


ARTICLE

# When a Free Act Costs a Motive: Clearing Consequentialism of Conflict

Austen McDougal 

Stanford University  
Email: [austendm@stanford.edu](mailto:austendm@stanford.edu)

## Abstract

Consequentialist theories that directly assess multiple focal points face an important objection: that one right option may conflict with another. Robert Adams raises an instance of this objection regarding the possibility that the right act conflicts with the right motives. Whereas only partial responses have previously been given, assuming particular views of the relation between motives and acts, an exhaustive treatment is in order. Either motives psychologically determine acts, or they do not – and I defend direct consequentialism on each assumption. Crucially, if motives determine acts, this may be compatible with the ability to act otherwise, but there remains a defense for consequentialism even on these assumptions. What clears consequentialism of conflict is not necessarily that the apparently right act is unavailable, but rather that its outcome is suboptimal once we account for necessary motives. Even if the agent remains free to perform the act, the act costs too much.

Given that life consists of more than merely performing acts, a moral theory that gives adequate direction for how to live must say more than merely how to act. It must also say something about other aspects of living, such as how to feel, how to be motivated, and so forth – each a focal point for moral assessment distinct from acts.<sup>1</sup> It is no objection that a theory with multiple focal points would be complex and difficult to comply with. After all, life is complex, and living morally was never supposed to be easy.

A natural approach for consequentialism is to assess each focal point directly by evaluating its outcome.<sup>2</sup> Consequentialist theories that directly assess multiple focal points face an important objection, however: that the right option for one focal point may conflict with the right option for another. This is the objection from normative conflict. An influential instance of the objection, due to Robert Adams, says that the direct assessment of both motives and acts is problematic since the right motives may conflict with the right act. Following the literature, I focus on motives and acts

<sup>1</sup>For an overview of focal points and the various options for assessing them, see Kagan (2000).

<sup>2</sup>And some, such as Greaves (2020: 425), would urge that the classical utilitarians themselves preferred this everywhere-direct approach.

in this article, although my response will suggest a general strategy on behalf of direct consequentialism for a variety of focal points.

The objection from normative conflict turns on significant assumptions about agency and has exercised a number of serious philosophers on both sides. Responding to the objection, defenders of direct consequentialism say that the apparently right act cannot be right after all if it requires the agent's motives to be different than they, in fact, are.<sup>3</sup> Those advancing the objection reply that it would be absurd for one's motives to limit what acts one is able to perform.<sup>4</sup> Both sides, however, have failed to appreciate how the latter point may be compatible with the former. Setting straight the various assumptions – and the prospects for consequentialism as a result – is an important task. Once we do so, I claim, direct consequentialism survives the objection no matter how the relation between motives and acts is understood.

I discuss Adams's example of Jack at the cathedral and elaborate on the objection that is supposed to arise in section 1. In section 2, I argue that if one's motives do not determine what acts one performs, then whatever tension exists between the right motives and the right act is not especially problematic for consequentialism. On the other hand, as I explain in section 3, if motives do determine acts, then the apparently right act cannot in fact be right because it is ruled out by the right motives. However, previous ways of articulating this response on behalf of consequentialism have made the further, unattractive assumption that the agent is unable to act otherwise. That need not be so, I argue in section 4. When motives determine acts, what clears consequentialism of conflict is that the apparently right act is suboptimal once we account for the necessary motives. That is, even if the agent remains free to perform the act, the act costs too much – something which consequentialists should care about.

## 1. Introducing normative conflict

Let the consequentialist standard for acts be

*act-direct consequentialism*: an act  $x$  is right iff  $x$  has the best outcome compared to every available alternative.<sup>5</sup>

Next, consider motives, which are “principally wants and desires, considered as giving rise, or tending to give rise to actions” (Adams 1976: 467). In other words, a set of motives is the set of conative attitudes on the basis of which one decides how to act in a particular instance. (Having one motive is compatible with having another, so the relevant alternatives for consequentialism to assess are *sets* of motives rather than individual motives. I sometimes revert to speaking loosely of a motive rather than a set of motives, on the simplifying assumption that the motive set comprises a single motive.) Consequentialists have considered a number of ways to assess motive sets. Some proposals would indirectly assess motive sets based on other focal points to

<sup>3</sup>See Feldman (1993), Dancy (1997), and Greaves (2020). Related arguments by Gruzalski (1986), Mason (2002), and Streumer (2003) also deny that the apparently right act is really right, but they nevertheless go on to abandon direct consequentialism in favor of (respectively) act consequentialism, life consequentialism, and combined act consequentialism.

<sup>4</sup>See Louise (2006) and Skidmore (2018). Parfit (1984: 31–37) also emphasizes this point, although he does not find Adams's objection decisive in the end.

<sup>5</sup>Cf. Pettit and Smith (2000), who argue that acts, no less than any other focal point in consequentialism, are to be assessed directly.

which they are related. For instance, Sidgwick suggests that the right motive is that which produces the right act.<sup>6</sup> However, that standard fails to take into account any value had by the motive that is not mediated by the resulting act. A motive to maximize value while on vacation, as in the example of Jack below, might detract overall from the experience despite producing the best available act in some cases. Insofar as consequentialism holds that “there is one ultimate moral aim: that outcomes be as good as possible,” assessing the outcomes of motives themselves seems more in line with the consequentialist spirit (Parfit 1984: 24). Accordingly, consider the following standard:

*motive-direct consequentialism*: a motive set  $x$  is right iff  $x$  has the best outcome compared to every available alternative.<sup>7</sup>

Adams’s central example illustrates how a motive can have consequences beyond those of the resulting act. In the example, Jack is thrilled to be visiting the cathedral at Chartres for the first time and is motivated solely to take in as much as possible during the day’s visit.<sup>8</sup> This motive will of course lead him to perform a number of acts, such as pausing to gaze at the cathedral’s façade when he first arrives, so the outcome of the motive includes the outcome of such acts. But the outcome of the motive also includes other things, such as the phenomenal quality of his experiences, in addition to whatever acts Jack performs. When he looks up at the façade, for example, he may do so with a certain abandon that shapes his experience in virtue of his motive to see as much as possible.<sup>9</sup> Jack would likely have a different experience were he to perform the same act of looking up on the basis of other motives. Jack might instead be motivated solely to bring about the best outcome, and although he would still enjoy the sight when he looks up, he would enjoy it less from this motive. Furthermore, while Adams emphasizes the effects on Jack, we can also suppose that there would be consequences for others perceiving Jack’s motives. The motive to see everything may be expressed in an enraptured look, whereas the motive to bring about the best outcome may be expressed in non-voluntary glances at his watch. A local resident of Chartres would presumably be charmed by the former and put off by the latter. So the outcomes of motives are not limited to what acts they produce.

A natural consequentialist move is to uniformly apply the test of value directly both to motives and to acts – call this simply *direct consequentialism*. The case of Jack is supposed to raise trouble for direct consequentialism. Specifically, Jack’s case is supposed to show that the right motive set may be the basis for the wrong act. Let the singleton motive to see everything be the right one for Jack to have throughout the day. Its outcome is better than that of any alternative motive set due to considerations mentioned earlier – Jack’s aesthetic experience, the effect on others, and so forth. Being motivated solely to see everything, Jack will stay late into the afternoon poring over entirely

<sup>6</sup>I draw here from Adams (1976: 467–69), who has a useful discussion of various positions adopted by the classical utilitarians regarding the assessment of focal points other than acts.

<sup>7</sup>Some consequentialists refrain from assigning deontic status to motives; see, e.g., Thornley (2022). I worry that this move faces a close relative to the problem from normative conflict: there remains conflict between right acts and the motives that are virtuous or optimal (or some such ethical notion). In any case, I follow the literature in supposing that motives are features of the agent that are to be assessed deontically.

<sup>8</sup>Although the example originates with Adams (1976: 470–72), I take the liberty to add some of my own details.

<sup>9</sup>Elsewhere, Adams (1999: chapters 7, 8, 13) discusses this phenomenon of being wholly devoted to an end.

forgettable parts of the cathedral, when his time would be better spent moving along and leaving earlier so that he can find a nice hotel room ahead of a relaxing evening. Apparently, the right thing to do is to leave early. But the motive to see everything does not support leaving early, so Jack seems unable to perform the right act from the right motive. Were he to leave early, he would do so on the basis of some other motive set, perhaps involving a motive to bring about the best outcome. And this other motive set, by stipulation, would not be right because it has a worse outcome than being motivated solely to see everything. Therefore, Adams concludes, the right act conflicts with the right motive. Call this *normative conflict*.

Set aside delicate questions regarding how fine-grained a set of motives can be, such as whether Jack can be motivated solely to see everything throughout most of the visit yet switch to another motive precisely when it would be best to move along and leave early. Adams (in personal communication) has questioned the desirability of such an arrangement, and Derek Parfit (1984: 34) rejects its possibility. I suspect that any view of motives will face this dilemma: either the motives in questions are too fine-grained to be psychologically possible for humans (as Parfit claims), or else they are coarse enough to give rise to consequences independent of the acts they produce, which creates space for normative conflict (as Adams might have in mind). To understand the latter horn, suppose that Jack is travelling with a companion who does not mind leaving early but would be upset by even a whiff of consequentialist motivation. We may bear such a case in mind for our purposes – namely, a case in which it would be wrong for Jack to have another motive even for a moment.<sup>10</sup> In brief, making motives more fine-grained is not a quick fix.

Examples of apparent normative conflict are easy to come by. Another widely discussed example comes from Parfit:

Clare could either give her child some benefit, or give much greater benefits to some unfortunate stranger. Because she loves her child, she benefits him rather than the stranger. (Parfit 1984: 32)

Suppose that it is right for Clare's motive set to involve stronger love for her child than for a stranger, given the importance of love in a parent–child relationship. Still, the right act in this particular situation is to benefit the stranger rather than her child.<sup>11</sup> There is again normative conflict between right motive and right act.

A little more argument is needed to fully spell out the objection from normative conflict, though, since there is as yet no logical inconsistency in the fact that the right motive may conflict with the right act. We should distinguish normative conflict – as a label for cases like Jack's and Clare's – from the *objection* to direct consequentialism that is supposed to follow, since there is some disagreement over the best form for the objection to take.

Some philosophers explicitly put the objection as one of logical inconsistency.<sup>12</sup> The most common route to inconsistency requires two further premises. The first is a

<sup>10</sup>Louise (2006: 81) offers remarks in a similar vein.

<sup>11</sup>It is worth noting that, for his part, Parfit does not take normative conflict to be an objection – not a decisive one, anyway – against direct consequentialism. He instead focuses on difficulties that arise when basing the objection on the principle that 'ought' implies 'can,' which I mention below. See especially Parfit (1984: endnote 14 for Part I).

<sup>12</sup>Greaves (2020: 431–35) works through various ways one might argue for inconsistency.

principle of agglomeration that says that, if  $S$  ought to  $\varphi$  and  $S$  ought to  $\psi$ , then  $S$  ought to  $\varphi$  and  $\psi$ . Thus, if  $S$  ought to have the right motive set and  $S$  ought to perform the right act, then  $S$  ought to have the right motive set and perform the right act. The second premise is the principle that ‘ought’ implies ‘can’, meaning that if  $S$  ought to  $\varphi$ , then  $S$  can  $\varphi$ . Thus, if  $S$  ought to have the right motive set and perform the right act, then  $S$  can have the right motive set and perform the right act. But, according to the objection,  $S$  cannot do both in cases of normative conflict. Note, in passing, that each step of the argument has faced resistance: serious doubts have been raised, independently, regarding the agglomeration principle and the principle that ‘ought’ implies ‘can.’<sup>13</sup>

Perhaps in recognition of those doubts, Adams prefers to put the objection differently. For him, normative conflict exposes direct consequentialism to a lack of “moral seriousness.” Someone in Jack’s position (and we are all in Jack’s position from time to time) cannot seriously hold direct consequentialism as a moral theory for both motives and acts. Even were he to recognize that leaving is the right act, he does perfectly well in being motivated solely to see all of the cathedral instead. Moreover, he need not be troubled about his single-minded motive, given that direct consequentialism approves of it. But to be complacent about his motivation to do what he knows is wrong – and unmotivated to do what he knows is right – is to fail to take such notions seriously. Indeed, Adams claims, this is tantamount to failing to treat morality at all seriously.

Underlying both forms of the objection is the thought that moral theories are supposed to offer applicable ideals for comportment. Humans certainly make mistakes, but if a moral theory cannot even give a best-case scenario in which one conforms to each of its injunctions, then something has gone wrong. The charge of logical inconsistency asserts that direct consequentialism does not always offer such a best-case scenario. We can buy logical consistency by rejecting agglomeration or ‘ought’ implies ‘can,’ but only at the cost of admitting that it is impossible to fully live up to direct consequentialism. Alternatively, the issue of moral seriousness emphasizes that, even if it were possible for an agent to conform to each of consequentialism’s injunctions, they would not be doing so reflectively. They would not be moved by the theory; it would be practically useless.

The importance of applicability and achievability can be debated, but I will assume that they are desiderata (albeit defeasibly so) for any moral theory. Thus, if the objection from normative conflict can be met, so much the better for direct consequentialism. My responses will address the objection in each of its forms, for I aim to undermine the general idea that there ever must be conflict between right motives and right acts.

Before proceeding, in order to parse the objection, we must consider two understandings of the relation between motives and acts. One is that the agent’s motives at a time fully determine what they do at that time:

*psychological determinism*: holding fixed the laws of psychology,  $S$  could not act otherwise without  $S$  being motivated otherwise, i.e.,  $S$ ’s act supervenes on  $S$ ’s motive set.

Notice that psychological determinism is a claim about the relation between motives and acts, not determinism writ large. Moreover, it might characterize some situations,

<sup>13</sup>For expressions of these doubts, predating Adams (1976), see Williams (1973) regarding agglomeration, and see Stocker (1971) regarding ‘ought’ implies ‘can.’

but not others. The defense of consequentialism in regard to any given case depends on whether determinism or nondeterminism obtains for that case, so I will take each in turn. I will first consider cases of nondeterminism – i.e., cases in which agents can act contrary to their motives. For example, on this understanding of Jack’s case, it is possible for him to intentionally leave early even while holding his motives fixed. Then, I will consider cases of determinism. On this understanding, Jack’s motives determine that he stays late at the cathedral.

## 2. Assuming nondeterminism

In this section, assume psychological nondeterminism: that is, for any cases considered here, acts are not determined by motives. The assumption that these are nondeterministic cases does not entail that no cases are deterministic. Perhaps motives determine acts in certain cases – like that of a parent caring for her child – but not in others – like that of a sightseer on vacation. In this section I tackle whatever cases are in fact non-deterministic. I begin with the issue of logical consistency and continue on to that of moral seriousness.

The charge of normative conflict becomes slippery under nondeterminism. Jack’s motive to see everything is the right motive, and his act of staying late is the wrong act. Keeping in mind that he could have left early without a change in motivation, why think that staying late is anything other than a failure on his part?<sup>14</sup> The argument from logical inconsistency does not go through, for he can jointly have the right motive and perform the right act. It adds little to point out that the right motive *causes* him to act wrongly (where causation is nondeterministic) or makes it difficult to do otherwise.<sup>15</sup> Given that he could have acted rightly while having the right motive, the fact that he fails to act rightly does not look like a problem for direct consequentialism. It looks like simply a moral problem with Jack.

The tension between motives and acts is problematic, some think, in virtue of a special relation between the two. We say that motives rationalize certain acts, those to which they stand in favorable relations; this emerges in Adams’s point that the right motives can amount to a lack of moral seriousness toward the right act.<sup>16</sup> James Skidmore presses the point in this way (speaking of the Clare example):

While her motives did not rule [the right] act out entirely, they strongly inclined her not to perform it and ultimately constituted the *reason* she did not ... [The right] set of motives does not encourage, but rather *discourages* the agent from doing the right thing. (Skidmore 2018: 218; emphasis original)

In other words, a motive is not just a brute cause. Rather, it causes the act because the agent acts *on the basis of* (the content of) the motive. And when the agent knowingly performs the wrong act without sufficient motives to the contrary, as Jack does, he fails to treat rightness and wrongness with proper gravity. On this interpretation, the objection from normative conflict is that direct consequentialism can approve of a lackadaisical attitude regarding wrong acts in some cases.

<sup>14</sup>Gruzalski (1986: 773) and Mason (2002: 289) briefly argue to this effect.

<sup>15</sup>See, e.g., Skidmore (2018: 217). Shortly, I address the more challenging point, also raised by Skidmore, that motives rationally incline Jack toward the wrong act.

<sup>16</sup>I am grateful to an anonymous referee for urging me to focus on moral seriousness as the most challenging form of the objection in this section.

Being motivated to do something wrong does not on its own constitute a lack of moral seriousness, however. For example, being motivated to take photos of the cathedral does not preclude Jack from having concern for the cathedral's policy against cameras. Jack can still abide by the policy out of appropriate concern, which would defeat any presumption that he lacks moral seriousness. That is not to deny that he could refrain from photography for purely selfish reasons, in which case he does lack moral seriousness. I am just noting that he *could* refrain because he takes the rules seriously, wants to protect the artwork, or for some other reason suitably connected to rightness, despite being motivated to take photos.

To press the objection, one might be tempted to say that whether the agent takes morality seriously is settled by their motives. Jack is morally serious, for instance, only if his motives collectively favor the right act as such. (What is it for motives to collectively favor an act? Although I leave open the possibility of other accounts, a natural Humean position is that motives come in degrees of strength, such that they collectively favor the act with the greatest sum of the strengths of individual motives in its favor.) If Jack's motives collectively favor the wrong act, on the other hand, then he must not take morality seriously – or so the thought goes. Natural though the thought may be, it presupposes that all of the reasons for which Jack acts are captured in motives. Assuming that motives do not determine acts, as this section does, Jack can choose contrary to his motives. Even if his motives collectively favor the wrong act, he may yet do the right one. And so long as he does the right act intentionally, he may do so *for* some reason or other. Thus, the reason for which he acts need not correspond to his motives, which in the example count *against* leaving early.

If the reason for which he acts expresses a concern for morality, Jack cannot be accused of lacking moral seriousness. Views vary as to what kinds of reasons appropriately express moral concern (Adams and his interlocutors do not seem to come out in favor of any particular view). According to some, the reason should be explicitly about the act's rightness – an example of this kind of reason is that leaving early is the right thing to do – whereas on other views, the reason need only mention features of the act that make it right – such as the fact that the hours are dwindling for Jack to drive and find lodging.<sup>17</sup> Both support leaving early, so (in the absence of further considerations) Jack can do the right thing for either kind of moral reason.

Another issue left open by nondeterminism is what kinds of mental states or attitudes, if not motives, correspond to acting for a reason. Other candidates include states like belief, intention, and commitment – for example, a belief that leaving early is right, or an intention or commitment to do the right thing. However, we need not take a stand on these controversial questions in moral psychology. It suffices that, given nondeterminism, Jack can intentionally leave early despite his motives favoring staying. In particular, Jack can choose to leave early out of moral concern while being motivated solely to stay in order to see everything. Therefore, the right motives do not preclude moral seriousness.

In this section, I have argued that no problematic version of normative conflict arises in cases where motives do not determine acts. The problem of logical inconsistency and 'ought' implies 'can' disappears straightaway. One might instead worry that the lack of motivation to perform the right act constitutes a lack of moral seriousness. For reasons I explained above, that is an overly restrictive and tendentious view about what it takes

<sup>17</sup>In connection with moral seriousness and reasons, I am drawing on the literature about praiseworthiness and moral worth. For a helpful overview, see Singh (2020).



for an agent to exhibit moral seriousness. If it is possible that the agent intentionally performs the right act in spite of their motives, then the reasons for which they act may well exhibit moral seriousness. However, if all that seems implausible – if it seems that motives must ultimately be involved in order for the agent to act for a reason – I suggest that is because the reader assumes that acts must be determined by motives. That position will be addressed in the following sections.

### 3. Assuming determinism: Feldman's response

Assume that psychological determinism holds: acts do supervene on motives, so there is no way of acting otherwise while holding motives fixed. (Again, we may allow that this applies to only a subset of all cases.) If Jack's case is like this, it is impossible for Jack to have the right motive yet leave the cathedral early. On its face, the objection from normative conflict gets a stronger grip here. Conflict between right options that is truly unavoidable does deserve to be held against a moral theory.

Some philosophers think that the objection betrays a misunderstanding of consequentialism, however.<sup>18</sup> If a right motive determines the act that follows, then that act cannot be wrong, and Adams has misidentified the right and wrong acts for Jack. After all, built into the consequentialist formula for right acts is a restriction to *available* acts. Any act that is ruled out by the agent's motives (e.g., Jack's leaving early) cannot be right. Meanwhile, any act with no available alternatives (e.g., Jack's staying late) cannot be wrong.

Fred Feldman presents a version of this response to the objection. Although Feldman's account is both influential and instructive, I will argue that it is implausible because it implies that people are unable to perform many acts that clearly seem to be available to them. In so doing, I will set the stage for what I consider to be a more plausible determinist response in section 4.

Feldman develops his account in terms of "ruling out" accessible worlds. The worlds accessible to an agent at a time *t* are those possible worlds that correspond to the various ways of living available to the agent going forward from *t* (including motives they might have and acts they are able to perform).<sup>19</sup> A world is ruled out for the agent at *t* if it is not in the set of worlds accessible to them at *t*. Let the worlds be ranked according to overall value. This ranking provides an interpretation for the consequentialist criterion of "having the best outcome compared to every available alternative": for a motive or act to have the best outcome at *t* is just for the motive or act to occur in the best worlds accessible to the agent at *t*. Jack's motive to see everything, for instance, is the right motive because it occurs in the best worlds accessible to him throughout the example.

<sup>18</sup>In addition to Feldman (1993) and Dancy (1997), discussed below, see Gruzalski (1986: 771–77), Mason (2002: 290), and Streumer (2003: 242–44).

<sup>19</sup>The notion of accessibility can be interpreted to fit possibilist, actualist, or securitist views depending on how we spell out "ways of living available to the agent going forward from *t*." These may include all the possible ways in which the agent could live, or may be limited to the ways in which the agent would live given the choice in question, or may be limited to the ways that the agent could secure at *t*. Thus, the ruling-out response is neutral between possibilist, actualist, and securitist views, a point which is sometimes lost.

For instance, Dancy (1997) overlooks that point when he argues that leaving early is not the right thing for Jack to do on the grounds that its outcome *would* be worse. That explanation works only on the assumption of actualism.

On the other hand, Greaves (2020: 435) offers the ruling-out response only on behalf of possibilism. On behalf of actualism, she suggests denying agglomeration in order to avoid logical consistency. However, that does not avoid Adams's way of putting the objection in terms of moral seriousness (as discussed in section 1). My preferred way of articulating the ruling-out response is a more general response to the objection.



The accessible-worlds approach helps make sense of how an act might not be right despite appearing to be the best act available as of *t*. Consider, first, a less fraught example concerning future acts. Suppose that out of all of the noontime acts available to Jack, as of 9:00AM, the one with the best downstream consequences would be to eat lunch at his favorite creperie in Paris. Unfortunately, Paris is over an hour away from Chartres, and traveling back and forth for lunch would cut short his time at the cathedral. A naïve form of consequentialism might say that ordering lunch from the creperie at noon is the right act as of 9:00 since it is available and its consequences in a strict sense (i.e., those that are causally downstream) are better than those of any alternative. Intuitively, though, that is not the right act as of 9:00 because the best way for Jack to spend his day involves remaining near the cathedral, which rules out lunch at the Paris creperie. As of 9:00, the right thing to do at noon is instead, say, to order lunch from the bakery next to the cathedral. Feldman could explain this by reminding us that Jack remains near the cathedral in his best accessible worlds, and in all of those worlds he orders from the bakery. In addition, ordering lunch from the Parisian creperie is trivially not the right act as of 11:50 (supposing Jack remains in Chartres) because ordering lunch from the creperie at noon does not occur in *any* worlds accessible at 11:50. It simply is no longer an available option. There is no normative conflict here, neither as of 9:00 nor 11:50 nor any other time.

Similarly, as of 9:00, the best way for Jack to be motivated throughout the day is solely by the motive to see everything. Feldman argues that this motive

rules out all possible worlds in which he leaves the cathedral early ... at 2:00PM *there is no world accessible to Jack in which he leaves the cathedral early*. Hence, even in the best worlds accessible to him at 2:00PM, he stays too late. (Feldman 1993: 211; emphasis added)

Staying late in the afternoon is (and always was) the right thing for Jack to do. That is, Feldman argues that leaving early at 2:00PM is not right as of 9:00AM because Jack does not leave early in any of the best worlds accessible to him as of 9:00, in which he is motivated throughout the day to see everything. Additionally, given that he is motivated to see everything, leaving early is not right as of 2:00 because it is not even available as of 2:00. The analogy with the lunch options should be clear. An act may have the best downstream consequences yet at no point be right: at first, because it would be too costly for Jack to put himself in a position to perform the act, and later, because he has rightly comported himself in such a way that he is no longer able to perform the act.

But notice that this response relies on the assumption that Jack is *unable* to leave at 2:00. Psychological determinism says only that it is impossible for Jack to leave at 2:00 without his motives being otherwise; it does not trivially follow that he cannot leave at 2:00 *simpliciter*.<sup>20</sup> For that to follow, further assumptions would be needed. First, our account of ability to act would have to restrict Jack's ability to only those acts that

<sup>20</sup>I am not the first to point this out. Skidmore (2018: 215–17) has also noted this in our context. But Skidmore argues (mistakenly, as I will explain) that the ruling-out response works only if incompatibilism is true: "In short, Feldman, Dancy, and Mason are able to eliminate the alleged conflict in the cases of Jack and Clare only by reinterpreting them in such a way that the alleged optimal act that conflicts with the optimal motive ... is one that they cannot perform. This is accomplished by adopting an incompatibilist approach" (Skidmore 2018: 217). I'm grateful to an anonymous referee for pressing me to make clearer how my point extends beyond the existing literature.

supervene on available motives. Second, Jack would have to be unable to be motivated differently. Both assumptions can be questioned.

First, accounts of ability are extremely controversial, so Feldman's response suffers to the extent that it relies on a disputed account. For instance, a simple conditional account of ability regarding acts says that an agent is able to  $\varphi$  just in case the following holds: if the agent were to be decisively motivated to  $\varphi$ , they would  $\varphi$ . That conditional can be true regardless of whether determinism is true – regardless of whether it is psychologically possible, holding fixed how things in fact stand, for the agent to  $F$ . It is not obvious whether the ethically relevant sense of “available” is given by psychological possibility, a conditional analysis of ability, or some other account.<sup>21</sup>

Second, recall that psychological determinism is a claim about the relation between motives and acts, so it is silent on whether motives themselves are determined. *A fortiori*, psychological determinism is silent on whether agents have control over their motives. While I am not suggesting that motives are ever voluntarily chosen, there are resources for arguing that motives, like certain other attitudes, are subject to a rational form of control that is direct despite being non-voluntary.<sup>22</sup> Plausibly, the agent is actively involved in forming or eliminating motives; this activity is sensitive to reflection on the part of the agent; and the agent is appropriately held responsible for their motives. Because agents have leverage over actions via their motives, they would be able to act otherwise if they were able to be motivated otherwise. Although I cannot do full justice to these issues here, it is at least a viable position to hold that psychological determinism is compatible with ability to act otherwise.

More pointedly, for some critics, Feldman's incompatibilism is problematic because it commits us to saying that Jack is unable to leave at 2:00.<sup>23</sup> That sounds implausible for any ordinary sense of ability. (Contrast the observation that he *is* intuitively unable at 11:50 to order a noontime lunch from a creperie over fifty miles away.) Unless Jack is suffering from some unknown pathology, we should say that he can leave the cathedral immediately at any point throughout the day. To raise the stakes: Feldman's incompatibilism implies that agents do not act wrongly even when they act in vicious ways that seriously harm others.<sup>24</sup> This point is easily overlooked in Jack's case, where the consequences of staying late are minor inconveniences to himself. For a more extreme example, imagine that Cruella sees an opportunity to torture a lost puppy. Cruella's cruel motives determine that she will in fact torture the puppy, so per Feldman's approach, even if she possesses a normally functioning psychology, she cannot help but torture the puppy. And if that really is her only option, it cannot be wrong. Indeed, Feldman is committed to denying that anyone ever acts wrongly (on the supposition that psychological determinism holds). But many find this unacceptable. Insofar as we want to preserve everyday moral judgments, Feldman's ruling-out response goes too far.

<sup>21</sup>That is not to say that conditional accounts are without problems; Maier (2020) is a useful survey.

<sup>22</sup>See Smith (2015), Portmore (2019), and Achs (2020) for several different proposals to the effect that we enjoy direct control over motives (note that Smith and Portmore discuss desires, not motives, strictly speaking). Related work by Hieronymi (2006) on control over attitudes like belief and intention is also highly influential here. I discuss this further in McDougal (MS).

<sup>23</sup>In particular, see Louise (2006: 76–69) and Skidmore (2018: 215–17).

<sup>24</sup>As Skidmore (2018: 216–17) takes pains to emphasize.

#### 4. Assuming determinism: the compatibilist response

A response to the problem from normative conflict will be more attractive if it does not rely exclusively on incompatibilism – and I will argue that such a version of the ruling-out response is available. It is unfortunate that Feldman seems to suggest that incompatibilism is essential to his response, and that others have read him in that way.<sup>25</sup>

Suppose that psychological determinism is compatible with the ability to act otherwise, in order to see that the response survives. As I will demonstrate, all of the following may hold at once: (a) an act is available to perform, roughly straightaway, so long as certain immediate conditions are met; (b) the act's downstream consequences are better than those of any alternative; yet (c) the act is suboptimal since the conditions necessary for the act cost more than its downstream benefits. Although the act is ruled out of the best accessible worlds, that does not mean the act is at any point ruled out of *all* of the accessible worlds. Let me offer a couple of examples that involve acts and other focal points before returning to motives.

First, consider an example in which one act requires another as a necessary condition. Suppose that Deirdre is on a walk with her two young children, currently carrying her daughter on her hip with one hand and walking alongside her son. Above all, the right thing for Deirdre to do is to continuously carry her daughter because setting her down would result in a massive tantrum. (I make the moral demands on parents quite heavy in this example.) Her son also wants to be picked up and will be fussy until he gets his way, but not as fussy as his sister would be if she were set down for any time at all. The problem is that, in order to pick up her son and put him onto her other hip, she needs both hands, so continuously carrying her daughter rules out picking up her son. That is not to say that she is unable to pick him up – she can do so straightaway (roughly speaking). She only has to drop her daughter, not as a consequence but rather as a condition of picking up her son. Given the cost of setting her daughter down, picking up her son is intuitively not the right thing to do. This is explained on the accessible-worlds approach since the best worlds accessible to her are worlds in which she continuously carries her daughter, and in these worlds she does not pick up her son. Thus, each of the following holds: (a) Deirdre is able to carry her son; (b) carrying him has better downstream consequences than the alternative of not carrying him; yet (c) carrying him is suboptimal since the conditions necessary for doing so (setting down his sister) would cost more than the downstream benefits of carrying him.

Available acts may also require costly mental states, and not just motives. Suppose that Evelyn's teenage daughter Joy has just made an irreverent joke that amuses Evelyn, allowing a precious moment of levity between them. Evelyn could prevent an uncomfortable moment down the road with Joy's uptight grandfather by warning Joy not to repeat the joke around him – and the warning would do Joy no harm, we may suppose – so the downstream consequences of warning her are better than the consequences of not doing so. The problem is that Evelyn has a complex regarding Joy's grandfather, such that she becomes upset any time the thought of him influences her. That would be the case if she decided to warn her daughter, so remaining amused rules out any warning. That is not to say that she is unable to warn Joy. Nothing

<sup>25</sup>Perhaps Feldman did not mean to imply this, but he certainly gives the impression of relying on it (see the block quote above). Others, like Skidmore, share that reading of him: “[Feldman and Dancy] adopt an *incompatibilist* approach to the weak form of motive-act determinism on which the cases rely ... the source of this dispute lies in a deeper disagreement between a compatibilist and an incompatibilist approach” (2018: 215; emphasis original).

prevents a decision to warn her, and if she were to decide to do so, then she would do so straightaway. It is just that her amusement would be spoiled then. Note that her amusement would be spoiled by her *decision* to warn Joy, not by the act of warning itself. (If she warned her out of habit, she would not be upset, but that kind of habit is not present in Evelyn's case.) The decision is a necessary condition for intentionally warning Joy, so warning her is intuitively not the right thing for Evelyn to do given what it would cost. Once again, this is borne out on the accessible-worlds approach since the best worlds accessible to her are worlds in which her amusement lasts, and in these worlds she does not warn Joy. Each of the following holds, as before: (a) Evelyn is able to warn her daughter not to repeat the joke; (b) warning her has better downstream consequences than the alternative of not warning her; yet (c) warning her is suboptimal since the decision necessary for warning her would cost more than the benefits of warning her.<sup>26</sup>

Finally, return to Jack. Recall that it is right for him to be motivated solely to see everything, but that leaving early would have better downstream consequences than staying late. Psychological determinism says that the motive to see everything rules out leaving early, and as we have seen, Feldman concludes that staying late is right on the grounds that Jack is unable to leave early. Suppose, however, that determinism is compatible with the intuitive thought that Jack can leave early, as I have sought to motivate.<sup>27</sup> No external restraint or internal pathology is holding Jack back. Of course, if Jack were to decide to leave early, then he would be motivated differently – that much follows from determinism – so leaving early is ruled out given the sole motive to see everything. Again, leaving early would not cause another motive, but another motive is a necessary condition of leaving early. Given the cost of an additional motive, leaving early is not the right thing for Jack to do. The best worlds accessible to Jack are worlds in which he is motivated solely to see everything, and he does not leave early in any of these worlds (although there are other accessible worlds in which he does leave early). That is because the following hold: (a) Jack is able to leave early; (b) leaving early has better downstream consequences than the alternative of not leaving early; yet (c) leaving early is suboptimal since the motives necessary for leaving early would cost more than the benefits of doing so.

It might appear that motives are morally privileged, given that the intuitive judgment regarding the right act, that leaving early is right, is overruled by the judgment regarding the right motive. If that is so, my response risks being arbitrary and unsatisfying (especially in the eyes of a direct consequentialist). The response does not privilege

---

<sup>26</sup>It is natural to wonder at this point how much distance remains between indirect consequentialism and the form of consequentialism that I am advocating. One difference is that indirect consequentialists tend to assess their primary focal points, such as decisions, at the level of type rather than token (e.g., the right decision is the one that *typically* results in the best consequences even if the consequences are worse in this particular case). See, for instance, Hooker (2003: chapter 4.2). A second difference is that indirect consequentialism lacks grounds for denying that an act is right when it follows from the right decision. In contrast, I maintain that a suboptimal act is not right if a better act could have followed from the same decision. While the point is moot in the current context – where the act supervenes on the decision/motives – the distinction makes a difference when the relation between act and prior focal point is nondeterministic. I address related concerns later in this section. I am grateful to an anonymous referee for helping me to think through the relation to indirect consequentialism.

<sup>27</sup>In further support of compatibilism, compare the following pair of claims from Adams: “Being in the right frame of mind for enjoying the visit, he [Jack] could not bring himself to leave the choir screen”, yet “it is still in his power to leave the choir screen if he wants to” (Adams 1976: 472–73).

motives in any systematic way, however. Motives and acts alike are assessed according to their entire outcomes – namely, the worlds in which they occur. What gives the appearance that motives are privileged, then? Motives are causally upstream of acts in the examples, so the downstream consequences of acts are a subset of all of the downstream consequences of motives. That asymmetry is natural in virtue of the psychological connection from motives to acts. The consequences of an act are also consequences of any underlying motives. There could, of course, be cases where the relation is reversed. For instance, we could imagine that an agent should drink one of several motive-inducing potions, and that drinking certain potions also has additional consequences separate from the induced motive. In that case, the act of drinking would appear privileged instead of the induced motive. Such examples are harder to come by, though. There is no regular, tight connection from acts to motives like there is from motives to acts. Motives may be privileged with respect to agency, accordingly, but there is no in-principle privilege with respect to consequentialism.

Another related worry is that consequentialism is supposed to be concerned with *promotion* in the sense of promoting the good.<sup>28</sup> If the conditions for acting are not promoted by the act itself, then they are not to be counted against the act (so goes the claim). I am happy to accept the antecedent if it relies on an ordinary sense of promotion, according to which what is promoted must be causally downstream of the act.<sup>29</sup> However, we should not presuppose that consequentialism must be firmly tied to promotion in that sense. For one, the assessment of options according to accessible worlds, as advocated by consequentialists like Feldman, would be prematurely disqualified from consequentialism. And his approach arguably does just as much work with its simpler toolkit, unburdened by folk notions like causation. Moreover, the accessible-worlds approach certainly fits the consequentialist spirit that outcomes be as good as possible. Its difference from causal approaches, then, seems to me the subject of a dispute within consequentialism: how best to understand what it means for a focal point to “have” an outcome. Whereas the causal approach takes the relevant relation to be promotion, the accessible-worlds approach takes a broader relation that might be described as *realization*. Jack leaving early does not cause a worse outcome, but it necessarily realizes a worse outcome, which consequentialism may properly care about and track via moral assessment.

In summary: this section argued that the ruling-out response need not assume that Jack is unable to leave early. According to compatibilism, leaving early is simply sub-optimal, and Adams misidentified it as the right act. Furthermore, since the right motives favor what is in fact the right act, Jack need not suffer from a lack of moral seriousness. This response thus has the virtue of at once defeating both versions of the objection from normative conflict, both moral seriousness and logical inconsistency.

<sup>28</sup>Thanks to an anonymous referee for raising this challenge.

<sup>29</sup>It is worth noting that much recent literature adopts a more technical notion of promotion, one that does not assume causation is essential but instead analyzes promotion in terms of probabilifying outcomes. Although certain probabilistic accounts do claim that the act must *cause* the increase in probability, others explicitly reject that claim; for examples of each, see Lin (2018) and Sharadin and Dellsén (2019), respectively. It is commonly observed that instantiation and constitution are alternatives to causation that also count as forms of promotion precisely because the probability of the outcome is raised compared to some baseline. Likewise, the relation of requiring as a condition could also count as a form of promotion in an artificial sense.

## 5. Conclusion

Direct consequentialism survives the objection from normative conflict no matter what assumptions we make regarding agency. Assuming that Jack's motives do not determine how he acts, then he fails to act rightly by staying late at the cathedral. He could and should have left early even while being motivated to see everything. On the other hand, assuming his motives do determine how he acts, then leaving the cathedral early cannot be right. The best worlds accessible to Jack throughout the day are those in which he is motivated solely to see everything, and leaving early is ruled out in those worlds. Crucially, this response is neutral regarding Jack's ability to leave early. Perhaps, in virtue of his motives, Jack is unable to leave early, but the ruling-out response need not be so committed. More plausibly, Jack is able to leave early – there are worlds accessible to him in which he leaves early – but those are not his best accessible worlds. This completes the defense against normative conflict no matter where we end up in the logical space of agency: nondeterminism or determinism, and incompatibilism or compatibilism.

I have focused on defending the direct assessment of motives and acts. Beyond that, the responses considered here are promising for direct consequentialism as applied to other focal points, too. My example of Evelyn, which considers decision-making as a focal point, suggests a template for how the ruling-out response might generalize. In all cases, I want to say, the right option is whatever *realizes* the best outcome compared to its alternatives. And if the options are determined by certain other conditions, that may limit which outcomes are best without necessarily limiting what is an option for the agent.

Consequentialist theories, whatever their other faults may be, are not guilty of imposing conflicting requirements. That is not to say that consequentialist theories are easy to live up to. Living rightly is a tricky business, and no one who takes consequentialism seriously should be tempted to think otherwise. Sometimes, doing the right thing may have consequences that mislead the agent down the road. But consequentialism is not so difficult as to require the impossible – the best of all accessible worlds is good enough.

**Acknowledgments.** I have benefitted greatly from feedback from Jorah Dannenberg, Taylor Madigan, Barry Maguire, Leif Wenar, and audiences at Stanford University. I also thank two anonymous reviewers for their terrific reports as well as the journal editor for suggestions regarding the phrasing of several key claims.

**Competing Interests.** The author declares none.

## References

- Achs, Rachel. 2020. Blameworthiness and Constitutive Control. *Philosophical Studies* 177(12): 3695–715.
- Adams, Robert Merrihew. 1976. Motive Utilitarianism. *The Journal of Philosophy* 73(14): 467–81.
- Adams, Robert Merrihew. 1999. *Finite and Infinite Goods: A Framework for Ethics*. Oxford: Oxford University Press.
- Dancy, Jonathan. 1997. *Reading Parfit*. Oxford: Wiley-Blackwell.
- Feldman, Fred. 1993. On the Consistency of Act- and Motive-Utilitarianism: A Reply to Robert Adams. *Philosophical Studies* 70(2): 201–12.
- Greaves, Hilary. 2020. Global Consequentialism. In Douglas W. Portmore (ed.), *The Oxford Handbook of Consequentialism*, 423–40. Oxford: Oxford University Press.
- Gruzalski, Bart. 1986. Parfit's Impact on Utilitarianism. *Ethics* 96(4): 760–83.
- Hieronymi, Pamela. 2006. Controlling Attitudes. *Pacific Philosophical Quarterly* 87(1): 45–74.
- Hooker, Brad. 2003. *Ideal Code, Real World: A Rule-Consequentialist Theory of Morality*. Oxford: Oxford University Press.

- Kagan, Shelly.** 2000. Evaluative Focal Points. In Brad Hooker, Elinor Mason, and Dale E. Miller (eds.), *Morality, Rules and Consequences: A Critical Reader*, 134–55. Edinburgh: Edinburgh University Press.
- Lin, Eden.** 2018. Simple Probabilistic Promotion. *Philosophy and Phenomenological Research* 96(2): 360–79.
- Louise, Jennie.** 2006. Right Motive, Wrong Action: Direct Consequentialism and Evaluative Conflict. *Ethical Theory and Moral Practice* 9(1): 65–85.
- Maier, John.** 2020. Abilities. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/win2020/entries/abilities/>.
- Mason, Elinor.** 2002. Against Blameless Wrongdoing. *Ethical Theory and Moral Practice* 5(3): 287–303.
- McDougal, Austen.** MS. Against Motive Fatalism. Unpublished manuscript.
- Parfit, Derek.** 1984. *Reasons and Persons*. Oxford: Oxford University Press.
- Pettit, Philip, and Michael Smith.** 2000. Global Consequentialism. In Brad Hooker, Elinor Mason, and Dale E. Miller (eds.), *Morality, Rules and Consequences: A Critical Reader*, 121–33. Edinburgh: Edinburgh University Press.
- Portmore, Douglas W.** 2019. Control, Attitudes, and Accountability. In David Shoemaker (ed.), *Oxford Studies in Agency and Responsibility Volume 6*, 7–32. Oxford: Oxford University Press.
- Sharadin, Nathaniel, and Finnur Dellsén.** 2019. Promotion as Contrastive Increase in Expected Fit. *Philosophical Studies* 176(5): 1263–90.
- Singh, Keshav.** 2020. Moral Worth, Credit, and Non-Accidentality. In Mark Timmons (ed.), *Oxford Studies in Normative Ethics Volume 10*, 156–81. Oxford: Oxford University Press.
- Skidmore, James.** 2018. Does ‘Ought’ Imply ‘Might’? How (Not) to Resolve the Conflict between Act and Motive Utilitarianism. *Philosophia* 46(1): 207–21.
- Smith, Angela M.** 2015. Attitudes, Tracing, and Control. *Journal of Applied Philosophy* 32(2): 115–32.
- Stocker, Michael.** 1971. ‘Ought’ and ‘Can.’ *Australasian Journal of Philosophy* 49(3): 303–16.
- Streumer, Bart.** 2003. Can Consequentialism Cover Everything? *Utilitas* 15(2): 237–47.
- Thornley, Elliott.** 2022. Is Global Consequentialism More Expressive Than Act Consequentialism? *Analysis* 82(1): 75–84.
- Williams, Bernard.** 1973. Ethical Consistency. In Bernard Williams (ed.), *Problems of the Self: Philosophical Papers 1956–1972*, 166–86. Cambridge: Cambridge University Press.