, but it is probable, I think, that arguments similar to the ones I present here apply to the appropriate forms of rule consequentialism too.

deontic status of historical actions – after all, their consequences are (at least for all we know) still unfolding, and will continue to unfold long after our deaths.[6]

In response to these challenges, some (such as Jackson) who are disposed to endorse consequentialism abandon objective consequentialism in favour of *subjective* consequentialism.[7] Like objective consequentialism, subjective consequentialism proposes that the deontic status of an action is determined by its place in the evaluative ordering of available actions. But according to subjective consequentialism, that evaluative ordering is on the basis of the *expected* value of each available action – where *expected value* is a technical notion, being the sum of the products of the subjective probability and the value of each possible outcome of an action: the expected value of *A*'s φ-ing is calculated by adding together the values achieved by taking each possible outcome of *A*'s φ-ing and multiplying the value of that outcome by its subjective probability.

This looks like an excellent way to accommodate – as objective consequentialism could not – the uncertainty about outcomes which seems an inevitable feature of a great many (and perhaps all) actions, for unlike the objective consequentialist, the subjective consequentialist doesn't need to know what the consequences *will be* in order to arrive at a conclusion about the deontic status of an action; all they need to know is which consequences are *possible*, and what the available evidence says – or perhaps even just what their beliefs say – about the probability of each possible outcome.

While subjective consequentialism seems be an advance on objective consequentialism, not all worries about the epistemic demands of consequentialism are assuaged by adopting this type of theory. For a start, working out deontic conclusions according to subjective consequentialist reasoning would require knowledge of what all the *possible* outcomes of an action are (or at least of what all the possible outcomes are which, in fact, one should assign a non-zero subjective probability to, given the evidence or one's beliefs). And in one way, that looks like a *more* demanding requirement than

[6]Frank Jackson, 'Decision Theoretic Consequentialism and the Nearest and Dearest Objection', *Ethics* 101 (1991), pp. 461–82; Frances Howard-Snyder, 'The Rejection of Objective Consequentialism', *Utilitas* 9 (1997), pp. 241–8; James Lenman, 'Consequentialism and Cluelessness', *Philosophy & Public Affairs* 29 (2000), pp. 342–70. As Hilary Greaves and others point out, epistemic concerns such as these assail all sorts of moral theories – including non-consequentialist ones – according to which it matters what the consequences of our actions are; which, according to some, means that such concerns assail all *plausible* moral theories. (Greaves, 'Cluelessness', *Proceedings of the Aristotelian Society* 116 (2016), pp. 311–39, at 312.) We shouldn't overestimate the extent to which the cluelessness objection presents a problem for *all* plausible non-consequentialist theories, though. According to some such theories, what matters for determining the *deontic* status of an action is not actually its consequences, but the agent's *expectation about* the consequences, or their *intentions with respect to* the consequences, and in neither case is there a particular epistemic worry of the type described here. It is true that according to some such theories the actual consequences of an action contribute to how *good* that action is; but their account of the *deontic* status of actions (their permissibility, being forbidden or required) is not tied, in the way characteristic of consequentialism, purely to facts about their evaluative properties (such as being (one of) the best actions). And while it might be plausible that a natural constraint on moral theories is that their accounts of what is permissible, required and forbidden should be apt to play an action guiding role in the deliberation of agents, it is not nearly so plausible that their accounts of what is good or bad (or how good or bad things are) should respect that constraint, *unless* those evaluative facts are supposed to determine the deontic ones: it might be implausible that we act wrongly in doing what we could not know is forbidden; but as far as I can see there is nothing at all implausible in the suggestion that we are often ignorant of how good something we have done or propose to do really is.

[7]Jackson, 'Decision Theoretic Consequentialism'. As Christian Piller reminded me, not all subjective consequentialists are motivated in this way.

objective consequentialism's requirement that we know just what will *actually* happen. Indeed, part of the bite of Lenman's 'cluelessness objection' to consequentialism (which below I will sometimes refer to as the *standard* cluelessness objection) is that the various consequences *and possible consequences* of any action are so various, so distant, and related in such complex ways to that action that we lack knowledge not only of what those consequences are, but also of what they *might* realistically be.

II. Indifference to the unforeseeable

Let's assume that the cluelessness objection is a serious one. How might a consequentialist deal with it? The most promising way is to start by adopting some form of subjective – rather than objective – consequentialism. But how will the subjective consequentialist deal with the fact that we are in no position to know what all (or even most) of even the *possible* consequences are?

Clearly, in at least some cases, we cannot have such knowledge. So the solution must be to give a principled reason for saying that the deontic status of an action depends (in so far as it depends upon the consequences at all) only upon the *foreseeable* consequences or possible consequences – otherwise, subjective consequentialism is no more plausible (in respect of this set of epistemic worries, at least) than objective consequentialism. And this is just the proposal defended by, for example, Hilary Greaves, who argues that:[8]

> *EVF*
> The expected value of an action is determined entirely via its foreseeable effects.

As Greaves explains, EVF can be justified by – and indeed seems only to be plausible on the basis of – a 'Principle of Indifference' governing rational credence.[9] This principle is controversial, but I will not discuss it here. And nor will I discuss the 'complex cluelessness' problems which Greaves goes on to present, which she does not think are so easily solved simply by appealing to EVF, or her proposals about what the problem of cluelessness might really amount to.[10] Instead, I want to argue, beginning in the next section, that appealing to EVF is not unproblematic even in perfectly ordinary cases, and that is because of the way in which it appeals to the foreseeable.

One thing to note straight away is that EVF wisely mentions *foreseeable* – and not *foreseen* – effects. The view that the expected value – and therefore, according to subjective consequentialism, the deontic status – of an action is determined by its *foreseen* effects has little to recommend it, at least in combination with the view that what is permissible or obligatory is determined by expected value, if only because it would licence more or less any judgement about what one is permitted to do, so long as one is careful to engineer the right epistemic circumstances for oneself: by ensuring that I don't know about or suspect any bad consequences of my actions, I would be permitted to perform them. This is not a consequence of EVF as stated, though, and so we should prefer EVF and its appeal to foreseeable, rather than foreseen, effects.

But EVF's appeal to foreseeable effects doesn't make for an attractive view in other ways. Or at least, as I shall argue in what follows, subjective consequentialism – i.e. the

[8]Greaves, 'Cluelessness'.

[9]For a precise statement of the Principle of Indifference, see Greaves, 'Cluelessness', p. 318. Greaves explains nicely what motivates EVF in 'Cluelessness', §2; and she discusses and defends the Principle of Indifference in 'Cluelessness', §§3–4.

[10]See Greaves 'Cluelessness', §§5–6.

view that the deontic status of actions is fully determined by their expected value – in combination with EVF entails either deontic indeterminacy or the possibility of arbitrary deontic variance. Neither result is attractive.

The moral facts display *deontic indeterminacy* iff it is neither the case that (a) it is permissible (for someone) to φ, nor that (b) it is not permissible (for someone) to φ. To say that our obligations or permissions are subject to *arbitrary deontic variance* is to say that whether we have an obligation or permission to φ depends upon something (i) variable, such that our obligations can change from one time or context to the next, and (ii) morally irrelevant. For example, a view which held that whether or not my actions are permissible depends upon who is evaluating them would represent our obligations as being subject to arbitrary deontic variance – and would, surely, be rendered utterly implausible on that account.

III. Unforeseeable consequences

Unforeseeable consequences are not *metaphysically indeterminate* states of affairs. They are just states of affairs which we have no *epistemic access* to: we cannot acquire sufficient evidence to epistemically ground beliefs about them; or we are in some other way unable to have beliefs about them which are in epistemically good shape (as opposed to just being right by luck); or perhaps we are in no position to have beliefs about them at all. If e is a merely unforeseeable consequence of φ-ing, then it is nonetheless a fact that e is a consequence of φ-ing.

There are many different causes – or, perhaps better, explanations – of unforeseeability. Sometimes, as Lenman argues, a state of affairs is so far in the future and mediated by so many complex facts that we simply don't have the mental capacity to know about it. But I want to focus on a different way in which a state of affairs can be unforeseeable to a particular person.

The unforeseeable is what it is *not possible to foresee*. Lenman and others are mainly concerned with what it is not possible for *anybody* to foresee, or what it is not possible for a person to foresee *in any circumstances*. But I want to focus instead on a rather different sort of case.

It is well known that possibilities do not 'agglomerate': from (i) *possibly* p and (ii) *possibly* q it does not follow that (iii) *possibly* $(p \,\&\, q)$. Since foreseeability is a matter of what it is *possible* to foresee, this fact about possibility has implications for foreseeability: (i) *p is foreseeable by A* and (ii) *q is foreseeable by A* do not jointly entail that (iii) A *could foresee both p and q*. And, as I shall now argue, it is not just that logic fails to guarantee that A could foresee both p and q if A could foresee each of p and q separately: there will be many actual cases in which A in fact cannot foresee both p and q, even though each of p and q are foreseeable to A.

As Lenman and many others have pointed out, some consequences of our actions (or of the actions of others) are so remote and/or mediated by complexity that no human being could foresee them regardless of how clever they are, how much effort they put into their inquiries, and how much time they spend on trying to find out the relevant facts. But there are many consequences (and other facts) which a person *could* foresee, but only if they spent more or less all their time and effort on trying to acquire the relevant knowledge. The vast majority of the actual consequences of our actions are such that if we can foresee them at all, it is only by spending more or less all our time and effort on investigating them.

Recall the comment attributed to E. M. Forster which is the epigram of this article. For our purposes, the point is that if some fact about the consequences of an action

could be foreseen by someone, but only if that person spent more or less all their available time and mental resources on gathering the relevant information, then that person would not have left the time and mental resources required to foresee *other* facts – which could be foreseen by them, *if* (but only if) that difficult to foresee fact were *not* foreseen (or, more precisely, if that person did not put their time and effort into trying to put themselves in a position to foresee that difficult-to-foresee fact). As I said, the possibility of (a particular person) foreseeing things does not agglomerate: it does not follow from the fact that p is foreseeable by A and the fact that q is foreseeable by A that it is possible for A to foresee *both* p *and* q – it might be that A could only foresee p if they didn't foresee q, and could only foresee q if they didn't foresee p. And the explanation for this, in many cases, is that A's foreseeing p requires so much of A's time and effort that they are left with insufficient time and effort to know what is required to foresee q, and vice versa.

A similar point applies, of course, to all sorts of things. In just the same way that the time and effort required to foresee p might exclude spending the time and effort required to foresee q, and vice versa, a person might have the potential to be a successful professional footballer, or a successful banker; but not a successful banker who is also a successful professional footballer, because each of those professions requires the sort of dedication which precludes the dedication required for the other.

It will be useful to introduce some jargon. I'll use *ultimately foreseeable* and *ultimately unforeseeable* to denote what is foreseeable or unforeseeable once a person has enacted some decision about which of p or q to investigate, in cases where foresight of p and foresight of q are each possible, but are not *com*possible. In some cases, p and q are both possible, but they are not *com*possible, meaning that they cannot be true together; and in some cases, p and q are foreseeable, but at least one of them is *ultimately* unforeseeable, meaning that they cannot be foreseen together.

In cases such as the ones I have been describing, it is both true and knowable that *something* is ultimately unforeseeable by A; and something more specific is both true and knowable, namely that either p or q is ultimately unforeseeable by A. But it is indeterminate *which* of p or q is ultimately unforeseeable by A, unless and until A embarks upon an effort to know p or an effort to know q, rendering A's available time and available effort insufficient for foreseeing the other fact, q or p.[11] This is what will be important in what follows. The crucial point to bear in mind, once again, is that both p and q might be foreseeable by A, but at most one of p or q is *ultimately* foreseeable by A, and it is indeterminate which, unless and until there is a fact of the matter about which of p or q A no longer has (or will lack) the time and/or available effort to come to foresee (because they are investigating the other).

IV. A problem with appealing to foreseeability

Suppose a particular action, A's φ-ing, has two actual effects, e_1 and e_2. And suppose that e_1 and e_2 are not jointly foreseeable, though each of e_1 and e_2 are foreseeable

[11] A's embarking on some *other* endeavour (aside from investigating p or q) which similarly takes the necessary time and effort away from investigating p or q would also, of course, settle which of p or q (or, in this case, both) is doomed to ultimate unforeseeability by A. (That is why I said, above, that *at least* one of p or q is ultimately unforeseeable.) But I leave aside that unproblematic complication here and in what follows.

individually. Suppose, further, that the values of e_1 and e_2 (and other relevant consequences) are such that if e_1 but not e_2 is included in our reckoning *A*'s φ-ing is permissible, whereas if e_2 but not e_1 is included in our reckoning *A*'s φ-ing is not permissible, according to consequentialism.

According to objective consequentialism, the deontic status of *A*'s φ-ing is unknowable, for knowledge of the deontic status of *A*'s φ-ing depends upon knowledge of e_1 *and* e_2. And that – knowledge of *both* – is knowledge that we cannot have.

As we saw, this leads some to reject objective consequentialism – driven by the thought that what is wanted is an *action guiding* normative theory – and endorse subjective consequentialism, which requires us to calculate not actual but *expected* value. But if we are subjective consequentialists, we must still deal with the fact that e_1 and e_2 are not jointly foreseeable, though each of e_1 and e_2 are foreseeable individually. Because each of e_1 and e_2 are foreseeable individually, we cannot simply say that both e_1 and e_2 are apt to be ignored in our calculation because they are the wrong sort of thing to be included in our reckoning. But because e_1 and e_2 are not *jointly* foreseeable, nor can we simply say – as subjective consequentialists – that both should be included in our reckoning, either.

The only option seemingly available to the subjective consequentialist in such a case is to say that one of *either* e_1 *or* e_2 are to be included in our reckoning. But which? According to EVF, we should take into account the foreseeable consequences, but not the unforeseeable ones. This seems not to help us out of our predicament, though, for each of e_1 and e_2 is foreseeable. In order to be helpful, EVF must distinguish between these two foreseeable consequences, and the natural (and, as far as I can see, only) way of doing so is by specifying that we should take into account all of the *ultimately* foreseeable consequences while ignoring all *ultimately* unforeseeable ones. Recall that 'ultimately' foreseeable and unforeseeable means what is foreseeable or unforeseeable once a person has enacted some decision about what to investigate.

But if it is not yet decided which of e_1 or e_2 is ultimately foreseeable and which is ultimately unforeseeable, then it is indeterminate which we should take into account, according to this modified version of EVF. So if the subjective consequentialist theory we are considering incorporates EVF, it is indeterminate, according to that theory, whether *A*'s φ-ing is permissible if it is not yet decided which of e_1 or e_2 is ultimately foreseeable and which is ultimately unforeseeable. And arguably it *isn't* yet decided which of e_1 and e_2 is ultimately foreseeable and which is ultimately unforeseeable (in advance of a decision being made and enacted about what to investigate), so the deontic status of *A*'s φ-ing is indeterminate, according to subjective consequentialism.

But what if it *is* decided which of e_1 or e_2 is ultimately unforeseeable, either because a decision about what to investigate has been made and enacted, or because even in advance of such events there is a fact of the matter about how the future will unfold in that respect? Then the problem is not deontic indeterminacy, but arbitrary deontic variance. What decides which of e_1 or e_2 is ultimately unforeseeable is a fact about how we direct our efforts and spend our time: if we decide to investigate e_1, e_2 is ultimately foreseeable and vice versa. But our deciding to concentrate our time and effort on investigating e_1 rather than e_2 is – or at least might be – a morally irrelevant fact. Both e_1 and e_2 are morally relevant facts; but my deciding to investigate one rather than the other might be motivated by prudential, pragmatic, aesthetic or other non-moral reasons, or no reasons at all. And regardless of the reasons (or lack of reasons) behind my choices here, the fact that I have made such a choice seems far removed from the sorts of considerations which bear upon the deontic status of the action I am

assessing. It is wildly implausible that the permissibility of *A*'s φ-ing depends upon a fact about how a deliberator has chosen to spend their time – especially when that deliberator is not *A*.[12] Surely no subjective consequentialist would welcome the news that the best version of their theory (the only one which avoids the original cluelessness objection, by limiting the relevant considerations to the ultimately foreseeable ones) tells us different things about the permissibility of *A*'s doing something depending upon whether we choose to think about one of the consequences of *A*'s actions rather than another.

And of course, there is absolutely no guarantee that we will all direct our efforts in the same ways; so what is ultimately unforeseeable for *A* (because *A* takes the time to find out about e_1, leaving no time to find out about e_2) might not be what is ultimately unforeseeable for *B* (who takes the time to find out about e_2, leaving no time to find out about e_1). What, then, does our subjective consequentialist theory say? Does it say that the deontic status of *A*'s φ-ing is different for *A* and for *B*? That relativistic conclusion would presumably not be welcomed by most subjective consequentialists. But then whose epistemic situation is to be privileged?

An obvious reply is that subjective consequentialists should consider *all* ultimately foreseeable consequences relevant to determining the deontic status of actions – ultimately foreseeable *by anyone*, that is. Then, since e_1 is ultimately foreseeable by *A* and e_2 is ultimately foreseeable by *B*, both e_1 and e_2 are apt for inclusion in the relevant calculations. But this reply is not at all unproblematic.

For a start, subjective consequentialism is often thought plausible precisely because the theory is supposedly action guiding for – because it delivers a conclusion when applied by – non-ideal epistemic agents who don't know all the facts. This is certainly taken by many to be the main advantage of subjective over objective consequentialism. But a form of subjective consequentialism which requires us to include in our reckoning all the relevant facts foreseeable by *anyone* no longer promises to be usable by particular agents or deliberators, who don't themselves know as much as everyone collectively knows. This form of subjective consequentialism is in the same boat as objective consequentialism, according to which there is a fact of the matter about the deontic status of *A*'s φ-ing, but not a fact which can be known by any particular agent (on the basis of consequentialist reasoning, at least). So, as far as the cluelessness objection and the action-guiding constraint go, subjective consequentialism would end up in the very same boat as objective consequentialism if it were to adopt this strategy to avoid relativism or arbitrary privileging of particular epistemic perspectives.

In any case, the problem would not be entirely solved by such means even if the subjective consequentialist were willing to live with a theory no less action-guiding than objective consequentialism for non-ideal agents: even if requiring the inclusion of *all* ultimately foreseeable consequences (i.e. foreseeable by anybody) fixed a deontic status for *A*'s φ-ing in a non-relativistic way *in the actual world*, there would remain a problem

[12]Couldn't what someone spends time thinking about make a difference to how much happiness there is in the world, for example, so that some (utilitarian) consequentialists would be happy to say that deontic variance on the basis of such facts isn't *arbitrary*? It could. But notice that in such cases, the morally significant difference is made not by the fact that someone spends their time thinking about some particular things *per se*, but by the happiness which results from that: it is the *happiness*, not the fact that someone has directed their attention in a particular way, which is morally salient. What I am suggesting is that it would be bizarre to think that in cases where the facts about how attention and effort are directed have no *further morally significant* consequences (such as increasing well-being or happiness), they could in themselves determine variations in deontic status in a non-arbitrary way.

of arbitrary deontic variance *across possible worlds*. Call this the *modal problem of arbitrary deontic variance*, or just the *modal problem* for short. Recall that what determines the ultimate foreseeability of some consequence for some particular deliberator is a fact (or set of facts) about what that deliberator investigates (and so, what they leave time and resources available for besides that investigation). It is obvious, I think, that those facts could have been different: most deliberators could have spent their time investigating things other than what they in fact investigate; and in many cases, their investigating different things would make a difference to what is, for them, ultimately foreseeable.

This claim, which I said I take to be obvious, does not rely upon any contentious assumptions about free will or determinism: someone who, in a non-actual possible world, ends up investigating something other than what they investigated in the actual world needn't do so because they have *freely chosen* to; and nor does determinism necessarily rule out such a possibility, for it might be explained by the fact that, in that non-actual possible world, the antecedent conditions (perhaps even going back to the initial state of the universe) are different from the actual ones. Whatever explains the fact that particular deliberators differ across possible worlds in respect of what they investigate, the fact that they *do* so differ is sufficient to generate the modal problem.

As I have described it so far, the modal problem is that given a fixed set of deliberators, difference across possible worlds in respect of how those deliberators spend their time and effort generates (according to the form of subjective consequentialism under consideration) deontic variance: *A*'s φ-ing might be permissible in the actual world (given that in the actual world deliberators investigate certain consequences, rendering others ultimately unforeseeable); but simultaneously in some non-actual possible world, *which differs only in respect of what those deliberators investigate* (and other facts which are required by this difference), *A*'s φ-ing might be impermissible. This is clearly modal deontic variance, and it is plausibly *arbitrary* modal deontic variance. It might be bad enough that subjective consequentialism ends up positing or allowing for modal deontic variance. But if that is sufficient reason to reject the theory, we don't need to appeal to cluelessness or ultimate unforeseeability to reject it. Pretty much any otherwise plausible form of consequentialism posits or allows for such variance, since the deontic status of each action or act depends upon the sorts of facts about consequences which are more or less bound to be different across possible worlds (even close-by ones). What is particularly troubling here – and what should arguably trouble even *consequentialists*, who are not averse to deontic variance *per se* – is the fact that what makes the difference to the deontic status of *A*'s φ-ing in different possible worlds is not just facts about what the consequences (or even character) of *A*'s φ-ing are in each of those worlds, but rather an apparently morally irrelevant difference between them in respect of what some deliberators (who might not be the agent of the action themselves, and might have nothing whatsoever to do with the action, reflecting upon it as an historical event thousands of years later) *think about*.

But the modal problem is even worse than I have described so far. Possible worlds differ not only in respect of what fixed populations of deliberators spend their time and effort doing, but also in respect of which deliberators exist. Populations of deliberators across possible worlds are fixed only artificially, and imposing restrictions upon which deliberators in a possible world count, or which possible words count (on the basis that only those with the same population of deliberators as exist in the actual world count) seems as deontically arbitrary as anything we've already seen. A possible world in which the population of deliberators is much larger than the one in the actual world might

have, in its set of consequences which are ultimately foreseeable *by someone*, consequences which are not ultimately foreseeable to anyone in the actual world. A possible world in which the population is much smaller might have fewer consequences ultimately foreseeable *by someone* in it, even if those deliberators who exist in both worlds do the same things. So if, as per the suggestion we are now considering, which seems like the only way for the subjective consequentialist to avoid relativism, the deontic status of *A*'s φ-ing depends upon what is ultimately foreseeable *by someone*, it seems that the deontic status of *A*'s φ-ing depends upon who is in the population of those deliberating – or, better, *available for deliberation* – in particular possible worlds. And that is another source of surely arbitrary modal deontic variance.

The simple solution, of course, would be to say that what matters for the deontic status of *A*'s φ-ing is not just everything that is ultimately foreseeable *by someone* in *any particular possible world*, but rather everything that is ultimately foreseeable *by someone* in *any* possible world: if a consequence of *A*'s φ-ing is ultimately foreseeable *in any possible world* by *any individual in that possible world*, then it matters *in every possible world* for the permissibility of *A*'s φ-ing. But this is an unsatisfying solution because, in so far as it avoids the sort of modal deontic variance I have just described, it comes at several serious costs.

Most seriously, it threatens to collapse subjective consequentialism into objective consequentialism; or at least to saddle subjective consequentialism with similar and equally objectionable features. Arguably, for *every* actual consequence of *A*'s φ-ing, there is *some* possible world in which that consequence is ultimately foreseeable by someone – even if that someone is the sort of super-deliberator which *only* exists in very distant possible worlds. But if so, the theory we are now considering requires us to take all of the actual consequences into account. And that looks *a lot* like objective consequentialism. The distinctive and arguably main appealing feature of subjective consequentialism, namely that it accommodates our real-world ignorance, has entirely gone, for the version of EVF upon which the theory under consideration relies tells us not to disregard *any* actual consequences (since they are all foreseeable in the relevant sense, i.e. they are each ultimately foreseeable by someone in some (far off) possible world), leaving us with just the problem of cluelessness with which our flight to subjective consequentialism began. So, subjective consequentialists who want to avoid the cluelessness objection by appealing to a version of EVF had better not argue that EVF should be interpreted such that no actual consequence which is ultimately foreseeable by someone in some possible world should be ignored in our (actual world) reckoning. On the other hand, *unless* they do this they face the modal problem I set out above.

Even fixing *the population* of deliberators by considering only those possible worlds in which the same deliberators exist as exist in the actual world would not suffice to solve or avoid the modal problem, since what is foreseeable partly depends upon how much time and cognitive ability those deliberators have. A population *P* in some non-actual possible world might not be able to know what that same population *P* could know in the actual world, simply because in that non-actual world the members of *P* are less intelligent or live much shorter lives, meaning that while they have the time and cognitive capacity to know *both* e_1 and e_2 in the actual world, they can know only e_1, for instance, in that non-actual world. But again, facts about the longevity of those deliberating about the permissibility of *A*'s φ-ing, or their intelligence, are surely morally irrelevant with respect to *A*'s φ-ing. To avoid such arbitrary deontic variance across possible worlds, we would need to restrict attention not only to those possible worlds in which the population is the same as in the actual world (or at least in which it is not

significantly larger or smaller), but to those possible worlds in which many of the facts about that population (such as their intelligence and longevity) are also fixed. And fixing *those* things would not deal with the modal problem in its first variety, discussed above, which concerns what even those same people, with their same intelligence, longevity etc. might variously end up spending their time and effort on in different possible worlds.

This last point shows why it doesn't help to say – as some might find it natural to say – that the only relevant deliberative perspective for assessing the deontic status of an action is that of *the agent*. The modal problem undermines that proposal because the agent might have directed their cognitive efforts differently; and in any case, the original, non-modal problem did not depend upon assuming that the deliberative perspectives of others are relevant – that consideration only *worsened* a problem which would be there anyway.

All in all, subjective consequentialists who put their faith in some version of EVF in order to avoid the cluelessness objection are in trouble. They cannot endorse any view which looks sufficiently like objective consequentialism to reintroduce the *standard* cluelessness objection. And I suspect that the attraction of principles like EVF is that they *seem* to more realistically limit the range of things we need to take into account in our deliberations about the deontic status of actions while still appealing to something, namely foreseeability, which is robust and stable enough across possible worlds, times and people to avoid relativism or arbitrary deontic variance (which would result from using being *foreseen* as the criterion for consideration). But this appearance is illusory, as I have shown: because possibility does not agglomerate, and foreseeability is a matter of the possibility of knowing, it is in some cases either indeterminate what counts as (ultimately) unforeseeable, or what determines the (ultimate) foreseeability of some consequences is some fact which is morally irrelevant and contingent, varying between both individuals and times, and possible worlds (generating arbitrary deontic variance).

V. Old wine in an unnecessarily fancy bottle?

It might be thought that the difficulty I presented in the previous section for subjective consequentialism is no different from, or does not add anything of substance to, a problem which could already be raised without mention of agglomeration and *ultimate* foreseeability and unforeseeability. What I'll call the *shallow problem* for subjective consequentialism is just that in any appeal to what is foreseeable to agents or deliberators, there is the danger of problematic deontic variance given that what is foreseeable is different for different people, at different times, and in different possible worlds. But there is something deeper about the problem I have been concerned with.

The shallow problem (which concerns just the contingent and variable nature of what is foreseeable to different people) could be addressed in a way that flows quite naturally from the set of concerns which motivate subjective consequentialism in the first place. As an improvement upon objective consequentialism, subjective consequentialism is motivated by the thought that it is unreasonable – or perhaps incoherent – to make the permissibility of a person's actions depend upon facts which they have no epistemic access to, and *could* have no epistemic access to even though they are not negligent (and perhaps to make the deontic status of someone's actions unknowable to others who have no epistemic access to the relevant facts). Thus subjective consequentialism supposedly strikes the right balance between (i) sensitivity to the situation and

capacities of agents and deliberators and (ii) robustness, in so far as the deontic status of actions does not depend upon just what *is foreseen* by agents or deliberators (which would be implausible, for the reason explained above), but depends rather upon facts about agents' and deliberators' situation and capacities which are not a matter of choice for them, as it is at least partly a matter of choice what is foreseen by them. The shallow problem is easily dealt with by pointing out that it is not necessarily *arbitrary* deontic variance if what is foreseeable to one person is not foreseeable to another (or is not foreseeable to that person in a different possible world), and thus the deontic status of an action varies: it is deontic variance, for sure; but in many cases it is not arbitrary because it is driven by differences in the situations and capacities of agents and deliberators, which are precisely the sorts of differences which, according to the subjective consequentialist, a moral theory should be sensitive to.

On the other hand, the problem I have been describing cannot be dealt with so easily. First, there is an added dimension, which is the possibility of not only (arbitrary) deontic variance, but also of *deontic indeterminacy*. The shallow problem does not involve indeterminacy, because there is no reason to think that it is *generally* true that there is no fact of the matter (yet) about what is foreseeable and unforeseeable. In the sorts of cases usually considered by those concerned with the epistemic dimension of consequentialism, the facts about what is foreseeable and unforeseeable are decided already by facts about agents' or deliberators' circumstances and/or capacities. But in the cases I have been concerned with, it might be indeterminate what is (ultimately) foreseeable and what is (ultimately) unforeseeable *even if the facts about agents' and deliberators' circumstances and capacities are fixed*, for the unforeseeability of some things depends, in ways I explained above, on individuals' *choices* which may not yet have been made.

Second, regardless of this extra dimension of possible indeterminacy, the problem I have been considering is harder to deal with than the shallow problem for the subjective consequentialist because individuals' choices about what to think about don't seem to be the sorts of things which plausibly constrain moral theories about the deontic status of actions in the way facts about individuals' circumstance and capacities (which, it might be noted, are typically beyond individuals' control) are. Subjective consequentialists – advisedly – don't pin much on what *is foreseen* by particular agents or deliberators, and a good reason for this is the fact that what is actually foreseen by an individual is, to some extent, in most cases a matter of choice (since one can generally fail to know something by ignoring it, or by cultivating negligence), and it is implausible that the deontic status of actions depends upon individuals' choices about what to ponder or which cognitive habits to cultivate such as to acquire or avoid knowledge. (The wrongdoer who is careful to ignore the morally relevant consequences of her actions, or other morally relevant facts, does not thereby do less wrong in virtue of lacking knowledge of those things.) But if it is appropriate to shy away from pinning much on what *is foreseen* for this reason, it is also appropriate to shy away from pinning much on the choices individuals make which are similarly morally irrelevant but apt to render some things unforeseeable which, so far as their circumstances and capacities are concerned, would not otherwise be unforeseeable.

Put another way, some of what is unforeseeable is unforeseeable more or less entirely because of our circumstances or capacities, and those unforeseeable facts are – from the perspective of typical subjective consequentialists – apt to be ignored for a good principled reason, namely that our circumstances and capacities are natural constraints upon moral theory. But no serious subjective consequentialist should consider our

choices similarly natural constraints upon moral theory, such as to limit what determines the deontic status of actions according to the choices we make, for the resulting moral theory would say that it is up to us what is permissible and impermissible to an implausible extent. Yet the problem I discussed above is precisely that subjective consequentialism and EVF seem to entail that just such facts about our choices *do* limit what determines the deontic status of actions. It turns out that while some unforeseeable facts can be quite properly ruled out by subjective consequentialists, some (the ones whose ultimate unforeseeability depends not upon just our circumstances and capacities, but crucially also upon our choices concerning how to direct our attention) cannot, for ruling them out requires us to recognize the choices we make as proper grounds for determining which facts are to be considered and which aren't. That is why the problem I have discussed involves *arbitrary* deontic variance, whereas thinking about foreseeability constraints in general raises the prospect of deontic variance, but not necessarily *arbitrary* deontic variance.

VI. 'Double' subjective consequentialism

According to what I'll call *double subjective consequentialism*, the deontic status of A's φ-ing depends upon the expected value ranking of A's φ-ing which results from applying EVF *and* the principle that what deliberators should investigate is itself a matter of doing what is best from an expected value point of view. It might be hoped that such a view avoids the incompossibility problem because there will be a determinate, non-arbitrary fact of the matter about whether to direct one's efforts towards e_1 or towards e_2, and therefore a determinate, non-arbitrary fact of the matter about the deontic status of A's φ-ing: if one should (by expected value lights) investigate e_1, then the deontic status of A's φ-ing is determined only by its evaluative ranking (modified by EVF) when e_1 but not e_2 is taken into account; and vice versa if investigating e_2 would have a higher expected value.[13]

It is true that treating decisions about what to investigate as subject to the same theory as decisions about what to do (as if investigating were not doing something!) is perfectly natural – and indeed so natural that it hardly seems fitting to call the view I am currently considering *double* subjective consequentialism at all: it is really just consistently applied subjective consequentialism. But therein lies the problem. For if investigating consequences is doing something (which it surely is), such as to render it under the purview of the standard subjective consequentialist theory, it is also apt to generate precisely the same incompossibility problem at *that* level.

If it is not already obvious why, let me explain. The double consequentialist proposal requires us to (1) first come up with an evaluative ranking for the various investigations we might go in for, and then (2) evaluatively rank (other) actions according to the expected value calculus modified by EVF – where EVF ambiguities such as the ones noted above are resolved in favour of whichever outcomes or consequences are foreseeable if we investigate what we ought to according to that first evaluative ranking at stage (1). But investigations have consequences too, and they *might* well be every bit as far-reaching and mediated by complexity as the consequences of other actions. So, at stage (1) we surely have to say that the kind of expected value calculus we require is one modified by EVF – for precisely the same reasons as motivated adopting EVF in the first place, above.

[13]Thanks to two anonymous referees who each pressed me to address this idea.

But, obviously, we are now going down the same road: investigating e_1 itself has some (possible) consequences, c_1 and c_2, each of which is foreseeable but which are not *jointly* foreseeable, and if we take c_1 into account but not c_2 then investigating e_1 is better than investigating e_2; but vice versa if we take c_2 into account but not c_1. In precisely the way we have seen before, deontic indeterminacy or arbitrary deontic variance threatens. It threatens initially at stage (1); but since stage (2) conclusions depend upon and are determined by stage (1) conclusions, it threatens at stage (2) as well.[14]

As I've already indicated, it would be naïve to respond that investigations into the consequences of (other) actions don't themselves have morally significant consequences. They very well might, and if you doubt that then consider for example the fact that learning things about what might happen if something is done can change our attitudes and feelings towards people, and therefore also our relationships (and so even who exists, given how people generally come to exist), and even *investigating* such things, even unsuccessfully, can modify our attitudes and feelings in ways that have consequences (as if those modifications weren't morally important in themselves). Merely focussing my attention, for example, on the sorts of ways in which a loved one *might* respond if I were to behave in a particular way (given what I know about them) can bring to salience facts about them, about me or about our relationship which would otherwise remain occluded or ignored, and their salience could modify my attitudes or feelings towards them, myself or our relationship – possibly with very real consequences for the way I treat them, and hence how our relationship develops or withers, and hence … etc. – and all of this could happen regardless of how successfully I predict what their actual response will be.

The fact that investigations (or deliberations, or however we characterize finding out about possible consequences) *might* have morally significant consequences is enough: as I explained above, the cluelessness problem for subjective consequentialists is that we need to know about consequences that there *might* be, even if they won't in fact occur.

It won't do, either, to abandon subjective consequentialism at stage (1) in favour of objective consequentialism about investigations: not only would that arbitrarily treat acts of investigating things as subject to different standards of moral evaluation from all other acts; it would also plunge us straight back into the standard cluelessness objection, since facts about what it is best to investigate might be cognitively or epistemically unavailable to us (as is always the danger with objective consequentialism), meaning that however subjective the consequentialism at stage (2), we could not know the deontic status of actions since stage (2) conclusions depend upon stage (1) conclusions.

(There is another problem for double consequentialism too. What do we say about optimific ties, in which investigating each of e_1 and e_2 are equally good? Standardly, subjective consequentialism says that it is permissible to do either. But that permits a person to decide whether or not their action is wrong, as above, on the basis of whether or not they choose to ignore bad consequences. And this is extremely implausible.)

A variant of the double consequentialist approach builds the time (or perhaps context) of A's φ-ing, and especially its time relative to deliberation about the possible outcomes, into the theory in a way which is plausible and important, but still not helpful for addressing the incompossibility problem. According to *time/context sensitive double consequentialism*, there is no saying whether A's φ-ing *per se* is permissible or

[14]The argument I have given here might ground a complaint about an *infinite regress* against subjective consequentialism. But I am more interested, here, to point out the way in which the incompossibility problem recurs *at each stage*, such that it causes trouble even if the chain of stages is finite.

obligatory; it is *A*'s φ-ing *at time t* (*given history h*) which is apt for deontic evaluation. So, we mustn't say simply that we are interested in whether *A*'s φ-ing is permissible or obligatory, and that in order to find out we had better investigate the consequences of *A*'s φ-ing. Rather, we must say that the deontic status of *A*'s φ-ing *after investigating the consequences* might be different from its deontic status before having done so.

Subjective consequentialists should recognize that new information can make a difference to what is permissible or obligatory. However, the cluelessness problems arise not in relation to new information, but in relation to the information we *could* have, but don't. As we've seen, we need a plausible principle which breaks the deadlock when outcomes are individually foreseeable but not jointly foreseeable. We must presumably say that in some cases (though perhaps not others) it would be wrong not to investigate some particular consequences, and therefore we ought to (1) investigate those consequences and then (2) do only what is permissible *after investigating those consequences*.

But this is just a version of the proposal we have already discussed in this section. Now the proposal is not just that the deontic status of *A*'s φ-ing depends upon what *would* be discovered if we investigated what we ought to (by consequentialist lights), but that the deontic status of *A*'s φ-ing depends upon what *has* been investigated. However, it cannot be simply that the deontic status of *A*'s φ-ing depends upon what *A* or anyone else has investigated, for as noted above that makes it too easy to determine the deontic status of actions simply by choosing to ignore bad (or good) outcomes. The idea must therefore be that the deontic status of *A*'s φ-ing is affected by facts about *the right* investigations having been carried out or not carried out. And here the familiar problem arises, for as explained above, subjective consequentialists have as much trouble specifying which the right investigations are as they do specifying what the right actions of other sorts are.

So, double consequentialism is no better off than 'one stage' consequentialism. In fact, double the consequentialism, double the trouble.

VII. Conclusion

If Lenman's standard cluelessness objection is a serious objection to objective consequentialism, then it sets up a serious objection to subjective consequentialism too. It is not that unforeseeable consequences pose the *immediate* problem – in respect of delivering knowable conclusions about the deontic status of particular actions – for subjective consequentialism that they pose for objective consequentialism. To avoid the immediate problem, subjective consequentialists can appeal to a principle of indifference to unforeseeable consequences, motivated by a general (though controversial) principle of rationality under uncertainty. But the problem is that indifference to unforeseeable consequences, such as is expressed by the EVF principle, raises a dilemma. Either the appealed-to principle is interpreted in a relatively narrow way, in which case (as I demonstrated above) it entails that in at least some cases the deontic status of actions is either indeterminate or arbitrarily variant. Or it is interpreted relatively broadly, to rule out these results, but now saddling subjective consequentialists with just the sorts of commitments which rendered objective consequentialism subject to the cluelessness objection, or at least to similar worries about the theory's ability to provide normative guidance to particular epistemically non-ideal agents in particular cases.

None of this is a problem if, as I believe, the cluelessness objection and similar objections based upon the assumption that a moral theory must be action guiding – in a

quite narrowly construed way – are not serious objections in the first place. But for those who take such objections seriously, subjective consequentialism ought to look no more attractive than objective consequentialism.[15]

[15]Thanks to Dylan Balfour, who started me thinking seriously about indifference principles, and to two referees and the editor of this journal, an audience at the University of York, particularly Christian Piller, and Rob Trueman for comments on a draft of this article.

Cite this article: Jay C (2019). Subjective Consequentialism and the Unforeseeable. *Utilitas* 1–17. https://doi.org/10.1017/S095382081900030X