

Mill and the Footnote on Davies

Christoph Schmidt-Petri

Published online: 14 August 2013
© Springer Science+Business Media Dordrecht 2013

1 Introduction

According to John Stuart Mill’s version of utilitarianism, “actions are right in proportion as they tend to promote happiness, wrong as they tend to produce the reverse of happiness.”¹ The second chapter of his *Utilitarianism* is devoted to rebutting various criticisms of this utilitarian theory. One of them is a variant of what today is called the ‘Demandingness Objection’. This objection, discussed in paragraph 19, is stated by Mill as follows: “They say it is exacting too much to require that people shall always act from the inducement of promoting the general interests of society...”²

Mill denies that this is a valid criticism of the utilitarian theory. He replies that utilitarianism does not in fact demand that people shall always act from the inducement of promoting the general interests of society, as the objectors claim. It only requires that they actually do promote the general interests of society. However, what motive induced the agent to perform an action is not part of the moral evaluation of that action, Mill claims:

... this is to mistake the very meaning of a standard of morals, and to confound the rule of action with the motive of it. It is the business of ethics to tell us what are our duties, or by what test we may know them; but no system of

¹ See *U II 2*, 55/210. References are by chapter and paragraph of *Utilitarianism*, the page numbers refer to the edition by Roger Crisp first (Roger Crisp, ed., John Stuart Mill, *Utilitarianism*, Oxford: Oxford University Press, 1998) and the *Collected Works* edition (John M. Robson, ed., *Collected Works of John Stuart Mill: Volume X: Essays on Ethics, Religion and Society*, Toronto: University of Toronto Press, London: Routledge and Kegan Paul, 1969) second.

² *U II 19*, 64/219.

ethics requires that the sole motive of all we do shall be a feeling of duty... the motive has nothing to do with the morality of the action, though much with the worth of the agent.³

So for the moral evaluation of an action, Mill seems to be saying, it is not important why it was performed. Mill tries to illustrate this point with several examples of which the following is of particular interest:

(Mill's example)

He who saves a fellow creature from drowning does what is morally right, whether his motive be duty or the hope of being paid for his trouble.⁴

This illustration of his theory failed to convince the Rev. J. Llewellyn Davies, who, according to Mill, objected to it in the following words:

(Davies's example)

Surely the rightness or wrongness of saving a man from drowning does depend very much upon the motive with which it is done. Suppose that a tyrant, when his enemy jumped into the sea to escape from him, saved him from drowning simply in order that he might inflict upon him more exquisite tortures, would it tend to clearness to speak of that rescue as 'a morally right action?'⁵

Mill does not think that this example shows that his theory is mistaken. Starting with the second edition of *Utilitarianism* as a book, published in 1864, Mill replies to Davies's expanded version of his example in a footnote, which unfortunately is not reprinted in all modern editions.

2 Mill's Reply to Davies

Mill's reply to Davies needs to be quoted in full. Mill writes:

- (1) I submit, that he who saves another from drowning in order to kill him by torture afterwards, does not differ only in motive from him who does the same thing from duty or benevolence; the act itself is different.
- (2) The rescue of the man is, in the case supposed, only the necessary first step of an act far more atrocious than leaving him to drown would have been.
- (3) Had Mr Davies said, "The rightness or wrongness of saving a man from drowning does depend very much" – not upon the motive, but – "upon the *intention*," no utilitarian would have differed from him.
- (4) Mr. Davies, by an oversight too common not to be quite venial, has in this case confounded the very different ideas of Motive and Intention.
- (5) There is no point which utilitarian thinkers (and Bentham pre-eminently) have taken more pains to illustrate than this.

³ U II 19, 64f/219.

⁴ Ibid.

⁵ As quoted by Mill in the footnote to U II 19, 65/219.

- (6) The morality of the action depends entirely upon the intention – that is, upon what the agent *wills to do*.
- (7) But the motive, that is, the feeling which makes him will so to do, when it makes no difference in the act, makes none in the morality: though it makes a great difference in our moral estimation of the agent...⁶

In this passage, Mill claims that Davies's example differs from his. Mill seems to be saying that in his own example the two actions at issue are in some sense identical whereas the action in Davies's 'tyrant' example is a third action which differs from those in his example. Hence Davies's example clearly cannot refute Mill's initial claim that two identical actions always have identical moral value, and that that moral value is independent of the agent's motive. Mill further adds that in Davies's example, in which, Mill obviously concurs, the tyrant's action does have a different moral value, the relevant difference is not to be found in any difference in the motive of the agent but instead in the intention of the agent. He goes on to say that Davies, though evidently not producing a good example to support his objection against utilitarianism, has therefore simply illustrated once again what utilitarians had been saying all along, namely that the moral value of actions depends on the intentions of agents, but not on the motive. In the end, Mill seems to believe, there is no substantive disagreement between himself and Davies. The problem is that Davies has misunderstood what Mill had said and hence mistakenly thought the tyrant's example was a counterexample to Mill's theory. But it's a misunderstanding rather than a counterexample, and Mill seems confident that he can explain where Davies went wrong.

However, Mill's reply nonetheless seems problematic, for several reasons. First, didn't Mill say just a few pages back that the moral value of an action depends on its effects? How then can it depend on the agent's intentions, as claimed in (6)? Even worse, how can it 'entirely' depend on them, as is also claimed in (6)? These seem to be plainly contradictory views of what determines the moral value of an action. Secondly, how is it possible that the motive may 'make a difference in the morality (of an action, presumably) if it makes a difference in the act', as apparently implied by (7), but nevertheless 'have nothing to do with the morality of an action', as previously claimed? That just seems very confused.

3 Expected Consequences and Praiseworthiness

One way of resolving the first problem, that is, the problem of squaring Mill's apparently contradictory assertions that the moral value of an action depends both on its effects and yet entirely on the agent's intentions, has been discussed by both Roger Crisp and Jonathan Dancy.⁷ The solution would be to reconcile the two apparently important concepts of the effects of actions and the intentions of agents

⁶ U II 19, footnote, 65/219f, numbering and all line breaks added, italics in original.

⁷ See Roger Crisp's editorial notes, "Notes to *Utilitarianism*" in Roger Crisp, ed., John Stuart Mill, *Utilitarianism*, Oxford: Oxford University Press, 1998, 111–150, pp. 121–123, and Jonathan Dancy, "Mill's Puzzling Footnote," *Utilitas* 12 (2000): 219–222.

by finding a place for each in the intended effects of actions. Two versions of this approach need to be distinguished:

A) On the first, Mill would be claiming that the moral value of an action does indeed depend on its effects, but only on some of them, namely on those intended by the agent. On this reading, those effects which do occur but were not intended are irrelevant, morally speaking.

B) On the second reading, the intended effects are not used to assess the moral value of an action, that is, whether it actually is right or wrong and to what degree, but its praiseworthiness, that is, whether it should be commended or condemned. These, though clearly related, are two different issues which Mill would be somewhat confused about.

A) The first reading is often called ‘expected consequence utilitarianism’. Abstractly speaking, this theory seems wholly implausible without some very serious modifications. It is implausible to claim that the moral value of an action merely depends on those of its consequences the agent has intended or foreseen rather than on all of them (assuming that it does depend on the consequences in some way).⁸ Surely we also need to consider what the agent ought to have intended or foreseen. After all, a negligent agent might only foresee a fraction of the consequences that actually occur and which a reasonable agent would have foreseen. But an action causing terrible suffering could not plausibly become less wrong just because the person performing it is careless at predicting its consequences or even culpably ignorant of what will happen as a result of his action – if anything, that makes matters even worse.

Now conceivably there might be a good reply to this objection, but if so it is not obvious at all how it would go.⁹ So if Mill actually held a version of expected consequence utilitarianism, we should expect him to say so clearly and discuss this very obvious objection too. Given that he doesn’t, we should resist ascribing this view to Mill, at least if a more plausible reading is available. I will argue below that there is.

B) On the second reading, the praiseworthiness but not the moral value of an action would be made dependent on the effects intended or foreseen by the agent.¹⁰ This is a *prima facie* plausible reading. It certainly seems relevant to our moral

⁸ Intention is the foresight of consequences, according to Mill in his comments to his father’s *Analysis of the Human Mind*, see John Stuart Mill, “James Mill’s Analysis of the Human Mind” in John M. Robson, ed., *Collected Works of John Stuart Mill: Volume XXXI: Miscellaneous Writings*, Toronto: University of Toronto Press, London: Routledge and Kegan Paul, 1989, 93–255, pp. 252f., as mentioned by Crisp, op. cit., p. 121, but see the further discussion of this below.

⁹ Expected consequence utilitarianism becomes somewhat more plausible if we calibrate it to the expectations of a ‘reasonable’ agent. But then an explanation of what might count as ‘reasonable’ in this sense would be of vital importance, and Mill is not offering any such explanation here. So it’s hardly plausible to ascribe that view to Mill either.

¹⁰ Praiseworthiness and moral value need not be identical for a utilitarian as it might sometimes promote the general happiness to, for instance, praise an unselfish sacrifice which, for some unfortunate reason, didn’t turn out to actually promote the general happiness. With hindsight, it might turn out that some other action would have had better consequences. Such an unsuccessful action might still be praiseworthy, as praising it might encourage others to perform similar yet successful actions, thus promoting the general happiness. The details would depend on the circumstances of the case.

assessment whether an agent intends to bring about, for instance, pleasure or pain. In the example at hand, the consequences of ‘killing by torture’ intended by the tyrant would clearly be much less desirable than those intended by the benign rescuer, so it would appear understandable if we scorned the tyrant’s action and praised the benign rescuer’s action. And we would do so even if by some unfortunate chain of events none of the intended effects were to materialise. Thus we can easily accept that the tyrant’s action seems so much less praiseworthy than the other, and how this seems to depend on the intended rather than the actual effects.

However, Mill’s previously quoted assertion that “He who saves a fellow creature from drowning does what is morally right, whether his motive be duty or the hope of being paid for his trouble”¹¹ seems to conflict with this reading. In this quotation from Mill, both actions, if praised at all, surely get praised equally. No distinction is made between them here. But this is a problem for the second reading because the selfish rescuer and the benign rescuer clearly expect different events to occur – we may plausibly assume that the selfish rescuer would not have saved the fellow creature from drowning if he hadn’t expected to actually be rewarded financially.¹² So these two rescuers do not just differ in motive but must also be presumed to differ in what they expect or intend to happen, that is, they must differ in intention. If Mill, in this example, was assessing the praiseworthiness of the actions by looking at what the agents expect to happen, he would have had to claim that these actions are not equally praiseworthy – and he seems to believe the exact opposite.

This approach founders by assuming that it is the praiseworthiness of the action that is assessed by looking at the intentions of the agent. Mill of course distinguishes between assessing agents and their actions. He agrees that both actions as well as agents may be assessed, and sometimes assessed differently; this is the very point of his own example.¹³ He clearly believes, as stated in (7), that our opinion of the agent’s motive is relevant for the moral estimation of the agent. So the motive we ascribe to the agent does influence our assessment of the agent (and maybe the praiseworthiness of the agent might indeed also depend on the agent’s intention). But this is very different from claiming, as the second reading does, that the praiseworthiness of the action depends on the agent’s intention. Actions and agents

¹¹ U II 19, 65/219.

¹² I think we may suppose that the selfish rescuer does not merely hope to be rewarded, though *ex hypothesi* he does that too. If he really is selfish, the hope itself would not be enough to motivate him to jump in the water. Because jumping in the water means he is, were he not to get a reward, definitely making himself worse off – he is certainly getting wet and might even himself risk his life; in any case he is not anyway jumping in the water for his own pleasure – he also needs to think there is a sufficiently high probability that his hopes become reality, that is, that he does get a reward which more than compensates him for his getting wet and taking the risk. It would be irrational – for a selfish person – to accept a certain loss in his own welfare if the probability of a reward is so low that realistically speaking, a reward may not be expected. So, if he is selfish and jumps in the water, he must expect to actually be rewarded. His hoping reflects the fact, to be discussed below, that he cannot be entirely sure to be rewarded.

¹³ I thank an anonymous referee for helping me clarify this issue.

need to be treated separately. Because Mill praises both actions equally, Mill's example is not compatible with the second reading either.

This means that neither of the two readings above, which make the moral value of an action or the praiseworthiness of an action (rather than of the agent) depend on the consequences intended by the agent, gives us a satisfactory reading of Mill's reply to Davies. Quite the contrary, working through them highlights even more difficulties. This approach to Mill's example seems to leave us with a new problem, the third: given that Mill's example is intended to illustrate that both the benign and the selfish rescuer do what is morally right, and granted that their intentions (in the sense of 'expectation') are not the same, how do we square this example with Mill's remark (6) in the footnote, that the intention does make a difference to the moral value of the action? Why then doesn't it do so in this example? After all, if the agents' intentions are different, as I have just argued, and the intention makes a difference to the moral value of the action, as Mill says, that difference in intention should make a difference to the moral value of these two actions even if, apparently, it makes no difference to the actions' praiseworthiness.

4 Complex Actions

Jonathan Dancy suggests a different interpretation.¹⁴ Of crucial importance is his emphasis on Mill's statement in (2) above: Mill claims that the tyrant's rescue is not even the same action as the rescue by the benign rescuer, as Davies mistakenly assumes. The tyrant's rescue is not a single complete action at all but only part of a more complex action. As Dancy correctly highlights, Mill says the rescue is a "necessary first step of an act" in (2).¹⁵ This illustrates why, just as Mill claims, Davies has not managed to produce an example in which two identical actions which however differ in motive are performed. As a result, we are under no pressure to explain the apparent and undeniable difference in their moral value by appealing to the difference in the agents' motives, as Davies would have it. The motives are indeed different, just as Davies says, but so are the actions. But with what justification does Mill claim that the actions are different? How are actions individuated, according to Mill?

The reason why the two actions are different, Dancy contends, might be due to the fact that the agents also have different intentions. And this would explain how the moral value of an action depends on the agent's intention, as Mill had claimed. The tyrant may plausibly be taken to intend three things, namely (to put it roughly) 'rescue + torture + killing'. By contrast, the benign rescuer merely intends a single 'rescue'. So if "the identity of the act is fixed by the intention with which one does it", as Dancy hypothesizes, the tyrant's and the benign rescuer's actions are certainly two very different kinds of actions.¹⁶ The latter is a rescue, the former is what may be called, using legal terminology, a rescue 'in coincidence with' or 'in

¹⁴ See Dancy, *op cit.*

¹⁵ *Op. cit.*, p. 221.

¹⁶ *Loc. cit.*

unity of act with' torture and murder. This unity of act of the tyrant's complex action would then render it impossible to evaluate the rescue-component of the tyrant's action separately. As the components of this complex action are inherently intertwined, we are not able to compare what we may loosely call 'the tyrant's rescue' with the 'benign rescue'. We are de facto always comparing an actual 'rescue' (by the benign rescuer) with a 'rescue in unity of act with torture and murder' (by the tyrant). It is unproblematic to agree that if these actions, the latter one being a complex one consisting of several components, are actually the actions being compared, they should not have the same moral value.

Though I believe this interpretation is on the right track, it is difficult to consider it entirely adequate.¹⁷ Even though this interpretation seems to explain the difference in our assessments of the tyrant vs. the benign rescuer (Davies's example), it seems to have some problems with our assessment of the benign vs. the selfish rescuer (Mill's example). Dancy's interpretation faces a problem similar to the one discussed in the previous section: Isn't it also the case that the selfish rescuer intends to be rewarded for his trouble whereas the benign rescuer has no such intention? If so, and if intentions determine the identity of an action, as they do on Dancy's reading, the benign and the selfish actions are also not identical – the first one is a simple 'rescue' and the other a complex one, namely a 'rescue + get rewarded' (or something similar). But it would be very odd for Mill to offer as a solution to Davies's example a move that implies that his own example shows the exact same defect he attributes to Davies's, namely, that the identity of actions has not been distinguished carefully enough. Why doesn't the very same objection apply to his own example? On that theory, his example too mistakenly equates two different actions.

In what follows, I will slightly modify Dancy's approach so that this problem doesn't arise. For the moment, it suffices to accept that neither of the approaches considered by Crisp and Dancy is entirely satisfactory.

5 Actions and Their Consequences

In this and the next two sections, I present my alternative to these readings. Let's start from the possibly simplistic working hypothesis that the moral value of an action depends on its actual effects only. Let's call it 'simple act utilitarianism'. Do any of Mill's statements under consideration contradict this simple act utilitarianism? To find out, we need to examine the moral evaluation of the three types of rescue (benign, selfish, tyrant) implied by Mill, find out what role intentions play in the determination of the moral value, and, finally, how exactly motives matter, if at all.

It seems clear that the two initial actions, the selfish and the benign rescue, are taken to have the same moral value by Mill. This is compatible with simple act utilitarianism because the actual consequences of the two rescue actions are the same, as I will now argue.

¹⁷ I should add that Dancy himself seems sceptical whether Mill would have endorsed this interpretation, but Dancy's reasons for being sceptical are different from mine.

Are the consequences of the two actions really the same? One could object that I have just argued, in the previous section, that at least the two agents don't think so. But the fact that the selfish rescuer intends and expects to receive a reward for his troubles doesn't make any difference to the actual consequences of his action. His intention (or hope or expectation) by itself clearly doesn't make the reward any likelier and certainly doesn't cause it to come about. Whether he gets a reward at all depends first and foremost on whoever might want and also be in the position to reward the rescue.¹⁸ Though the rescue is a necessary condition for being rewarded for it, it is so far from being sufficient that it seems entirely natural not to call any reward an effect of the rescuer's action. In any case, if someone were to want to reward the selfish rescuer, that person would certainly also want to reward the benign rescuer. So, *ceteris paribus*, either both the benign and the selfish rescuer get a reward or neither of them does. Whether or not the selfish rescuer gets a reward, then, is no more a consequence of his rescue action than it would be a consequence of the rescue action of the benign rescuer. As far as the consequences of these two rescue actions are concerned, they are certainly identical, whether they get the reward or not.

But the same can hardly be said of the tyrant's action. Though up to a certain point in time the sequence of events constituting the rescue of the drowning person will look indistinguishable to an onlooker, the events ultimately unfold very differently. Only in the selfish and benign cases will the rescued victim derive a net benefit from having been rescued. In these two cases, his life may be assumed to go on as before, and as far as we can tell, that life was better than death by drowning. In the tyrant's case, on the other hand, this doesn't quite seem to be the case. The victim may well have preferred to drown rather than to first get 'rescued' and then tortured and killed somewhat later. Mill clearly assumes that drowning is preferable to being 'rescued', tortured and killed when he says, in (2), that "The rescue of the man is, in the case supposed, only the necessary first step of an act far more atrocious than leaving him to drown would have been."¹⁹ So in the tyrant's case, the 'rescue' doesn't improve the 'rescued' person's life in the long run, quite the contrary, from an outside perspective, even drowning appears preferable to what he will be going through. This makes Mill's claim that the tyrant is performing a completely different action sound rather plausible.

But we still don't know how exactly Mill is individuating actions here. Mill says in the *System of Logic* that actions are individuated by the intentions of the agent and the action's consequences:

Now what is an action? Not one thing, but a series of two things: the state of mind called a volition, followed by an effect. The volition or intention to produce the effect, is one thing; the effect produced in consequence of the intention, is another thing; the two together constitute the action.²⁰

¹⁸ Because the example wouldn't make much sense otherwise, I assume that the rescue is not legally obligatory and also carries no legally due reward.

¹⁹ U II 19, footnote, 65/219.

²⁰ In John M. Robson, ed., *Collected Works of John Stuart Mill: Volume VII: A System of Logic, Ratiocinative and Inductive, Books I–III*, Toronto: University of Toronto Press, London: Routledge and Kegan Paul, 1974, at p. 55.

If an action is individuated by the intention of the agent and the action's consequences, actions may differ because their consequences differ. And it is entirely plausible that in Davies's example the difference between the two actions might be due to the differences in their consequences – and not, as Dancy had claimed, appealing to the idea of complex actions, to the difference in the agent's intentions. As just shown, the consequences of the tyrant's 'rescue' action are very different from those of the other two rescue actions. That difference in consequences certainly suffices to justify Mill's claim, in (1), that the tyrant's action is different from those of the benign and selfish rescuers.

So Dancy rightly emphasizes Mill's claim that the tyrant's action is different. But this claim doesn't compel us to stipulate the existence of complex actions. If both intentions and consequences of an action determine its identity, there is no need to appeal to the tyrant's amalgamating intention to explain that difference in identity which might simply be due to the difference in the action's consequences.²¹ And if we are not thus compelled to assume that this amalgamation is taking place, we can just as well assume that Mill is talking not about the complex action consisting of rescue + torture + killing, but only about the 'rescue', even in the tyrant's case. That makes all three 'rescue' actions comparable again.

As far as the moral evaluation of the three actions is concerned, the situation then is as follows. The benign and selfish rescue actions have the same consequences: continued existence of the rescued person, promoting that person's happiness (compared to letting him drown). Hence morally speaking, they are both equally right. The potential reward, given that it is not a consequence of either rescue action, is neither here nor there for the morality of these actions. The 'rescue' action of the tyrant has somewhat different consequences though, namely suffering caused by torture and subsequent death, clearly decreasing the 'rescued' person's happiness. This difference in consequences both turns the tyrant's action into a different action than those of the other rescuers – I found it necessary to call them 'rescue actions' and the tyrant's a 'rescue' action (with inverted commas) – and explains why that action, given that its impact on the overall happiness is negative, is wrong rather than right. So far, simple act utilitarianism is coherent with Mill's claims.

6 Intentions and Actions

Nothing has been said about the first problem yet (the relevance intentions have for the moral value of actions, according to Mill, which might appear problematic given his claim that the effects matter). The problem arose because Mill had emphatically asserted that the "morality of the action depends entirely upon the intention."²² The previous section might be thought to be in denial about this assertion as it seems to have been arguing that Mill believed the morality of an action to depend on its actual consequences only, and this might seem to be entirely disregarding the

²¹ As will become clear, the tyrant's intentions still play an important role. My view differs from Dancy's mostly in that the tyrant's intention is not taken to create a complex action.

²² *U II 19*, footnote, 65/219.

intention of the agent. So far the argument also ignores what above I have called the third problem, namely, how to square Mill's example of the benign and selfish rescue actions (which seemed to display different intentions but have now been claimed to have identical moral value), with this proclaimed relevance of intentions for the moral value of actions.

In (6) above, Mill seems to be equating the 'intention' of an action with what the agent 'wills to do'. That is quite different from the interpretation Crisp relies on. Crisp appeals to Mill's commentary to his father's *Analysis of the Human Mind*, where Mill writes:

Intention, when we are said to intend the consequences of our actions, means the foresight, or expectation of those consequences, which is a totally different thing from desiring them.²³

This passage clearly clashes with Mill's description of intention in (6), where he says: "The morality of the action depends entirely upon the intention – that is, upon what the agent *wills to do*."²⁴ It's not clear which of the two interpretations Mill 'really' believed in. In the absence of further evidence either way I think it is reasonable to use the reading to be found in *Utilitarianism* to explain the passage in *Utilitarianism* at issue, if that would allow us to reach a coherent interpretation – even if not of Mill's views in general, which might be incoherent, but only of his views in *Utilitarianism* (which so far still appear incoherent too).²⁵ So let's accept it for the purpose of the argument.

Using this conception of intention, we may somewhat rhetorically rephrase the question thus: How does the morality of an action depend on what the agent 'wills to do' (or wants to bring about)? Now the answer seems obvious: in general, what the agent wants to bring about determines what happens as a result of the agent's actions. And what happens is what the moral evaluation of actions is all about. So it is clear how, in some sense, the moral value depends 'entirely' on the intention of the agent in the examples: If the benign rescuer pulling the drowning person out of the pond wants to save him, the person will be saved. If the tyrant wants to torture and kill him, he will be tortured and killed. Clearly, the difference in intention brings about a difference in the effects. It is also the only reason required to explain why the effects differ. Thus the effects depend 'entirely' on the intention of the agent: they would have been different if the intention had been different, as Davies's example aptly illustrates (albeit confusedly, as Mill points out).²⁶ If the moral value

²³ Mill, op. cit. 252f., italics in original.

²⁴ U II 19, footnote, 65/220, italics in original.

²⁵ For a related argument about the relevance of Mill's comments to his father's book, see Christoph Schmidt-Petri, "On an Interpretation of Mill's Qualitative Utilitarianism," *Prolegomena* 5 (2006): 165–177.

²⁶ The preceding only holds in general, unfortunately. Other people or events might interfere with the outcomes of an action or the agent might be unable for other reasons to bring about the effects he intended. This is an important issue in its own right and it is not clear what Mill would say about it. Importantly, however, it doesn't seem to be the issue between Mill and Davies. Both Mill and Davies seem to assume that in the cases under discussion, people manage to bring about what they want to bring about. For purposes of clarifying the overall structure of his theory, focusing on clear cases is a legitimate strategy for both Mill and Davies. His objection does not turn on intentions and effects coming apart. It

of an action is determined by its consequences, and the consequences depend entirely on the intention of the agent, the moral value also depends ‘entirely’ on the intention of the agent. The first problem is solved.

So is the third problem: The benign and the selfish rescuer do not differ in intention in this sense. Though they do differ in what they expect (the sense of ‘intention’ used by Crisp) or hope to happen, they do not differ, as argued in the preceding section, in what they want to bring about nor in what happens as a consequence of their action. If the intentions of these two rescuers do not differ, and nor do the consequences, then they must be performing the same action, on Mill’s theory of action. Consequently it would be correct to claim that these actions have the same moral value even if it is also true that the moral value of actions depends entirely on the agent’s intentions. It does indeed, via the production of effects, but because the intentions and effects are identical in this case so is the moral value of the two actions.

7 Motives and Character

If the preceding reading is acceptable, Davies’s example has been shown to be a non-starter. The tyrant’s action is morally worse than the others because its consequences are worse, and the consequences come about because the tyrant wants to bring them about. So Davies is entirely right with his observation motivating the objection – that it would not tend to clearness to call the tyrant’s action ‘morally right’ – but that’s because, unlike the other two cases, it’s not actually morally right at all, it’s wrong. There is no need to invoke the tyrant’s motive to agree to this obvious fact, which Mill would never have wanted to deny.

The differences to the benign rescue case couldn’t be more striking. But the difficulty with Mill’s original example which Davies’s unhelpful exaggeration was intended to bring to the fore still persists. Davies, after all, is putting his finger on a genuine problem for Mill’s theory. Mill, in order not to be stuck with having to say, implausibly, that there is literally no difference between what is happening in the benign rescue case and the selfish rescue case, says the difference does not lie in the action or its moral worth but in the motive of the agent. The motive, according to (7), is the “feeling which makes him will to do” something, and “makes a great difference in our moral estimation of the agent” having “nothing to do with the morality of the action, though much with the worth of the agent.”²⁷ The idea seems to be that when we compare the selfish rescue and the benign rescue, as far as the moral evaluation of the action is concerned, we ultimately consider only the consequences. But any action is performed by an agent, and if we know the motive of the agent, that is, the feeling that motivates him to do what he does – for instance,

Footnote 26 continued

would be uncharitable to infer from this that neither Mill nor Davies realised that not all actions turn out as intended. Because I am also merely clarifying the overall structure of Mill’s theory, it would be equally uncharitable to claim that the interpretation of the footnote I am offering is intended to suggest that Mill’s theory is exhaustively described as simple act utilitarian in this sense.

²⁷ *U* II 19, footnote, 65/219f and *U* II 19, 65/219.

benevolence or selfishness – we are also in the position to assess the character of the agent. And that is an entirely separate but maybe equally interesting issue.

Probably for good reason, Mill doesn't specify how precisely this character evaluation should take place, but he does indicate that it is relevant that someone might have "a good or bad habitual *disposition* – a bent of character from which useful, or from which hurtful actions are likely to arise."²⁸ So it seems character traits or behavioural dispositions are assessed by looking at what kind of actions they would typically give rise to and the extent to which these actions promote the general happiness. For instance, selfishness will only give rise to rescue actions if a reward is to be expected whereas benevolence will plausibly give rise to rescue actions even without the expectation of such a reward. So benevolence will, if rescue actions were regularly necessary, contribute more steadily to the promotion of the general happiness than selfishness. So at least in such scenarios, we hope for benevolence rather than selfishness. The general idea is clear enough.

But this way of integrating the agents' motives into the picture invites the following objection.²⁹ If intentions are morally relevant to the assessment of actions because they give rise to effects which undoubtedly are relevant, then why are not motives also relevant to the assessment of actions, given that they give rise to actions and thus to effects too? Motives seem just one step further away from the effects than intentions. Mill's assertion that intentions are relevant yet motives are not then seems highly arbitrary. It's absolutely undeniable that both are.

The answer to this further problem is simply to admit that motives are also relevant to the moral value of actions in some sense, but to claim that they are relevant in a sense that is different from the sense in which intentions are relevant to the moral value of actions. And the sense in which motives are relevant is, unlike the sense in which intentions are relevant, not one which allows one to say that the moral value of actions depends on the motives of the agent. This needs further explanation.

It's clear that motives do make some difference a utilitarian like Mill should care about because they influence which actions are performed.³⁰ They are relevant to the moral value of actions because, as a general rule, which actions are performed depends on them. In this sense, it is fair to say that motives are highly relevant to the moral value of actions. We just need to hypothetically compare a world full of benevolent people with one full of selfish people.

But it's also clear that motives do not necessarily make a difference to the consequences of the actions that do get performed. Mill's example illustrates precisely this scenario. The selfish and the benign rescue, we had seen, are brought about by different motives. However, they are nonetheless identical actions: neither is there a difference in intention nor in consequences, which, according to the intention/consequence-conception of action mentioned in the *System of Logic* are the features that individuate actions; consequently, nor is there a difference in their respective moral value. In this sense, motives are not relevant to the moral value of

²⁸ U II 19, footnote, 65/219f, italics in original.

²⁹ Cf. Crisp, op. cit., pp. 122f.

³⁰ Cf. Dancy, op. cit., p. 222.

actions. We might prefer benevolent people over selfish ones because they will rescue people even in some circumstances when selfish ones won't, but if the circumstances are such that both do rescue, the good effects brought about are independent of their motives.

Motives, then, may make a difference to what action is chosen (as illustrated in Davies's example) but any particular action may be chosen out of various motives (as illustrated in Mill's example). As a result, it would not be accurate to say that if the motive of an action had been different, its effects would have been different too. Clearly that's not true generally speaking, as Mill's example illustrates, even if it is true in some sense in Davies's example – only in some sense, though, as here, because the tyrant has other intentions, the action itself also turns out to be different. And so that's the sense in which the moral value of an action does not depend on the motive with which it is done: It is possible to vary the motive which causes the action to take place without varying the action itself and hence without varying its moral value. But it is not possible to vary the intention without varying the moral value of an action because the intention will determine the effects which in turn determine the moral value.

To put it differently: some motive is required for any action, whether good or bad, to come about. But for the moral value of the action that does come about, it doesn't matter what motive it was that made it come about. What matters, morally speaking, are the action's consequences. If these are different because the motive gives rise to a different intention, the moral value will be different – but so will be the action.³¹ This is the solution to the second problem.

8 Conclusion

There is a view compatible with everything Mill says in these passages that can deal with all three problems. It's a simple act utilitarianism in which the moral value of an action is determined by its actual consequences. On this view, the consequences of an action, what happens, depends on what the agent wants to bring about, that is to say, they depend on the agent's intentions. Therefore the moral value of an action depends, given that on simple act utilitarianism it entirely depends on the action's consequences, entirely on the intentions of the agent – at least in the simple cases at issue between Mill and Davies in which people manage to bring about what they want to bring about. In these cases, varying the intention varies the consequences and thus the moral value. The simple act utilitarian nevertheless recognises that motives are also relevant – not necessarily to the consequences of the actions that do get performed but to which actions get performed at all. Hence motives are of paramount importance in theory and practice. However, as actions are individuated by the agent's intentions and the action's consequences, without appealing to the motive giving rise to it, any particular action might be brought about by different

³¹ The fact that the actions are different not because the tyrant's motive is different (though plausibly that too is different since apparently it is neither benevolence nor selfishness) but because his intentions and hence the consequences of his action are different is what Davies seems not to realise. Davies's example is one where the motive "makes a difference in the act" as mentioned in (7), *U* II 19, footnote, 65/219.

motives. Mill's example illustrates this scenario. Hence it is possible to vary the motive due to which an action is performed without varying the action's consequences, that is, without varying its moral value. In this sense, the moral value of an action does not depend on the motive with which it is done. Hence Davies's objection is indeed unfounded and Mill's footnote is rescued.³²

³² I want to thank Luc Bovens, an anonymous referee, and audiences in London and Regensburg for help with this paper.