

Mill's act-utilitarian interpreters on *Utilitarianism* chapter V paragraph 14

Dale E. Miller 

Philosophy and Religious Studies, Old Dominion University, Norfolk, VA, USA

ABSTRACT

In the fourteenth paragraph of the fifth chapter of *Utilitarianism*, J. S. Mill writes that 'We do not call anything wrong, unless we mean to imply that a person ought to be punished in some way or other for doing it; if not by law, by the opinion of his fellow-creatures; if not by opinion, by the reproaches of his own conscience.' I criticize the attempts of three commentators who have recently presented act-utilitarian readings of Mill – Roger Crisp, David Brink, and Piers Norris Turner – to accommodate this passage.

ARTICLE HISTORY Received 14 November 2016; Accepted 23 January 2017

KEYWORDS J. S. Mill; act utilitarianism; rule utilitarianism; sanction utilitarianism; blame; guilt

1. Introduction

The question of what moral theory John Stuart Mill holds is keenly debated. While I have previously made positive contributions to this debate (see especially 2010, 79–110), my present contribution will primarily be a negative one, since I shall offer some criticisms of the work of three interpreters who have recently put forward act-utilitarian readings of Mill: Roger Crisp, David Brink, and Piers Norris Turner. My criticisms will focus on something very specific, namely their attempts to make sense of one passage, paragraph fourteen of chapter five of *Utilitarianism* (which I will hereafter call V14). The relevant portions of this lengthy paragraph run as follows:

For the truth is, that the idea of penal sanction, which is the essence of law, enters not only into the conception of injustice, but into that of any kind of wrong. We do not call anything wrong, unless we mean to imply that a person ought to be punished in some way or other for doing it; if not by law, by the opinion of his fellow-creatures; if not by opinion, by the reproaches of his own conscience. This seems the real turning point of the distinction between morality and simple

CONTACT Dale E. Miller  demiller@odu.edu

© 2017 Canadian Journal of Philosophy

expediency. It is a part of the notion of Duty in every one of its forms, that a person may rightfully be compelled to fulfil it. Duty is a thing which may be exacted from a person, as one exacts a debt. ... Reasons of prudence, or the interest of other people, may militate against actually exacting it; but the person himself, it is clearly understood, would not be entitled to complain. There are other things, on the contrary, which we wish that people should do, which we like or admire them for doing, perhaps dislike or despise them for not doing, but yet admit that they are not bound to do; it is not a case of moral obligation; we do not blame them, that is, we do not think that they are proper objects of punishment ... I think there is no doubt that this distinction lies at the bottom of the notions of right and wrong; that we call any conduct wrong, or employ, instead, some other term of dislike or disparagement, according as we think that the person ought, or ought not, to be punished for it; and we say, it would be right, to do so and so, or merely that it would be desirable or laudable, according as we would wish to see the person whom it concerns, compelled, or only persuaded and exhorted, to act in that manner. ([1861] 1969, 246)

I will be asking two questions about the act-utilitarian readings of Mill that I consider: first, whether they are consistent with this passage, and, second, whether we can adopt them without attributing unnecessarily implausible or otherwise unappealing positions to Mill. In each case, I will show that one or both questions must be answered in the negative.

Admittedly, this might not by itself establish that the act-utilitarian reading of Mill must be abandoned. While Mill has much to say that bears on the choice between moral theories, different remarks seem to point in different directions and some seem to be outright contradictory. Any interpretation will therefore fit awkwardly with certain passages, impute a view to him that is unpalatable in some respect or other, or both. The best that those of us who write on his moral philosophy can realistically hope to accomplish is to show that our favored reading has fewer and smaller imperfections than the alternatives. Still, V14 does seem to be particularly important; it appears in the essay that is generally regarded as our best source for understanding the mature Mill's views on ethics, and the length at which he dilates on the point that he makes therein suggests that he considers it especially significant. Even Daniel Jacobson, who contends that Mill's interpreters tend to attach far too much importance to *Utilitarianism*, takes V14 to be the key to understanding Mill's moral philosophy.¹ So if Mill's act-utilitarian interpreters should prove unable to accommodate the passage adequately this would be a major strike against their reading, if not by itself a fatal flaw.

2. Lyons on V14

Before turning to Mill's act-utilitarian interpreters, I will first say a bit about the work of one who rejects this way of reading Mill and does so largely on the strength of V14. This is David Lyons, whose influential account of what Mill is doing in V14 is worked out in a series of papers that began to appear in the

mid-1970s. I start here because Crisp, Brink, and Turner are all reacting to Lyons's account, and in fact they all accept much of it.

As Lyons reads V14, Mill here presents a metaethical analysis of the concepts of morally wrong action and moral duty or obligation in terms of the justifiability of punishment. More specifically, and more distinctively, Lyons takes Mill to analyze these concepts in terms of the justifiability of one particular variety of punishment, namely the 'internal sanction' that is imposed by the agent's own conscience in the form of guilt; guilt is a warranted or justified response to immoral behavior even when the other kinds of punishment which Mill mentions are not.

When Mill traces a connection between 'the idea of penal sanction' and the concept of wrong conduct, he clearly stretches the former to cover not just external threats and penalties of an informal, extralegal nature, but even guilt feelings and self-reproach ... Furthermore, Mill suggests that self-reproach is not just one among a number of alternative forms of 'punishment,' the justification of which is connected with the idea of wrong conduct, but rather that it is the minimal, essential sort of 'punishment' so linked with the idea of immorality. Mill says that external sanctions may or may not be justified for wrong conduct, but that guilt feelings are always warranted when one acts wrongly. ([1982] 1994, 129)

In consequence, Lyons writes, 'wrongness is conceptually connected with justified guilt feelings, but only contingently or synthetically connected with external sanctions' ([1982] 1994, 130).

As I noted previously, Lyons's reading of V14 is a crucial premise in his argument for the conclusion that Mill is something other than an act utilitarian. For Mill, Lyons maintains, sanctions 'assume coercive rules' ([1976] 1994, 55). This may appear to be true only of legal sanctions. However, on Lyons's reading, Mill understands the conscience as an enforcer of rules; he thinks that we experience guilt when we have violated a rule that we have 'internalized,' or at least when we perceive ourselves to have done so. So the internal sanction presupposes coercive rules, too. And so do the informal social sanctions, at least insofar as they are expressions of blame, since typically we blame others for transgressing more or less the same coercive social rules that we feel guilt for breaking ourselves. Blame and guilt do differ in an important respect, however: while we can choose not to act on our feelings of blame toward someone else, and so refrain from imposing external sanctions on her, our feelings of guilt 'operate more or less automatically' ([1976] 1994, 56). While there will be some variations between individuals, Lyons takes Mill to believe that by and large people who live in the same social milieu will internalize the same rules. So to say that an agent should feel guilt for a given action is to say that there should be what Lyons calls a 'coercive social rule' that prohibits the action, that is, that the members of the agent's society should internalize such a rule.

So far nothing about the position that Lyons is attributing to Mill is distinctively utilitarian. However, on Lyons's reading Mill applies a utilitarian criterion to determine what coercive rules should be current in a society. A society is

warranted in having a particular moral code, which is to say a particular set of coercive rules, just if this would yield more happiness than would its having some other code instead. This is Bentham's theory of punishment, which Lyons reminds is one of the parts of Bentham's thought that Mill finds most satisfactory, applied to coercive social rules ([1976] 1994, 52; see [1833] 1969, 11). For an action to be wrong, therefore, is for it to be prohibited by the moral code which it would be optimal for people to internalize. While guilt is the only form of punishment that is conceptually linked to immorality, on this reading of Mill, other forms of punishment – informal social sanctions or formal legal penalties – may also be justified for morally wrong actions. Whether they are will depend on whether the benefits of their employment exceed the costs. (It might sometimes maximize happiness for a society to pass laws that prohibit some type of behavior without bringing it about that its members internalize coercive social rules that prohibit it. In these instances, Lyons takes Mill's view to be that while this behavior should be illegal, it is not immoral ([1976] 1994, 57)).

Lyons observes that 'Mill's view resembles many modern rule-utilitarian theories in that the relevant rules are (as it were) "ideal", not actual. Mill's theory refers to justifiable rules, but not only to rules that are actually accepted or "in force" ([1976] 1994, 62). Still, Lyons himself refuses to call Mill a rule utilitarian.² Some more recent commentators who follow closely in Lyons's wake, however, including Wendy Donner (2009, 33–55) and myself, have not shared his reticence on this point.

3. Crisp

Crisp reads Mill as an act utilitarian largely, if not solely, on the strength of the second paragraph of the second chapter of *Utilitarianism* (II2) (1997, 95–99). It is here that Mill famously describes utilitarianism as the creed which 'holds that actions are right in proportion as they tend to promote happiness, wrong as they tend to produce the reverse of happiness' ([1861] 1969, 210). Crisp, like Brink and Turner, reads Mill as a 'sophisticated' act utilitarian, one who distinguishes between the theory's moral standard and the 'decision procedure' that it tells agents to follow in deliberating about what to do. According to Crisp, Mill recommends that we generally follow the rules of 'customary morality' when choosing between alternative courses of action, since they are reasonably good guides to choosing optimal actions; indeed, 'Crisp's Mill' believes that 'customary morality is solidly grounded on the principle of utility' (1997, 114). Nor are these rules to be mere rules of thumb to which no feelings are attached. These rules are already 'inculcated deep within us,' as Crisp puts it, which is to say that they are widely internalized in Lyons's sense, and Crisp's Mill regards this as desirable (1997, 109). (The area in which customary morality is least satisfactory from an act-utilitarian standpoint may be the paucity of the sacrifices that it demands from us for the benefit of strangers, and Crisp suggests that Mill

may be downplaying the theory's demandingness in *Utilitarianism* so as not to frighten his readers away (1997, 115)). There are two ways in which Crisp's Mill believes that explicitly utilitarian reasoning should figure in our moral thinking. First, when the rules of customary morality conflict, as they frequently will, we should set them aside and try to work out which of the actions open to us would maximize happiness (1997, 110). Second, Mill would have us periodically employ utilitarian reasoning to look for ways in which customary morality could be improved from a utilitarian standpoint, so that we can then make strategic use of 'praise and blame' to bring it about that the next generation internalizes a better moral code than has ours (1997, 131). In this way customary morality will gradually improve.

Crisp's Mill, then, holds a view much like that of R. M. Hare, according to which we should internalize rules that are distinct from the act-utilitarian injunction to maximize happiness but should periodically reflect 'in a cool hour' about how the rules that we have internalized might be improved (1981, 25–64). One difference is that Crisp's Mill apparently believes that someday humanity will have progressed to the point at which a non-utilitarian decision procedure is largely obsolete; people who are sufficiently intelligent and impartial will be able to guide their behavior through the direct application of act-utilitarian thinking: 'They will have a pretty good idea of when happiness has been maximized or not,' Crisp says, 'and will no longer need much or any of customary morality, having moved beyond it' (1997, 131). This does not mean that they will have become beings of the sort whom Hare calls 'archangels,' able to foresee perfectly all of their actions' ramifications and so invariably produce optimal consequences. But they will reliably act in ways that maximize expected happiness or utility, and hence their actions will be 'subjectively' if not always 'objectively' right from the act-utilitarian perspective.

When he turns his attention to V14, Crisp agrees with Lyons that in this passage Mill is 'analysing the notion of wrongness itself' and that on his analysis to call an action wrong means that it should be punished (1997, 128). Crisp claims that Mill would have done better to have said something quite different in V14, since he thinks that the analysis that the passage contains is vulnerable to an 'open question' argument, but he apparently finds the passage too clear to be construed differently. Crisp acknowledges that at first glance the combination of this analysis with act utilitarianism appears self-contradictory. An action might be suboptimal, and therefore wrong in act-utilitarian terms, and yet the act of punishing the agent might also be suboptimal; the act utilitarian would therefore apparently have to say that the punishment both is and is not justified and that the action both is and is not wrong. But Crisp asserts that the automatic operation of the internal sanction saves Mill's view from this obvious sort of internal inconsistency.

The principle of utility, as a practical principle, applies to human conduct. It does not govern consciences directly, since, unlike legal punishment or blaming others,

conscience is not something over which we have control. Thus there is no imaginable case of an agent's failing to maximize happiness to which Mill would be forced to retract any attribution of wrongness. (1997, 129)

Note that, unlike Lyons, Crisp does not maintain that Mill's analysis of wrongness is framed specifically in terms of the internal sanction; he says only that Mill believes that to say that an action is morally wrong means that some punishment or other is warranted for it. Guilt is simply the one form of punishment whose imposition does not require some 'punisher' to act, and so it is the one form of punishment that an act utilitarian can consistently say is warranted for every suboptimal action.

Crisp admits that Mill's conjoining of the act-utilitarian moral standard with V14's analysis of moral concepts does still have some 'peculiar implications.' These include the facts that we will never know when to feel guilty, since we can never be certain whether we have actually maximized happiness, and that we should feel guilty in cases in which we chose the action which had by far the highest *expected* happiness, even though unknown to us there was something we might have done instead whose *actual* consequences would have been better (1997, 130). Crisp suggests that the latter of these worries can at least be ameliorated if we take Mill to be thinking about subjective rather than objective wrongness in V14; as Crisp draws this distinction, an action is subjectively wrong if it fails to maximize expected happiness and objectively wrong if its actual consequences are suboptimal. Making this move probably does something to address the former worry as well.

But there is a more serious problem for Mill on Crisp's reading whose force it does nothing to blunt, which is that Crisp's attempt to free the combination of views that he attributes to Mill from inconsistency by appeal to the internal sanction fails. Recall first that as Crisp reads Mill all that it means to say that an action is wrong is that some punishment or other for it is justified. Crisp, I have already noted, does not follow Lyons in taking Mill to analyze wrongness specifically in terms of the internal sanction. Mill might be able to say without inconsistency that guilt is warranted for every suboptimal action. On Crisp's interpretation, though, Mill is still committed to saying that an action is wrong as long as *any* punishment is warranted for it. So, consider a case in which an individual did something which maximized happiness and which the individual knew would maximize happiness, but for which it would now maximize happiness to punish the individual via some external sanction. Jamie Lannister, let us say, having sworn a solemn oath to protect the reigning monarch no matter what, kills the mad king Aerys Targaryen to prevent him from incinerating the city of King's Landing and all of its inhabitants. It may well be optimal to punish Jamie, at least by informal social sanctions like using the insulting nickname 'Kingslayer.' This might, for instance, reinforce the social prohibitions against killing and oath-breaking. Punishing Jamie might still be optimal even if the circumstances of his act were generally known, as is not the case in the *Game*

of *Thrones* universe (as of this writing, anyway), since this might discourage people from trying to calculate for themselves whether their killing someone or breaking an oath would produce a positive net balance of happiness and we are notoriously bad at such utility calculations. But if the act-utilitarian principle is the appropriate standard for us to use in judging whether punishing Jamie is warranted, and if by this standard his punishment *is* warranted, then given the analysis in V14 we must conclude that he was wrong to kill King Aerys. This conclusion contradicts Mill's putative act utilitarianism, so the view that Crisp assigns to Mill is still internally inconsistent.³

It might look like Crisp could easily eliminate this inconsistency with a minor repair to his view; all that would be needed, one might think, is for him to follow Lyons in saying that Mill's analysis of wrongness in V14 is framed specifically in terms of the justifiability of the internal sanction. Yet the fundamental problem would remain. Recall that Crisp's Mill believes that the rules we should have internalized are those of our society's customary morality. Over time we can hope that customary morality gradually improves, but our internalization today of the rules that are current in our customary morality is part of what Crisp's Mill believes is 'solidly grounded' on a utilitarian foundation. This means that we ought to feel guilty about violating these rules, not about failing to maximize happiness or expected happiness. Crisp recognizes that there is an issue here; another of the peculiar implications he lists of the view that he attributes to Mill is that when we do something wrong and feel guilty about it we will usually be feeling guilty for the wrong reason. Suppose I tell a lie, for example, and this fails to maximize happiness. For an act utilitarian it is the failure to maximize that is responsible for my action's being wrong, Crisp observes, but I will feel guilty not about that but about lying as such.

But casting the issue in this way misstates it, or at least it distracts attention from a larger issue. If Crisp's Mill analyzes what it is to call an action wrong in terms of the justifiability of punishing it via the internal sanction, and if he believes that we should be punished via the internal sanction for actions that violate the rules of customary morality, then he is committed to the claim that actions contrary to those rules are wrong. This is true even when those actions do maximize happiness, as actions forbidden by customary morality sometimes will, and again this contradicts Mill's putative act utilitarianism. Even if Crisp's Mill says that optimal actions that violate the rules of customary morality are only subjectively wrong, while they are still objectively right, this would still contradict his putative act utilitarianism since Crisp says that for an act-utilitarian actions are subjectively right if they maximize expected happiness. Customary morality will not permit every action that maximizes expected happiness.⁴ Similarly, if Crisp's Mill analyzes what it is to call an action wrong in this way and if he believes (as it seems clear that he does) that we should internalize *only* the rules of customary morality, then he must believe that any actions permitted by these rules are morally permissible. Yet our customary morality permits very

many actions that maximize neither actual nor expected happiness. Yet again, this contradicts Mill's putative act utilitarianism, since act utilitarianism permits no actions that fall short of optimality.

Crisp might appeal here to his contention that Mill believes that in a genuinely civilized world, one that hopefully lies in our future, customary morality as we know it will no longer exist.⁵ Perhaps when he analyzes the notion of what it is for an action to be wrong in terms of the justifiability of punishment, Mill could mean that actions are wrong if they are the sort of actions for which the inhabitants of this genuinely civilized world would be punished by the internal sanction, even if in the world as it is the rules which people should feel guilty about violating are rather different. And perhaps in this world people would internalize only a single coercive social rule, namely one that requires them to maximize expected happiness.

I do not believe that V14 can bear this construction. Nothing in Mill's language there does anything to suggest that when he says, for instance, that a person who acted wrongly is a 'proper object of punishment,' what he really means is that she would be a proper object of punishment in a better and hopefully future world in which humanity has made great moral and mental advancements. On a straightforward reading of the passage it says instead that a person has a duty to engage in some line of conduct in the world as it is just if that conduct can 'rightfully' be 'exacted' from her in the here and now.

But there is another problem with trying to read V14 in this idealized way. Mill had a model for what people with highly cultivated natures, the sort who no longer needed anything resembling our customary morality, would be like: his eventual wife, Harriet Taylor Mill. In an essay on marriage written expressly for her consumption, Mill describes how a society of people with higher natures like hers would order their behavior.

By following their own impulses under the guidance of their own judgment, they would find more happiness, and would confer more, than by obeying any moral principles or maxims whatever: since these cannot possibly be adapted beforehand to every peculiarity of circumstance which can be taken into account by a sound and vigorous intellect *worked by a strong will*, and guided by what Carlyle calls 'an open loving heart'. Where there exists a genuine and strong desire to do that which is most for the happiness of all, general rules are merely aids to prudence, in the choice of means; not peremptory obligations. (Mill [1832–3?] 1984, 39)

While Crisp may be right to say that Mill hopes that someday it will be possible to trust people to work out how to maximize happiness without the need of a decision procedure, what this passage shows is that Mill does not conceive of such a state of society as one in which people internalize a single rule that requires them to maximize happiness. Rather they would internalize no rules at all, in Lyons's sense of feeling guilty for violating the rules and blaming others who do so. There would be no need for coercive social rules backed by sanctions of either the internal or external variety, since people would spontaneously

follow 'their own impulses under the guidance of their own judgment.' Given V14 we would expect Mill to say just what he in fact does say about such a world, namely that in it there would be no need for 'any moral principles or maxims whatever,' including the act-utilitarian principle if this is understood as a moral rule: 'morality would not exist at all as morality, since morality and inclination would agree.' So if we try to give V14 an idealized reading, one according to which Mill believes that actions are wrong here and now if they are forbidden by the rules that the inhabitants of the genuinely civilized world for which he hopes would do best to internalize, then we are forced to draw the absurd conclusion that Mill does not believe that *any* actions are wrong here and now.⁶

So the view that Crisp assigns to Mill is internally inconsistent, inasmuch as it entails contradictory evaluations of the moral standings of some actions, and its incoherence is fairly obvious. We might say on this basis that this view is inconsistent with V14. Even if we do not say this, however, we must at the very least say that considerations of charity militate strongly against embracing Crisp's reading.

4. Brink

From Crisp I will turn now to Brink. Brink too accepts much of Lyons's reading of V14, agreeing that Mill is here offering an analysis of what it means for an action to be wrong in terms of the justifiability of punishment or sanction. They differ on two points, though, only one of which Brink acknowledges. The explicit disagreement is with Lyons's claim that Mill takes sanctions to 'assume coercive rules.' Brink writes that he sees 'no evidence that Mill wants to introduce rules or principles into his formulation of the utilitarian standard' (2013, 102). The implicit disagreement concerns Lyons's assertion that Mill's analysis in V14 is framed specifically in terms of the internal sanction. Brink writes as if he and Lyons agree that according to Mill in chapter V of *Utilitarianism* an action is wrong if any sort of sanction is warranted for it.

Brink further agrees with Lyons that, in chapter V of *Utilitarianism*, Mill intends for a utilitarian criterion to be used to judge when a punishment is warranted. He therefore labels the moral theory that Mill propounds in this chapter 'sanction utilitarianism.' And 'Because sanction utilitarianism is a species of indirect utilitarianism,' he writes, 'it is inconsistent with act utilitarianism' (2013, 101). There's no reason to think that it will be optimal for agents to be punished if and only if they have performed a suboptimal action. Yet because Brink takes the rest of Mill's corpus to offer abundant evidence that he subscribes to act utilitarianism – the passages that Brink adduces include but are not limited to II2 – he concludes that there is an 'ambivalence' at the heart of Mill's understanding of morality. Even within *Utilitarianism* itself, Brink's Mill puts forward two inconsistent utilitarian moral theories.

It may not seem appropriate to number Brink among Mill's act-utilitarian interpreters, since this is only one of the theories that he attributes to Mill. But Brink is clear that he regards act utilitarianism as Mill's considered view and sanction utilitarianism as an ill-considered afterthought. Brink's claim that act utilitarianism is Mill's considered view hangs in part on his belief that sanction utilitarianism is a thoroughly unattractive theory, so that it would be uncharitable to suppose that Mill's commitment to it is deep. Brink says that sanction utilitarianism both fails to offer the advantage over act utilitarianism that it might at first glance seem to and that it suffers from some serious defects. The advantage that the theory might seem to have is that it can account for the existence not only of morally obligatory and morally wrong actions but also of actions that are morally permissible without being morally required and of supererogatory actions. An action is permissible for a sanction utilitarian if neither its performance nor its omission are optimally blamed; it is supererogatory if it is permissible and its consequences are sufficiently good (2013, 104). (For 'blame' here understand 'performing' or expressing blame, e.g. imposing at least informal external sanctions, as opposed merely to taking an attitude of blame.) Brink does not deny that sanction utilitarianism can account for these four distinct moral standings, but he asserts that act utilitarianism can do so as well, or at least that it can offer what he calls a 'pragmatic reconstruction' of this fourfold distinction (2013, 105). Specifically, the act utilitarian can distinguish between actions that it is optimal to blame agents for performing, actions that it is not optimal to blame agents for performing, actions that it is optimal to blame agents for omitting, and actions that it is not optimal to blame agents for omitting but *is* optimal to praise them for performing. Brink takes the fact that act utilitarians can assign some actions to this last category, for instance, to show that the theory can adequately account for our commonsense conviction that some actions go beyond the call of moral duty.

One of the defects that Brink finds in sanction utilitarianism is that it suffers from a version of what has come to be known as the 'wrong kind of reason' problem. More specifically, Brink alleges, sanction utilitarianism offers the wrong kind of reason for why someone should be punished. According to this theory, Brink says, an action is wrong in virtue of the justifiability of a sanction for it rather than the reverse. This, he observes, 'inverts what many would regard as the usual dependency between wrongness and sanction' (2013, 107). In other words, commonsense or customary morality is retributivist. Of course, act utilitarians reject this dependency altogether – they are the paradigmatic non-retributivists – but Brink remarks that at least the act utilitarian and retributivist 'agree on the need for an account of an action's wrongness that is independent of its suitability for sanction.'

The other defect is due to what Brink calls sanction utilitarianism's 'hybrid structure.' Sanction utilitarianism offers an indirect utilitarian standard for judging whether actions are right or wrong; they are wrong if it would be optimal

to sanction them. Yet, Brink says, 'imposing sanctions is itself a kind of action, and we can ask whether the imposition of a particular sanction would be right or wrong' (2013, 108). But this, Brink continues, implies that for this one specific kind of action the sanction utilitarian is committed to employing both a direct utilitarian standard – imposing the sanction is wrong just if imposing it would not be optimal – and an indirect utilitarian standard – imposing the sanction is wrong just if punishing someone for imposing it would be optimal. But these standards will sometimes, perhaps often, lead to different answers, so sanction utilitarianism is inconsistent.

No work is needed to show that Brink's interpretation attributes an unappealing position, or rather an unappealing combination of positions, to Mill; the claim that Mill advocates two plainly inconsistent moral theories even within one essay is pretty damning. I hasten to add here that Brink is by no means an uncharitable reader of Mill; he reaches his conclusion that Mill's utilitarianism is 'ambivalent' with reluctance, not glee. He contends, however, that the evidence that Mill subscribes to act utilitarianism provided by passages other than V14 is even more compelling than the convincing evidence that V14 provides for his sanction utilitarianism, so that this conclusion is regrettably forced upon us. And of course he takes the same evidence to support his claim that Mill is more committed to act than to sanction utilitarianism.

I agree with Brink that in V14 Mill's advocating a sanction utilitarianism that is inconsistent with act utilitarianism. I disagree about whether the available textual evidence requires us to read Mill as an ambivalent utilitarian who is most strongly committed to act utilitarianism, but challenging Brink on this point would require detailed discussion of a number of passages scattered throughout Mill's corpus; this would take me far beyond the scope of the present discussion. So my criticism of Brink's reading here will be limited to a brief rejoinder to his negative assessment of sanction utilitarianism. If I can show that Brink is too hasty in writing this view off as far less attractive than act utilitarianism then this will go some distance toward showing that he is also too hasty in concluding that act utilitarianism is Mill's considered view, even if he is right about Mill's ambivalence.

Recall that Brink both overlooks Lyons's interpretative claim that Mill intends to frame his analysis of what it means to call an action wrong in terms of the justifiability of the internal sanction specifically and challenges Lyons's claim that for Mill the justifiability of individual applications of sanctions depends on the justifiability of coercive social rules. There is an important passage from a piece of correspondence that Mill penned in 1859 that I take to provide strong evidence both that Lyons's claims are correct and that they are connected. (Note that Lyons himself doesn't cite this letter.) Mill writes that:

Now as to the still more important subject of the meaning of *ought*. ... I believe that the word has in some respects a different meaning to different people. We must first distinguish between those who have themselves a moral feeling – a

feeling of approving & condemning conscience, & those who have not, or in whom what they *may* have is dormant. I believe that those who have no *feeling* of right & wrong cannot possibly *intue* the rightness or wrongness of anything. They may assent to the proposition that a certain rule of conduct is right; but they really mean nothing except that such is the conduct which other people expect & require at their hands. ... This you will probably agree with, & I will therefore pass to the case of those who have a true moral feeling, that is, a feeling of pain in the fact of violating a certain rule, quite independently of any expected consequences to themselves. It appears to me that to them the word *ought* means, that if they act otherwise, they shall be punished by this internal, & perfectly disinterested feeling. ([1859] 1972, 649)

In this letter, Mill is even more explicit than he is in V14 about offering an analysis that conceptually ties wrongness to the operation of the conscience. He is also explicit here that the conscience operates by enforcing internalized rules; the experience of guilt is an instance of 'a feeling of pain in the fact of violating a certain rule.' Admittedly, the specific account that Mill offers here is not precisely the analysis that he offers in V14; it explicates what it means to call an action wrong in terms of the speaker's expectations about when she *would* experience guilt rather than in terms of her beliefs about when people *should* experience it. Certainly the account found in *Utilitarianism* is the more plausible of the two, but given that the letter was composed during the period in which Mill was drafting *Utilitarianism*; it is reasonable to read the essay in light of it and take the former to contain a revised version of what is found in the latter. So the specific version of sanction utilitarian that Mill puts forward in chapter V of *Utilitarianism* might accurately, if not elegantly, be called 'rule-governed internal-sanction utilitarianism.'

Once we recognize that Mill understands what it is for an action to be wrong specifically in terms of the justifiability of the internal sanction, Brink's hybrid structure objection to Mill's sanction utilitarianism no longer goes through. The objection depends on the claim that, as Brink's puts it, 'imposing sanctions is itself a kind of action.' But, as Crisp recognizes, the internal sanction is different from legal or informal social sanctions in that its imposition does not require an action. So this objection cuts no ice against Mill's version of sanction utilitarianism.

This leaves Brink's 'wrong kind of reason' objection, to which I shall make two replies. First, the sort of sanction utilitarianism that Lyons finds in Mill is in fact not obviously vulnerable to the wrong kind of reason objection that Brink advances. The view that Lyons takes Mill to hold does not say that a particular action is wrong just if it would be optimal for the agent to feel guilty about it; it does not treat the consequences of her experiencing guilt as the right kind of reason for judging her to have acted wrongly. Utilitarian considerations only enter into the justification of the coercive social rules which distinguish right from wrong actions. When it comes to judgments about particular actions and agents, the view regards the wrongness of an action – its contravention of a

justified rule – as the right reason to sanction the agent, so the usual dependency is maintained. The criminal law provides a useful analogy here. The commonsense view is that the fact that the consequences of sending an individual to jail would on balance be desirable is the wrong kind of reason for doing so. The right kind of reason is that she has been found guilty of violating a criminal law for which the penalty is incarceration.⁷ But there is no comparable commonsense consensus that legislators should not rely on cost-benefit analyses in deciding what behaviors to criminalize. Most people, I think, would say that the consequences of proposed laws are at least among the right kinds of reasons for legislators to attend to, and some – perhaps even many – might say that they are the only relevant considerations.⁸

Second, there *is* a moral theory that is obviously at odds with commonsense morality's views about the right kinds of reasons for punishment, and it is act utilitarianism. So given Brink's assumption that conventional morality correctly captures the *right* kind of reason for punishment, act utilitarianism must justify punishment by appeal to a reason of the wrong sort. Nothing could make this clearer than Brink's own pragmatic reconstruction of the familiar fourfold distinction between moral standings found in commonsense morality. Although Brink assumes that it will be 'typically' be optimal to blame agents only for actions that were egregiously suboptimal, and so wrong in act-utilitarian terms, he does not and cannot assert that this will always be the case (2013, 105). (He acknowledges that some actions that it will be optimal to praise will actually be actions that were suboptimal, i.e. wrong in act utilitarian terms.) Brink's proposed reconstruction thus highlights the fact that for act utilitarians the answer to the question of whether we should blame (i.e. punish) someone is only contingently connected to the moral standing of whatever it is that we would be blaming her for. But this flies in the face of, if it does not invert, 'what many would regard as the usual dependency between wrongness and sanction,' since it denies that the justifiability of sanctions depends upon the wrongness of the conduct. For the act utilitarian, the fact that the consequences of, for instance, sending a person to jail would on balance be desirable is precisely the *right* kind of reason to incarcerate her, which is why one longstanding objection to act utilitarianism is that it can require punishing the innocent. Sanction utilitarianism seems if anything to be closer to commonsense morality in this regard than is act utilitarianism; at least sanction utilitarianism and commonsense morality agree there is *some* conceptual connection between an action's suitability for sanction and its wrongness. Of course, an act utilitarian will rightly say that simply to assume that conventional morality accurately captures the right kind of reason for punishment begs the question against her theory. I do not mean to endorse this assumption. My point is only that when Brink endorses it, in the course of arguing that Mill would do better to reject sanction utilitarianism and become an act utilitarian *simpliciter*, he weakens rather than strengthens his own case.⁹

So even if we accept that Brink's pragmatic reconstruction suffices to show that act utilitarianism can account for our familiar variety of moral standings as well as sanction utilitarianism can, making the theories equally attractive in that regard, it remains the case that neither of Brink's objections to sanction utilitarianism gives us any reason to reject the theory. One of the objections – that the theory fails to honor commonsense morality's notions about what counts as the right kind of reason for punishing someone – seems to have more force against act than (the relevant version of) sanction utilitarianism, if it has any force at all. So it appears that if we must regard Mill as an ambivalent utilitarian then it would be kinder to his reputation for us to take his considered view to be sanction utilitarianism, if the text could possibly permit this. And if Brink should turn out to be so far wrong about the text that it is possible for us to regard Mill as a thoroughgoing sanction utilitarian, then this would be kinder still.

5. Turner

In company with Crisp, Turner reads Mill as a consistent or 'unambivalent' act utilitarian. In company with both Crisp and Brink, Turner maintains that Mill's act utilitarianism is of the sophisticated variety that distinguishes between the act-utilitarian moral standard and the decision procedure that theory recommends to agents. But Turner's approach to V14 is novel.

According to Turner, Mill assigns different meanings at different times to the word 'morality' itself and to moral terms such as 'right' and 'wrong,' and it is only from context that we can judge which set of meanings to attach to the language in a given passage. Sometimes when Mill talks about morality he is talking about what is ultimately or fundamentally choiceworthy, by which I take Turner to mean what we have the most reason to do all things considered (2015, 723–724). When Mill is using moral language in this sense, right actions are choiceworthy and wrong ones are not. At other times, though, Mill uses the word 'morality' to refer to blameworthiness, and when he is speaking in this way, right actions are not blameworthy and wrong ones are. So rather than depicting Mill's views as substantively inconsistent, as does Brink, Turner maintains that apparent substantive inconsistencies can be explained away as inconsistencies in Mill's manner of speaking. For Turner Mill is not an ambivalent utilitarian so much as an ambiguous one, we might say, or at least one who does not explicitly disambiguate key terms. That is a less damning charge against Mill as a thinker; indeed, it may not be damning at all as long as there is some principle that guides Mill's use of terminology and he avoids equivocating.

On Turner's reading, act utilitarianism is Mill's theory of morality in the sense of choiceworthiness, and Mill's using moral language in this sense in passages that lend themselves to an act-utilitarian interpretation such as II2. In at least some passages that do not lend themselves to an act-utilitarian reading, however, Turner contends that Mill has shifted his meaning, so that he is actually

referring to blameworthiness rather than choiceworthiness (2015, 737–741). This is true of V14 in particular, and also of Mill's discussion in the *System of Logic* of the Art of Life's 'department' of 'Morality' ([1843] 1974, 949–950). So Turner agrees with Lyons inasmuch as he takes V14 to offer an analysis of one sense of 'morally wrong action.'

But how are these two senses of 'morality' related? According to Turner, morality in the sense of blameworthiness 'is best understood as evaluating actions in terms of whether a person has failed to do what is reasonably expected of him given epistemic and other limitations, and this evaluation accounts for the appropriateness of holding that person accountable' (2015, 740). An action for which an individual is blameworthy is one 'with regard to which the individual would have no right to complain if he were punished for performing it, because he had failed to do what others would reasonably expect him to do. It is in this sense that he is blameworthy.' Turner acknowledges that Mill believes that for beings like us following the best available decision procedure will largely be a matter of complying with a set of rules, although he denies that this will always be the case; according to Mill, on Turner's reading, 'an agent may be blameworthy for *not* violating a rule,' any rule, when she could reasonably be expected to know that doing so would yield a better outcome (2015, 741). Even when it comes to self-blame, via the conscience, Turner doubts that the violation of some ideal rule is necessary and sufficient for punishment to be warranted. This doubt stems at least in part from doubt about whether Mill regards the operation of the conscience to be entirely governed by internalized rules, the letter to Ward notwithstanding; Turner suggests that 'Mill takes conscience to be sensitive enough to track obvious exceptions to general rules' (2015, 743).

Turner's reading of Mill is insightful and inventive, but it is not unproblematic. Note first that on straightforward textual grounds it may not be consistent with V14. It is essential to Turner's reading that Mill sometimes talks about actions being morally wrong without this connoting that punishing the agent would be justified, namely when he uses 'wrong' to mean not choiceworthy. But in V14, Mill says explicitly that there is no such sense of 'morally wrong': 'the idea of penal sanction ... enters ... into the conception ... of *any* kind of wrong.' Admittedly, this point may depend on a too-literal reading of this line, in which Mill is noting that the conceptual connection between wrongness and the justifiability of punishment does not hold only in the case of injustices. So I shall not put too much weight on it. But there is a second problem with Turner's interpretation that I take to be far more damaging.

This is that on Turner's reading Mill's view contains an especially glaring and egregious example of the wrong kind of reason problem that I discussed in the previous section. Suppose the question arises of whether to punish someone by an external sanction. As Turner reads him, Mill must say that this course of action is choiceworthy – it is what we have the most reason to do, all things considered – just if it would be optimal. Whether the individual in question did

anything blameworthy – that is, whether she failed to do her duty in virtue of failing to live up to others' reasonable expectations of her – is irrelevant. Not only is its being 'appropriate' to punish her not a necessary condition of her punishment's being warranted, it is not even germane to the question.¹⁰ It is hard to know what to make of Turner's use of 'appropriate,' divorced as it is from any connection to our reasons for action. For Turner's Mill, after all, the only considerations that count as reasons in favor of punishing someone are the desirable consequences that this would have, if any; the only considerations that count as reasons against are the undesirable consequences. If you are a thoroughgoing act utilitarian, one who says that punishment is warranted just if it will be optimal, then when people protest that this is the wrong kind of reason for punishment at least you can accuse them of begging the question against you, as already noted. But if you adopt an account of when punishment is warranted according to which this depends on what the agent did in the past, then the fact that punishing her would produce better future consequences is the wrong kind of reason for punishment by your own lights.

We should demand *very* strong textual evidence before concluding that this is Mill's position. And in fact the text, or at least V14 itself, seems to tell in the other direction; the clear implication of V14 seems to be that an individual who is not blameworthy *would* be entitled to complain about being subjected to punishment. It would be quite surprising for Mill to believe that punishing her over her justified complaint could nonetheless be choiceworthy – and stranger still for him to believe that the fact that her complaint is justified does not even count against its choiceworthiness.

6. Conclusion

Clearly my conclusion is that none of Mill's act-utilitarian interpreters have as yet found a compelling way of coming to terms with *Utilitarianism* V14. As I conceded at my commencement, this alone is not enough to let anyone draw the further conclusion that the act-utilitarian interpretation ought to be rejected. Despite its importance, V14 is just one of many stretches of text that bear on the debate how we should answer the question of what moral theory Mill holds. Apart from II2 I have said nothing about the passages that may seem to favor the act-utilitarian interpretation, and there are several. The lines which do this most clearly might be these from 'Taylor's Statesman':

The evil of departing from a well-known and salutary rule is indeed one momentous item on that side of the account; but to treat it as equal to infinity, and as necessarily superseding the measurement of any finite quantities of evil on the opposite side, appears to us to be the most fatal of all mistakes in ethical theory. (Mill and George [1837] 1977, 638)

There is also a letter to John Venn, in which Mill writes that 'I agree with you that the right way of testing actions by their consequences, is to test them by

the natural consequences of the particular action' (Mill [1872] 1972, 1881). And there are more.

Now of course those who dislike the act-utilitarian reading have ways of explaining away these passages. II2, for instance, seems if anything to express a scalar view that is rather different from what is today considered standard act utilitarianism (cf. Crisp 1997, 96–97). 'Taylor's Statesman' is a very early ([1837] 1977) essay co-authored with George Grote, one that even the editor of Mill's *Collected Works* relegates to an appendix, so the line that I quoted from it may not represent Mill's position, let alone his considered and mature view. The Venn letter may be more relevant to Mill's views on prudential than moral evaluation.¹¹ But we kid ourselves if we are entirely satisfied with these explanations. As I said at the outset, there is little prospect of our ever finding a moral theory that is both very plausible in its own right and a perfect fit with all of Mill's ethical writings. But while there may be no reading of Mill's moral theory on which we can rest in complete comfort, V14 remains an especially sharp thorn in the side of those who settle on the act-utilitarian interpretation.

Notes

1. Jacobson (2003) argues that *Utilitarianism* is intended as an ecumenical statement of doctrine that is common ground between utilitarians rather than a statement of Mill's views specifically. He also (2008) argues for a novel interpretation of Mill's utilitarianism according to which Mill does not believe that 'purely self-regarding' actions can be wrong, for reasons that are independent of any utility calculations, and V14 is an essential part of his argument. I do not mean to insinuate that there is any inconsistency between these articles. The earlier piece focuses on the first four chapters of *Utilitarianism*, but since we know that chapter V was composed apart from the rest of *Utilitarianism* it is entirely possible that it was written with a different purpose in view (see Robson 1969, cxxiii–cxxv).
2. Lyons gives two reasons for declining to call Mill a rule utilitarian, neither of which is persuasive. One is that Mill takes the disutility of sanctions into account in determining which coercive social rules are justified; Lyons says that most rule utilitarians don't take sanctions into account, but he never gives any reason why they could not, and indeed he acknowledges that at least one – Richard Brandt – does (1977, 122). The other is that Mill arrives at his view through, in part, 'considering the logic of the central moral concepts' ([1978] 1994, 79). Again, Lyons gives no reason why this should disqualify Mill from being a rule utilitarian. There may be instances in which it does make sense to take the patterns of reasoning that would lead someone to endorse moral theories into account when classifying them, rather than attending only to their standard for judging actions right or wrong. This would let us say that distinct theories are extensionally equivalent, as Lyons famously shows is true of 'primitive' rule utilitarianism and act utilitarianism (Lyons 1965, 119–135). But why Lyons should believe that the fact that Mill's reasoning includes some analysis of moral concepts should differentiate his moral theory from rule utilitarianism is mysterious. (See also Gaus 1980, 268).

3. Of course, this would not hold true if Crisp took Mill to be saying only that part of what we mean when we 'call anything wrong' is that it is justifiably punished. In other words, my objection assumes that in V14 Crisp takes Mill to be asserting the biconditional 'An action is wrong if and only if it is justifiably punished.' If instead he takes him to assert only the conditional statement 'If an action is wrong then it is justifiably punished,' then my objection is inapplicable. The fact that Mill says in V14 that when we call an action wrong we 'mean to imply' that the agent should be punished might even seem to lend support to this weaker reading, since it might suggest that Mill does not take himself to be giving a full account of the meaning of 'morally wrong action.' However, there is nothing in Crisp's text to suggest that he intends to claim that Mill asserts only the conditional statement in V14 and not the biconditional. On the contrary, he writes that Mill's 'phrase "mean to imply" here is best taken as equivalent to "mean", since in 5.14 he explicitly says he is analysing the 'notions' of right and wrong ... According to Mill, when one says that an action is "wrong" one means that a ought to be punished by law, opinion or conscience' (Crisp 1997, 128). Moreover, it would be difficult to reconcile this way of reading V14 with the text, Mill's slack use of 'mean to imply' notwithstanding, since at the end of the paragraph it becomes clear that he means to assert the biconditional: 'we call any conduct wrong, or employ, instead, some other term of dislike or disparagement, according as we think that the person ought, or ought not, to be punished for it.' I am indebted to Roger Crisp and Ben Eggleston for discussion on this point.
4. We should not forget that Crisp's Mill says that even now we should follow the act-utilitarian principle when the rules of customary morality conflict. But there is no reason to expect that every time a rule of customary morality forbids a maximizing action it will be in conflict with another rule of the same kind.
5. I allude here to Mill's use of 'civilization' in 1861, 231–233.
6. It is also worth noting that a society in which the only rule that is widely internalized is the act-utilitarian principle is probably not one for which a utilitarian would wish. First, such a society might face the following unpleasant choice: either use draconian (i.e., severely painful) measures to make people experience compunction and blame strongly enough for them to be motivated to make the sacrifices that this principle requires or allow them to constantly experience some level of guilt for failing to live up to the principle. Second, it might be difficult for people who have internalized only this rule to coordinate their behavior, and so many benefits of social cooperation might be lost. How, for example, can I make plans based on your promise to assist me, if I know that when the time comes you will feel obligated to break your promise should this seem to you to be happiness-maximizing? (See Hodgson 1967, 38–50.) Of course, this second point may also raise a problem for Mill's depiction of a 'post-moral' society in the 'Marriage' essay.
7. We might think it is permissible for a judge not to jail a person who has broken such a law if the law permits some alternative punishment whose consequences would be better, of course.
8. It is possible that the type of sanction utilitarianism that Lyons takes Mill to hold is open to some other, subtler wrong kind of reason objection. Stephen Darwall, for example, has argued that this is true of indirect rule-consequentialist theories, and he would certainly include the view that Lyons attributes to Mill within the scope of the objection. Note, though, that Darwall maintains that direct act-consequentialist theories are equally vulnerable to the same objection (2006, 103–104).

9. An anonymous reviewer rightly notes that the view that Crisp attributes to Mill may also be vulnerable to a version of the wrong kind of reasons problem, although space did not permit me to explore this possibility.
10. It is possible that the type of sanction utilitarianism that Lyons takes Mill to hold is open to some other, subtler wrong kind of reason objection. Stephen Darwall, for example, has argued that this is true of indirect rule-consequentialist theories, and he would certainly include the view that Lyons attributes to Mill within the scope of the objection. Note, though, that Darwall maintains that direct act-consequentialist theories are equally vulnerable to the same objection (2006, 103–104).
11. Or so Ben Eggleston and I have argued (2008).

Acknowledgments

I am grateful for the valuable comments that I have received from David Brink, Roger Crisp, Ben Eggleston, Daniel Jacobson, Piers Norris Turner, an anonymous reviewer for this journal, and an audience at the 2016 meeting of the International Society for Utilitarian Studies in Lille, France.

Notes on contributor

Dale E. Miller is Professor of Philosophy at Old Dominion University. His other work on Mill includes his book *J. S. Mill: Moral, Social and Political Thought* (Cambridge: Polity, 2010), *John Stuart Mill and the Art of Life* (co-edited with Ben Eggleston) (New York: Oxford, 2011), and *A Companion to Mill* (co-edited with Christopher Macleod) (Malden: Wiley-Blackwell, 2017).

ORCID

Dale E. Miller  <http://orcid.org/0000-0002-8946-0040>

References

- Brink, David. 2013. *Mill's Progressive Principles*. Oxford: Clarendon Press.
- Crisp, Roger. 1997. *Routledge Philosophy Guidebook to Mill on Utilitarianism*. London: Routledge.
- Darwall, Stephen. 2006. *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge: Harvard University Press.
- Donner, Wendy. 2009. "Mill's Moral and Political Philosophy." In *Mill*, 13–143. Malden: Wiley-Blackwell.
- Eggleston, Ben, and Dale E. Miller. 2008. "Mill's Misleading Moral Mathematics." *Southwest Philosophy Review* 24 (1): 153–161.
- Gaus, Gerald. 1980. "Mill's Theory of Moral Rules." *Australasian Journal of Philosophy* 58 (3): 265–279.
- Hare, R. M. 1981. *Moral Thinking: Its Levels, Method, and Point*. Oxford: Oxford University Press.

- Hodgson, D. H. 1967. *Consequences of Utilitarianism: A Study in Normative Ethics and Legal Theory*. Oxford: Clarendon Press.
- Jacobson, Daniel. 2003. "J. S. Mill and the Diversity of Utilitarianism." *Philosophers' Imprint* 3 (2): 1–18.
- Jacobson, Daniel. 2008. "Utilitarianism without Consequentialism: The Case of John Stuart Mill." *Philosophical Review* 117 (2): 159–191.
- Lyons, David. 1965. *Forms and Limits of Utilitarianism*. Oxford: Oxford University Press.
- Lyons, David. 1977. "Human Rights and the General Welfare." *Philosophy and Public Affairs* 6 (2): 113–129.
- Lyons, David. [1976] 1994. "Mill's Theory of Morality." Reprinted in *Rights, Welfare, and Mill's Moral Theory*, 47–65. Oxford: Oxford University Press.
- Lyons, David. [1978] 1994. "Mill's Theory of Justice." Reprinted in *Rights, Welfare, and Mill's Moral Theory*, 67–88. Oxford: Oxford University Press.
- Lyons, David. [1982] 1994. "Benevolence and Justice in Mill." Reprinted in *Rights, Welfare, and Mill's Moral Theory*, 109–146. Oxford: Oxford University Press.
- Mill, J. S. [1832–3?] 1984. "On Marriage." In *Collected Works of John Stuart Mill*, edited by J. M. Robson, vol. XXI, 35–49. Toronto: University of Toronto Press.
- Mill, J. S. [1833] 1969. "Remarks on Bentham's Philosophy." In *Collected Works of John Stuart Mill*, edited by J. M. Robson, vol. X, 3–18. Toronto: University of Toronto Press.
- Mill, J. S. [1843] 1974. In *A System of Logic. Collected Works of John Stuart Mill*, edited by J. M. Robson, vol. VII–VIII. Toronto: University of Toronto Press.
- Mill, J. S. [1859] 1972. "Letter to William George Ward." In *Collected Works of John Stuart Mill*, edited by J. M. Robson, vol. XV, 646–650. Toronto: University of Toronto Press.
- Mill, J. S. [1861] 1969. In *Utilitarianism. Collected Works of John Stuart Mill*, edited by J. M. Robson, vol. X, 203–259. Toronto: University of Toronto Press.
- Mill, J. S. [1872] 1972. "Letter to John Venn." In *Collected Works of John Stuart Mill*, edited by J. M. Robson, vol. XVII, 1881–1882. Toronto: University of Toronto Press.
- Mill, J. S. and George Grote. [1837] 1977. "Taylor's Statesman." In *Collected Works of John Stuart Mill*, edited by J. M. Robson, vol. XIX, 617–647. Toronto: University of Toronto Press.
- Miller, Dale E. 2010. *John Stuart Mill: Moral, Social and Political Thought*. Cambridge: Polity.
- Robson, John M. 1969. "Textual Introduction." In *Collected Works of John Stuart Mill*, edited by J. M. Robson, vol. X, cxv–cxxxix. Toronto: University of Toronto Press.
- Turner, Piers Norris. 2015. "Rules and Right in Mill." *Journal of the History of Philosophy* 53 (4): 723–745.